



UNIVERSIDADE FEDERAL DA BAHIA – UFBA
DEPARTAMENTO DE ENGENHARIA ELÉTRICA E DE
COMPUTAÇÃO - DEEC
PROGRAMA DE PÓS-GRADUAÇÃO EM ENGENHARIA
ELÉTRICA - PPGE



ESCOLA POLITÉCNICA

CONJUNTO DE CLASSIFICADORES NEURAIIS ALIMENTADOS POR
INFORMAÇÃO ESPECIALISTA PARA FILTRAGEM *ONLINE* EM
UM DETECTOR DE PARTÍCULAS

Edmar Egidio Purcino de Souza

Tese de Doutorado apresentada ao Programa
de Pós-graduação em Engenharia Elétrica,
PPGEE, da Universidade Federal da Bahia,
como parte dos requisitos necessários à
obtenção do título de Doutor em Engenharia
Elétrica.

Orientadores: Eduardo F. de Simas Filho

Paulo C. M. de Abreu Farias

TESE DE DOUTORADO

Salvador
Outubro de 2021

de Souza, Edmar Egidio Purcino

Conjunto de Classificadores Neurais Alimentados por Informação Especialista para Filtragem *Online* em um Detector de Partículas/Edmar Egidio Purcino de Souza.
– Salvador: UFBA/PPGEE, 2021.

XIX, 137 p.: il.; 29, 7cm.

Orientadores: Eduardo F. de Simas Filho

Paulo C. M. de Abreu Farias

Tese (doutorado) – UFBA/PPGEE/Programa de Engenharia Elétrica, 2021.

Referências Bibliográficas: p. 113 – 124.

I. de Simas Filho, Eduardo F. *et al.*. II. Universidade Federal da Bahia, PPGEE, Programa de Engenharia Elétrica. III. Título.

CONJUNTO DE CLASSIFICADORES NEURAIS ALIMENTADOS POR
INFORMAÇÃO ESPECIALISTA PARA FILTRAGEM *ONLINE* EM UM
DETECTOR DE PARTÍCULAS.

Edmar Egidio Purcino de Souza

TESE SUBMETIDA AO CORPO DOCENTE DO PROGRAMA DE PÓS-GRADUAÇÃO EM ENGENHARIA ELÉTRICA DA UNIVERSIDADE FEDERAL DA BAHIA (PPGEE / UFBA) COMO PARTE DOS REQUISITOS NECESSÁRIOS PARA A OBTENÇÃO DO GRAU DE DOUTOR EM ENGENHARIA ELÉTRICA.

Examinada por:



Prof. Eduardo Furtado de Simas Filho, D.Sc.
Orientador (PPGEE/UFBA)



Prof. Paulo C. M. de Abreu Farias, D.Sc.
Coorientador (PPGEE/UFBA)



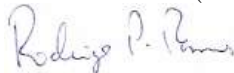
Prof. Antônio Carlos Lopes Fernandes Júnior, D.Sc.
Avaliador Interno (PPGEE/UFBA)



Prof. Anderson Dourado Sisnando, D. Eng.
Avaliador Externo (UFRB)



Prof. Romis Ribeiro de Faissol Attux, D.Sc.
Avaliador Externo (UNICAMP)



Prof. Rodrigo Pereira Ramos, D.Sc.
Avaliador Externo (UNIVASF)

SALVADOR, BA - BRASIL
OUTUBRO DE 2021

Agradecimentos

Agradeço aos meus pais Edvaldo e Magnólia e à minha irmã Mayara, por todo incentivo dado na persistência em continuar desde os primeiros anos de escola. Meus pais, minha família, sem vocês eu não teria chegado em lugar algum.

Agradeço a Laiane, minha esposa, pelo amor, carinho, paciência e compreensão. Com suas palavras doces (e as severas também), eu recuperei forças pra continuar. Obrigado pelo que fez por mim, durante todos esses anos. Se existe alguém que sabe como foi esse caminho, com certeza é você.

Agradeço aos meus orientadores Prof. Eduardo Simas e Prof. Paulo Farias, e ao Prof. Seixas, pelas discussões, conversas, ideias, motivações e inúmeras correções/ajustes dos trabalhos desenvolvidos desde o mestrado. Em 2012, na turma da disciplina Processamento Estatístico de Sinais do PPGEE/UFBA, ministrada pelo Prof. Simas, eu pude dar os primeiros passos que me levaram a contribuir para um dos maiores e mais complexos experimentos científicos já construídos. Prova de que a educação com a pesquisa científica pode nos levar a caminhos impensáveis.

Aos amigos do time do Ringer, Juan, João, Micael e Gabriel, pelo trabalho em equipe, discussões, implementações e suporte durante o desenvolvimento deste trabalho. Valeu time!! Obrigado por recepcionar tão bem no LPS/COPPE e a inestimável ajuda. Aos amigos do Laboratório de Sistemas Digitais da UFBA, pelo convívio, debates técnicos e as doses intermináveis de café para suportar o desafio. Aos amigos do CHC, Adjan, Diego, Julio, Luidson e Tiago, pela amizade de mais de 20 anos, gargalhadas e apoio para manter a sanidade mental. Aos amigos da Área1, Emílio, Laion, Lucas, Murilo e Vandrê pela amizade e motivação.

Agradeço a colaboração ATLAS, pela disponibilização dos dados e de toda a infraestrutura para análises e desenvolvimentos. Agradeço ao Prof. Fernando Gabriel Monticelli (UNLP/AR) pela paciência, orientação e condução do meu processo de qualificação para afiliação na colaboração. Agradeço também ao Prof. Chris Meyer (Indiana/EUA) pelas contribuições no processo de qualificação e toda a confiança depositada nos desenvolvimentos em qualidade de dados para o *TrigEgamma*.

À mãe natureza.

Agradeço ao Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) e a FAPESB, pelo suporte financeiro.

Resumo da Tese apresentada à PPGE/UFBA como parte dos requisitos necessários para a obtenção do grau de Doutor em Engenharia Elétrica (D. Eng.)

CONJUNTO DE CLASSIFICADORES NEURAI ALIMENTADOS POR
INFORMAÇÃO ESPECIALISTA PARA FILTRAGEM *ONLINE* EM UM
DETECTOR DE PARTÍCULAS

Edmar Egidio Purcino de Souza

Outubro/2021

Orientadores: Eduardo F. de Simas Filho
Paulo C. M. de Abreu Farias

O ATLAS é um dos principais experimentos do LHC, tendo como objetivo investigar os constituintes fundamentais da matéria e suas formas de interação. No LHC, partículas são colididas a cada 25 ns, podendo alcançar uma energia de até 14 TeV, e durante as colisões, é gerado um grande volume de dados (≈ 70 TB/s). Elétrons são partículas de interesse do experimento, porém são mascarados por um intenso ruído de fundo composto por jatos hadrônicos, que podem apresentar semelhante perfil de deposição no calorímetro (medidor de energia altamente segmentado). Para lidar com o grande volume de informação, o ATLAS utiliza um sistema de seleção *online* de eventos, para remover a parcela de informação não relevante e preservar as assinaturas de interesse. O NeuralRinger é o método padrão do experimento para classificação de elétrons na etapa rápida de calorimetria, que descreve o perfil de deposição de energia das partículas por meio de anéis concêntricos, gerados em torno da célula mais energética. Os sinais em anéis alimentam um conjunto de classificadores neurais para realizar a decisão elétron/jato. Uma informação usualmente utilizada em calorimetria para discriminação entre elétrons e jatos é o espalhamento lateral do chuveiro. Devido aos diferentes tipos de interação dessas partículas com a matéria, jatos apresentam perfil mais largo. Visando a melhoria de desempenho do NeuralRinger, neste trabalho é proposta uma etapa de pré-processamento especialista, que destaca as diferenças do perfil lateral do chuveiro, facilitando a discriminação. Métodos de treinamento de redes neurais foram desenvolvidos para ajustar de modo integrado os pesos da rede neural e os coeficientes da etapa de pré-processamento. Os métodos propostos foram avaliados utilizando dados experimentais do ATLAS. Através da abordagem proposta, foi possível alcançar desempenho superior ao NeuralRinger, reduzindo a aceitação de falsos elétrons.

Abstract of Thesis presented to PPGEE/UFBA as a partial fulfillment of the requirements for the degree of Doctor of Electrical Engineering (D. Eng.)

ENSEMBLE NEURAL CLASSIFIERS FED BY EXPERT INFORMATION FOR ONLINE FILTERING IN A PARTICLE DETECTOR

Edmar Egidio Purcino de Souza

October/2021

Advisors: Eduardo F. de Simas Filho

Paulo C. M. de Abreu Farias

ATLAS is one of the main experiments of the Large Hadron Collider (LHC), which aims to investigate the building blocks of matter and their forms of interaction. At the LHC, particles are collided every 25 ns, reaching an energy of up to 14 TeV, and during collisions, a large amount of data is generated (≈ 70 TB/s). Electrons are particles of interest to the experiment, however they are masked by an intense background noise composed of hadronic jets, which may have a similar deposition profile in the calorimeter (highly segmented energy meter). To handle the large volume of information, ATLAS uses an online event selection system, to remove the non-relevant information and preserve the interesting signatures. The NeuralRinger is the standard method for electron classification in the fast calorimetry step, which describes the energy deposition profile of particles with concentric rings, generated around the most energetic cell. The ring-shaped signals feed an ensemble neural classifiers to make the electron/jet decision. An information commonly used in calorimetry, to electrons/jets discrimination is the lateral shower width. Due to the different interaction types of these particles with matter, jets have a wider profile. Aiming at improving the NeuralRinger performance, in this work an expert pre-processing step is proposed, which highlights the differences in the lateral shower profile of rings-shaped signals, facilitating the discrimination. Neural network training methods were developed to tuning the neural network weights and the preprocessing step coefficients in an integrated way. The proposed methods were evaluated using experimental data from ATLAS. Through the proposed approach, it was possible to achieve superior performance to the NeuralRinger, reducing the acceptance of false electrons.

Sumário

Lista de Figuras	ix
Lista de Tabelas	xiv
1 Introdução	1
1.1 Motivação	2
1.2 Objetivo	5
1.3 Principais Contribuições desta Tese	5
1.4 Organização do Documento	7
2 Física de Partículas e o Experimento ATLAS	8
2.1 O Modelo Padrão da Física de Partículas	10
2.2 Grande Colisor de Hádrons	13
2.3 O Experimento ATLAS	16
2.3.1 Breve Introdução à Calorimetria	18
2.3.2 Características Gerais do Experimento ATLAS	20
2.4 Seleção <i>Online</i> de Elétrons no ATLAS	27
2.4.1 Primeiro Nível de Filtragem	28
2.4.2 Segundo Nível de Filtragem	29
2.5 Algoritmo <i>NeuralRinger</i>	33
2.5.1 Breve Histórico de Contribuições ao <i>NeuralRinger</i>	35
2.6 Monitoração e Qualidade de Dados no Sistema de Filtragem <i>Online</i>	37
3 Sistemas Inteligentes de Classificação	42
3.1 Classificação de Sinais	42
3.1.1 Problemas de Detecção Binária	43
3.2 Redes Neurais Artificiais	48
3.2.1 Treinamento de Redes Neurais com Algoritmos Bio-Inspirados	52
3.3 Aplicações de Aprendizado de Máquina em Problemas de Física de Altas Energias	58

4 Metodologia Proposta e Conjuntos de Dados	61
4.1 Métodos de Treinamento Propostos	62
4.1.1 Treinamento usando Algoritmos Bio-Inspirados	63
4.1.2 Treinamento usando <i>Backpropagation</i>	71
4.1.3 Procedimento de Validação Cruzada	72
4.2 Descrição dos Conjuntos de Dados Utilizados	74
4.3 Organização para Exibição dos Resultados	77
4.4 Parâmetros de Avaliação de Desempenho	78
5 Resultados	81
5.1 Comparação de Desempenho entre os Métodos de Treinamento	82
5.2 Ajuste da Função de Saída do Neurônio pelo Nível de Empilhamento	102
5.3 Comentários e Discussão	105
6 Conclusões e Trabalhos Futuros	108
Referências Bibliográficas	113
A Produção Científica	125
A.1 Publicações em Periódicos	125
A.2 Artigos Aceitos em Periódicos	127
A.3 Artigos em Conferências	129
A.4 Resumos em Conferências	132
B Qualificação para Afiliação no Experimento ATLAS	134
B.1 Descrição da Tarefa de Qualificação	134
B.2 Monitoração <i>Online</i> nos Algoritmos de Reconstrução	135

Lista de Figuras

2.1	Representação gráfica do modelo padrão das partículas e interações fundamentais da matéria. Extraído de [42].	11
2.2	Representação gráfica da instalação subterrânea do LHC no CERN, e alguns de seus experimentos localizados em pontos de colisão de feixes. Extraído de [8].	14
2.3	Planejamento em linha do tempo para a operação do LHC com seu aumento gradual da energia do centro de massa de suas colisões e os prazos reservados em janelas de parada para atualizações dos experimentos (<i>Long Shutdown</i>). Extraído de [44].	15
2.4	Complexo de etapas de aceleração de feixes de prótons no CERN, que compõem o acelerador LHC. Extraído de [45].	16
2.5	Representação gráfica do experimento ATLAS e seus subsistemas. Extraído de [7].	17
2.6	Sistema de coordenadas empregado pelo ATLAS. Extraído de [46].	18
2.7	Esquemático da composição do sistema de calorimetria do ATLAS. Extraído de [7].	21
2.8	Corte transversal do ATLAS, mostrando os subsistemas e as interações ideais de algumas partículas. Extraído de [48].	22
2.9	Representação gráfica da granularidade das camadas do barril do ECAL em $ \eta = 0$ e um esboço da geometria de acordeão. Extraído de [49].	23
2.10	Estrutura de acordeão do barril eletromagnético. A figura na parte superior é uma vista de um pequeno setor do calorímetro em um plano transversal aos feixes do LHC. Extraído de [49].	24
2.11	Representação gráfica de um módulo do TileCal. Extraído de [52].	26
2.12	Sistema ATLAS TDAQ na <i>Run2</i> mostrando os componentes relevantes para o acionamento, bem como a leitura do detector e o fluxo de dados. Extraído de [54].	29

2.13	Esboço de um conjunto de células do calorímetro que o algoritmo de janela deslizante aplicado no <i>L1Calo</i> para realizar a seleção de elétrons. A região central em verde representa a <i>RoI</i> , a região de isolamento EM é marcada em amarelo e a região central e de isolamento HAD é marcada em roxo. O valor de energia das TT na região central é somada para fornecer a energia local. Extraído de [55].	30
2.14	Sequências de <i>trigger</i> simplificadas de elétrons para obtenção de dados de colisão próton-próton durante a <i>Run2</i> . Adaptado de [18].	31
2.15	Ilustração de diferentes configurações de anéis para diversas camadas dos calorímetros do ATLAS. A célula mais energética (r_1) é mostrada em vermelho, enquanto os anéis ao seu redor alternam entre branco e preto. Extraído de [18].	34
2.16	Distribuição da largura efetiva do chuveiro (R_p) unidimensional, para chuveiros produzidos por elétrons e píons, ambos em 80 GeV, avaliados no experimento SPACAL. Extraído de [20].	36
2.17	Diagrama esquemático do fluxo de trabalho na <i>Run2</i> para a avaliação da qualidade no ATLAS. Os histogramas fornecidos pelo sistema de DQ, apresentam distribuições de incluem cadeias de <i>trigger</i> previamente selecionadas para análises de qualidade de dados. Extraído de [65].	38
2.18	Visualizador <i>web</i> para comparação das distribuições de variáveis de calorimetria entre o conjunto de dados de referência (pontos azuis nos histogramas) e os dados gerados pela versão do código fonte de análise em validação (traços pretos). Na coluna da esquerda estão listadas as cadeias de <i>trigger</i> selecionadas para serem monitoradas em ambiente <i>offline</i>	40
3.1	Ilustração de um problema de classificação entre duas hipóteses (decisão binária).	44
3.2	Ilustração que exemplifica os possíveis resultados de um sistema binário de classificação.	46
3.3	Curvas ROC de dois classificadores binários hipotéticos distintos e os respectivos histogramas de saídas.	47
3.4	Diagrama em blocos do método de aprendizagem supervisionada para sistemas de classificação.	48
3.5	Grafo de fluxo de sinal ressaltando os detalhes do neurônio de saída k . Adaptado de [75].	49
3.6	Fluxograma de um GA genérico com as principais operações.	53

3.7	Ilustração de uma partícula com posição $\mathbf{p}^{(l)}$ que se move no espaço de busca, sendo influenciada pela partícula do enxame com melhor função custo ($\mathbf{g}_{best}(n)$) e da melhor posição ocupada pela partícula l na iteração n ($\mathbf{p}_{best}^{(l)}$). Os coeficientes cognitivo (c_1) e social (c_2) ponderam a influência de $\mathbf{p}_{best}^{(l)}$ e $\mathbf{g}_{best}(n)$ respectivamente na definição da velocidade. O coeficiente de inércia $w(n)$ ajusta o passo de atualização da velocidade $\mathbf{v}^{(l)}(n+1)$. Adaptado de [101].	55
3.8	Ilustração do movimento provocado a cada geração pelo método CMA-ES aos indivíduos de uma população em um espaço de busca hipotético. Adaptado de [104].	56
4.1	Fluxo de processamento do sistema de classificação proposto.	62
4.2	Ilustração dos anéis gerados em uma das camadas do calorímetro EM, com a indicação das distâncias unitárias entre os anéis concêntricos.	63
4.3	Ilustração do fluxo de processamento para computação da função objetivo e otimização dos parâmetros da etapa de pré-processamento.	64
4.4	Composição das variáveis de um indivíduo (rede neural) a ser otimizado pelos algoritmos bio-inspirados.	66
4.5	Ilustração do fluxo de execução método de treinamento proposto, utilizando o Algoritmo Genético desenvolvido.	67
4.6	Ilustração do fluxo de execução método de treinamento proposto utilizando o algoritmo PSO implementado.	69
4.7	Ilustração do fluxo de execução do método de treinamento proposto com o algoritmo CMA-ES.	70
4.8	Ilustração do mapeamento R_p multidimensional aplicado aos sinais em anéis. A energia e a distância ao centro de cada anel é ponderada com os coeficientes α e β respectivamente. Estes coeficientes são ajustados juntamente com os pesos da rede neural, a cada iteração de treinamento com o algoritmo <i>backpropagation</i>	72
4.9	Fluxograma do procedimento de validação cruzada e utilização dos conjuntos de dados experimentais utilizados para o projeto dos classificadores neurais.	73
4.10	Distribuição dos conjuntos de dados experimentais de 2017, em função da energia (a) e da coordenada η do ATLAS.	75
4.11	Razão de energia hadrônica (R_{HAD}) para (a) elétrons e Jatos com $E_T \geq 50$ GeV e (b) $15 \leq E_T[\text{GeV}] < 20$, do conjunto de dados experimentais de 2017.	76

4.12	Sinais em anéis da faixa $E_{T2} \times \eta_0$ (média e desvio padrão) para elétrons (a) e jatos hadrônicos (b) do conjunto de dados experimentais de 2017.	77
4.13	Ilustração do efeito do empilhamento (linha de cor magenta) provocado pela incidência de partículas que geraram os sinais em cor preta e vermelha, espaçadas em 50 ns. Extraído de [124].	79
4.14	Distribuição do índice de empilhamento de sinais $\langle \mu \rangle$, para a subdivisão do espaço de fase $E_{T2} \times \eta_0$.	80
5.1	<i>Boxplot</i> da taxa de falso alarme para o ponto de operação <i>Tight</i> ($P_D[\%] = 99, 83$), encontrada para cada método de treinamento bioinspirado avaliado, em função do número de neurônios na camada oculta. Resultados obtidos usando o conjunto de dados da subdivisão do espaço de fase $E_{T2} \times \eta_0$.	83
5.2	<i>Boxplot</i> da taxa de falso alarme para o ponto de operação <i>Tight</i> ($P_D[\%] = 99, 83$), encontrada para o método RingRp, em função do número de neurônios na camada oculta. Resultados obtidos usando o conjunto de dados da subdivisão do espaço de fase $E_{T2} \times \eta_0$.	84
5.3	Perfil (média e desvio padrão) dos sinais em anéis do espaço de fase $E_{T2} \times \eta_0$ (a) e o perfil dos sinais em anéis processados pelo mapeamento R_p (b). A esquerda é apresentado o perfil de anéis para assinaturas de elétrons e a direita o perfil de anéis para assinaturas de jatos.	87
5.4	Variação dos valores de α e β ocorrida durante o procedimento de otimização do classificador neural treinado pelo método RingRp, que obteve melhor desempenho do índice SP. Espaço de fase $E_{T2} \times \eta_0$.	88
5.5	Perfil (média e desvio padrão) dos sinais em anéis do espaço de fase $E_{T2} \times \eta_2$ (a) e o perfil dos sinais em anéis processados pelo mapeamento R_p (b). A esquerda é apresentado o perfil de anéis para assinaturas de elétrons e a direita o perfil de anéis para assinaturas de jatos.	90
5.6	Variação dos valores de α e β ocorrida durante o procedimento de otimização do classificador neural, que obteve melhor desempenho do índice SP. Espaço de fase $E_{T2} \times \eta_2$.	91
5.7	Perfil (média e desvio padrão) dos sinais em anéis do espaço de fase $E_{T2} \times \eta_4$ (a) e o perfil dos sinais em anéis processados pelo mapeamento R_p (b). A esquerda é apresentado o perfil de anéis para assinaturas de elétrons e a direita o perfil de anéis para assinaturas de jatos.	93

5.8	Varição dos valores de α e β ocorrida durante o procedimento de otimização do classificador neural, que obteve melhor desempenho do índice SP. Espaço de fase $E_{T2} \times \eta_4$.	94
5.9	Distribuição da grandeza $W_{\eta 2}$ para elétrons e jatos hadrônicos, na subdivisão do espaço de fase $E_{T2} \times \eta_0$ (a) e $E_{T2} \times \eta_4$ (b).	100
5.10	Distribuição da grandeza $W_{\eta 2}$ para elétrons e jatos hadrônicos, na subdivisão do espaço de fase $E_{T2} \times \eta_2$ (<i>crack</i>).	101
5.11	Histogramas bidimensionais para os sinais de saída do classificador neural para elétrons (a) e jatos hadrônicos (b). O classificador atua no espaço de fase $E_{T2} \times \eta_0$ no ponto de operação <i>Tight</i> . A curva em azul representa o ajuste do limiar de decisão da rede em função do nível de empilhamento, para manter a taxa de $P_D[\%] = 99,18$.	103
5.12	Histogramas bidimensionais para os sinais de saída do classificador neural para elétrons (a) e jatos hadrônicos (b). O classificador atua no espaço de fase $E_{T2} \times \eta_0$ no ponto de operação <i>VeryLoose</i> . A curva em azul representa o ajuste do limiar de decisão da rede em função do nível de empilhamento, para manter a taxa de $P_D[\%] = 99,18$.	105
B.1	Diagrama com o fluxo de operação dos algoritmos implementados para monitoração da etapa de reconstrução no HLT (etapa precisa).	136
B.2	Arquivo de teste para verificação dos histogramas de saída das variáveis selecionadas para monitoração.	137

Lista de Tabelas

2.1	Descrição da Intensidade Relativa, partícula mediadora, <i>spin</i> e massa relacionadas com as três forças fundamentais apresentadas no Modelo Padrão da Física de Partículas.	12
2.3	Número de anéis concêntricos gerados em cada camada do calorímetro.	34
2.4	Faixas de E_T e η definidas para a operação do conjunto de redes neurais do <i>NeuralRinger</i> no HLT do ATLAS.	35
2.5	Lista de algumas das principais grandezas utilizadas pelo ATLAS, para a descrição do formato do chuveiro (<i>shower shapes</i>) calculadas pelos algoritmos de reconstrução.	39
3.1	Matriz de confusão para um prolema de classificação binária.	47
4.1	Resumo dos principais parâmetros utilizados no treinamento dos classificadores.	65
4.2	Características do GA implementado.	68
4.3	Parâmetros do algoritmo PSO implementado.	69
4.4	Parâmetros do algoritmo CMA-ES implementado.	71
4.5	Resumo dos principais parâmetros utilizados para o treinamento dos classificadores neurais com <i>backpropagation</i> .	72
4.6	Dados da luminosidade de pico, luminosidade integrada, valores de pico e médios da taxa de empilhamento durante as colisões de prótons em 2017. Extraído de [18].	75
4.7	Número de exemplos contidos no conjunto de dados experimentais de 2017 divididos por faixas de E_T e η (azul para elétrons e vermelho para jatos hadrônicos).	77
4.8	Descrição da notação utilizada para indicar os métodos de classificação de elétrons com $E_T \geq 15$ GeV avaliados neste trabalho.	78
5.1	Comparação de desempenho entre os classificadores avaliados para o ponto de operação <i>Tight</i> . Os classificadores foram treinados para a subdivisão do espaço de fase $E_{T2} \times \eta_0$.	84

5.2	Comparação de desempenho entre os classificadores avaliados para o ponto de operação <i>VeryLoose</i> . Os classificadores foram treinados na faixa $E_{T2} \times \eta_0$.	85
5.3	Valores (média e desvio padrão) dos coeficientes α e β encontrados pelos métodos de treinamento propostos. Os classificadores foram treinados na faixa $E_{T2} \times \eta_0$.	86
5.4	Comparação de desempenho entre os classificadores avaliados para o ponto de operação <i>Tight</i> . Os classificadores foram treinados na faixa $E_{T2} \times \eta_2$.	89
5.5	Comparação de desempenho entre os classificadores avaliados para o ponto de operação <i>VeryLoose</i> . Os classificadores foram treinados na faixa $E_{T2} \times \eta_2$.	89
5.6	Comparação de desempenho entre os classificadores avaliados para o ponto de operação <i>Tight</i> . Os classificadores foram treinados na faixa $E_{T2} \times \eta_4$.	92
5.7	Comparação de desempenho entre os classificadores avaliados para o ponto de operação <i>VeryLoose</i> . Os classificadores foram treinados na faixa $E_{T2} \times \eta_4$.	92
5.8	Resultados obtidos para o treinamento dos classificadores neurais em todo o espaço de fase. São indicados os valores para o ponto de operação <i>Tight</i> encontrados para as taxas de P_D , P_F e SP (média e desvio padrão), dos classificadores neurais com melhor desempenho em cada sorteio da validação cruzada.	96
5.9	Resultados obtidos para o treinamento dos classificadores neurais em todo o espaço de fase. São indicados os valores para o ponto de operação <i>VeryLoose</i> encontrados para as taxas de P_D , P_F e SP (média e desvio padrão), dos classificadores neurais com melhor desempenho em cada sorteio da validação cruzada.	98
5.10	Média e desvio padrão dos coeficientes α e β encontrados pelo método RingRp dos classificadores neurais com melhor desempenho de índice SP nas inicializações de treinamento.	99
5.11	Valores dos indicadores de desempenho $SP[\%]$, $P_D[\%]$ e $P_F[\%]$, usando todo o conjunto de dados disponível como entrada para os classificadores do <i>ensemble</i> , com taxa de detecção ajustada para o ponto de operação <i>Tight</i> .	101
5.12	Valores dos indicadores de desempenho $SP[\%]$, $P_D[\%]$ e $P_F[\%]$, usando todo o conjunto de dados disponível como entrada para os classificadores do <i>ensemble</i> , com taxa de detecção ajustada para o ponto de operação <i>VeryLoose</i> .	102

5.13	Taxas de P_D e P_F obtidas após o ajuste da função de saída dos clas-	
	sificadores neurais em função da taxa de empilhamento, para o ponto	
	de operação <i>Tight</i> .	104
5.14	Taxas de P_D e P_F obtidas após o ajuste da função de saída dos clas-	
	sificadores neurais em função da taxa de empilhamento, para o ponto	
	de operação <i>VeryLoose</i> .	105

Símbolos e Abreviaturas

Símbolos

c	Velocidade da luz no vácuo
E	Energia
E_T	Energia Transversa
E_T^{miss}	Energia Transversa perdida
E_{Ratio}	Razão de energia
eV	Elétron-Volt
e^-/j	Elétron/Jato
fit	função aptidão
f_1	fração da energia reconstruída na primeira camada do calorímetro
fb^{-1}	femtobarn inverso, unidade de luminosidade do feixe de partículas
Hz	Hertz: Unidade de medida de frequência que indica ciclos por segundo
L	Luminosidade
m	massa
p_T	Momento transverso
R_η	Vazamento de energia para o calorímetro hadrônico
R_p	Grandeza que indica a largura efetiva do chuveiro
$W_{\eta 2}$	Grandeza que indica a largura do chuveiro
T	Tesla: Unidade da densidade de fluxo magnético
η	Pseudorapidez (Coordenada do ATLAS)
θ	Ângulo polar (Coordenada do ATLAS)
ϕ	Ângulo azimutal (Coordenada do ATLAS)
$\langle\mu\rangle$	Número médio de interações por cruzamento de feixes
γ	Símbolo utilizado para indicar a partícula Fóton
$p - p$	Colisão próton-próton

Abreviaturas

ATLAS	<i>A Toroidal LHC ApparatuS</i> (Experimento do LHC)
CERN	Centro Europeu para Pesquisa Nuclear
CMA-ES	<i>Covariance Matrix Adaptation - Evolution Strategy</i>
CMS	<i>Compact Muon Solenoid</i>
CPU	<i>Central Process Unit</i>
CTP	<i>Central Trigger Processor</i>
EM1	Primeira camada do calorímetro eletromagnético
EM2	Segunda camada do calorímetro eletromagnético
EM3	Terceira camada do calorímetro eletromagnético
FPGA	<i>Field-Programmable Gate Array</i>
GA	<i>Genetic Algorithm</i>
HEP	<i>High Energy Physics</i>
L1	Energia da segunda camada eletromagnética do ATLAS
FPGA	<i>Field-Programmable Gate Array</i>
HAD0	primeira camada do calorímetro hadrônico
HAD1	segunda camada do calorímetro hadrônico
HAD2	terceira camada do calorímetro hadrônico
HEP	<i>High Energy Physics</i>
HLT	<i>High Level Trigger</i>
ICA	<i>Independent Components Analysis</i>
LHC	<i>Large Hadron Collider</i>
L1	Primeiro nível de filtragem do ATLAS
L2	Segundo nível de filtragem do ATLAS
MAGIC	<i>Major Atmospheric Gamma Imaging Cherenkov Telescopes</i>
MSE	<i>Mean Square Error</i>
P_D	Probabilidade de Detecção
P_F	Probabilidade de Falso Alarme
PSO	<i>Particle Swarm Optimization</i>
PS	Camada Pré-amostrador
RNA	Rede Neural Artificial
ROC	<i>Receiver operating characteristic</i>
RoI	Região de interesse
SP	Índice Soma-Produto
TDAQ	<i>Trigger and data acquisition</i>

Capítulo 1

Introdução

O reconhecimento de padrões é uma tarefa realizada cotidianamente por seres vivos. Inspirado no modo como organismos inteligentes realizam o processo de tomada de decisão, sistemas computacionais para reconhecimento de padrões têm sido desenvolvidos, sendo empregados em diversas áreas do conhecimento. Em casos extremos nos quais os eventos de interesse são raros e a quantidade de atributos a serem analisados é grande, o sistema de reconhecimento de padrões precisa ser projetado adequadamente. Sistemas de classificação que atuam neste contexto enfrentam desafios em sua concepção, pois precisam operar com elevada eficiência de detecção e baixas taxas de falso alarme. Estes desafios são ainda mais destacados quando a frequência de entrada de informação é elevada pois além de preservar a informação de interesse, é necessário atender possíveis restrições temporais para tomada de decisão. Estes aspectos podem ser observados em distintos problemas científicos, como a classificação de imagens espaciais [1], classificação de sinais em radioastronomia [2], detecção de raios-gama por telescópios [3,4] e identificação de corpos celestes [5]. Nestes casos, o ruído de fundo (sinais das classes que não são interessantes) apresenta elevada ocorrência e em muitos casos pode apresentar assinatura semelhante à classe que se deseja observar.

Para atender à eficiência requerida, o projeto do sistema de classificação deve considerar as informações especialistas do problema, que podem estar relacionadas ao procedimento de extração de atributos discriminantes dos sinais, ou até mesmo das taxas de eficiência mínima requerida para operação. Dentre os sistemas de classificação com destaque relevante na literatura recente, Redes Neurais Artificiais (RNA) apresentam um vasto leque de aplicações em problemas de reconhecimento de padrões. Características como: alto grau de paralelismo, capacidade de reconfiguração, generalização e velocidade de execução, permitem sua utilização em distintos problemas científicos [6].

A elevada taxa de produção da informação, presença de grande massa de dados

de ruído de fundo e raridade de ocorrência em assinaturas de interesse são características presentes no experimento ATLAS (*A Toroidal LHC ApparatuS*) [7], localizado em um dos pontos de colisão do acelerador de partículas LHC (*Large Hadron Collider*) [8], instalado no CERN (*European Organization for Nuclear Research*), na fronteira entre a Suíça e a França. O ATLAS é um experimento que investiga os constituintes fundamentais da matéria, por meio das interações de feixes de partículas colididos pelo LHC com seus subsistemas de detecção [9]. Dentre o vasto programa de física do LHC, elétrons são partículas que apresentam destacado interesse para o experimento, devido a suas relações com modelos de física teórica, alguns deles ainda não verificados experimentalmente [10].

O ATLAS é dotado de um sistema de seleção *online* de eventos (filtragem) de elétrons com dois níveis sequenciais (*trigger*), usado para reduzir gradualmente o fluxo de dados [11]. O primeiro nível realiza uma seleção menos rigorosa de eventos, enquanto o segundo nível, que analisa apenas os eventos aceitos pelo nível anterior, é dotado de algoritmos com maior complexidade, realizando uma análise mais apurada para a tomada de decisão. Após a filtragem *online*, as informações do evento são armazenadas em mídia permanente, permitindo a análise final *offline*. Este sistema de filtragem baseia-se em estratégias de processamento distribuído, utilizando sistemas embarcados de alta velocidade no primeiro nível, e algoritmos de processamento rápido no segundo nível.

Na tomada de dados prevista para 2021-2024 (*Run3*), o LHC pretende elevar a energia envolvida na colisão, bem como aumentar progressivamente o número de colisões por cruzamento de feixes de prótons, o que exige a atualização dos sistemas de filtragem dos experimentos, para identificação das partículas de interesse [12]. Os dados gerados a cada colisão ocupam cerca de 1,5 MB, o que considerando a frequência de colisões a 40 MHz, corresponde a uma taxa de 60 TB/s e torna impeditivo o armazenamento de todos os dados gerados.

1.1 Motivação

Experimentos de Física de Altas Energias (ou HEP, do inglês *High Energy Physics*) têm como objetivo a investigação das propriedades e das formas de interação fundamentais entre as partículas elementares [13]. Em 4 de Julho de 2012, os experimentos ATLAS e CMS (*Compact Muon Solenoid*) [14], ambos integrados ao LHC, observaram experimentalmente com elevado nível de precisão uma partícula compatível com o bóson de Higgs previsto pelo Modelo Padrão [15], o que aprimorou o conhecimento sobre interações nucleares. As descobertas realizadas pelo ATLAS desde o primeiro ciclo de operação do LHC (*Run1* - 2010-2013) têm crescido co-

nhhecimento da composição do Modelo Padrão da Física de Partículas (ou SM - do inglês *Standard Model*). Com o objetivo de acumular uma maior quantidade de informação para verificação de processos físicos mais raros que o bóson de Higgs, o LHC vem elevando a taxa de colisões próton-próton. Para isso, o LHC pretende no período de tomada de dados nos anos de 2022-2025 (*Run3*) atingir 14 TeV, que é o máximo nível de energia de colisão previsto no projeto original, bem como aumentar o número de colisões por cruzamento de feixes.

Esta estratégia permite produzir mais eventos a serem observados, contudo acresce desafios para as áreas de eletrônica e processamento de sinais envolvidas nos subsistemas de detecção do ATLAS, devido a uma maior ocupação nos canais dos detectores, produzindo sobreposição de sinais de diferentes colisões (empilhamento). O elevado grau de sobreposição de sinais pode descaracterizar o formato esperado do sinal de deposição de energia nos sensores, podendo deteriorar a eficiência dos algoritmos de seleção de eventos. O LHC prevê operações além do ano de 2030, período no qual pretende-se continuamente realizar o aumento do número médio de colisões por cruzamento de feixe [16]. O aumento da ocupação dos canais, bem como as elevadas taxas de sobreposição de sinais nos sensores tornam o cenário desafiador para os sistemas de *trigger* do ATLAS. Caso o grande volume de eventos não relevantes venha a ser gravado para análise *offline*, a detecção da física de interesse, em geral envolvida com raros decaimentos pode ser comprometida. Este contexto acentua a necessidade do aprimoramento do sistema de seleção *online*, para aumentar a rejeição da parcela de informação não relevante [17]. Ainda, considerando a raridade dos eventos interessantes envolvidos, o sistema de filtragem precisa operar com elevados índices de eficiência, de modo a reduzir a probabilidade de um evento interessante ser rejeitado no processo de filtragem. Os elevados índices de empilhamento podem afetar a eficiência dos algoritmos de filtragem *online*, uma vez que os algoritmos de estimação de energia podem ser afetados e as assinaturas processadas pelos algoritmos de filtragem podem sofrer distorções. O aumento do volume de dados a ser processado pode comprometer os algoritmos envolvidos na filtragem *online*, por conta de limitações de taxa de transmissão para o processamento.

Na primeira etapa do segundo nível de filtragem *online* do ATLAS, o *NeuralTrigger* é o método utilizado para classificação de elétrons [18]. Este método utiliza o conhecimento especialista em calorimetria para a extração da informação discriminante do chuveiro de partículas desenvolvido nos calorímetros. Em sua etapa inicial um algoritmo de extração de atributos discriminantes é executado onde são gerados para cada camada do calorímetro uma quantidade de anéis em torno da célula mais energética. Em seguida, a energia medida em cada célula do calorímetro pertencente ao anel é somada, compreendendo em todas as camadas um total de 100 atributos (anéis) a serem utilizados como entrada para os classificadores neurais. O conjunto

de redes neurais utilizado é treinado por intervalos de energia e intervalos de posição de interação das partículas no sistema de calorimetria. A estratégia do conjunto, composto por 25 redes neurais, é motivada por mudanças nas respostas dos sensores para variações na energia e posição de interação, fornecendo assim classificadores especialistas por faixas de energia e posição de interação da partícula [19].

Extensas investigações nas áreas de calorimetria em altas energias indicam que a largura do chuveiro de partículas que se desenvolve no calorímetro [20] pode ser utilizada como informação para discriminação de assinaturas de elétrons e jatos. Jatos hadrônicos apresentam em geral um perfil de deposição de energia que se espalha mais lateralmente em relação à célula mais energética nas camadas do calorímetro, em comparação com as assinaturas de elétrons [21]. A utilização de grandezas que descrevem o espalhamento longitudinal e lateral são comumente empregadas em experimentos de física de partículas. Especificamente, no trabalho [22], o perfil lateral do chuveiro foi usado em um mapeamento multidimensional de anéis topológicos em uma investigação primária para condições previstas para o LHC. A inclusão desta informação especializada em calorimetria (largura do chuveiro), permitiu a melhoria na detecção de elétrons e redução do ruído de fundo considerando as restrições e condições de operação estimadas na época.

Outro requisito que ganha importância quando os eventos que se deseja detectar apresentam raridade de ocorrência são as restrições de operação, associadas às taxas de eficiência de detecção do classificador. No cenário com elevado volume de eventos, os classificadores precisam operar com pontos de operação que preservem a identificação das assinaturas de interesse (alta taxa de detecção), enquanto minimiza o quanto for possível a aprovação de falsos positivos. Embora os algoritmos de treinamento de redes neurais, baseados na retropropagação do erro tenham alcançado resultados importantes em problemas de classificação, o método *backpropagation* [23] otimiza os pesos da rede usando o erro médio quadrático entre a saída obtida pela rede e o alvo associado às classes. Este procedimento não leva em consideração as restrições de desempenho vinculadas a taxa de detecção ou a taxa de falso alarme esperada, para atender os requisitos da aplicação.

Considerando os aspectos citados, que envolvem a utilização do conhecimento especialista em calorimetria, assim como o uso das informações relacionadas as restrições de operação em termos de eficiência e taxas de falso alarme, esta tese de doutorado apresenta contribuições para a melhoria do método de classificação *NeuralRinger*. É proposta neste trabalho a adição de uma etapa de pré-processamento, que pondera o perfil dos anéis obtidos pelo *NeuralRinger*, usando a energia e a distância do anel à região de interesse. Esta estratégia permite destacar na base de anéis a informação especialista discriminante da largura do chuveiro de partículas no sinais medidos. Para realizar este pré-processamento (e definir os coeficientes

que ponderam a energia e a distância entre os anéis), métodos de treinamento de redes neurais são propostos nesta tese de doutorado, que ajustam simultaneamente os pesos da rede e os coeficientes da etapa de pré-processamento. Algoritmos de otimização bioinspirados e um método baseado em retropropagação foram adaptados para realizar o treinamento da rede e ajustar os coeficientes da etapa de pré-processamento.

1.2 Objetivo

Este trabalho tem como objetivo principal desenvolver novas estratégias para aprimoramento do *NeuralRinger*, algoritmo de identificação de elétrons utilizado no segundo nível de filtragem no ATLAS. Neste contexto, esta tese apresenta um novo método de treinamento de redes neurais utilizando algoritmos bio-inspirados, que efetua o treinamento dos classificadores para melhoria de desempenho em pontos específicos de operação. Um método de pré-processamento com informação especializada de calorimetria foi desenvolvido para destacar a informação lateral do chuveiro de partículas, permitindo um melhor desempenho na separação elétrons/jatos. Métodos de treinamento de redes neurais foram adaptados, para ajustar simultaneamente os pesos dos classificadores bem como coeficientes da etapa de pré-processamento. Os objetivos específicos deste trabalho são:

1. Desenvolver um método de treinamento para ajuste simultâneo dos pesos da rede e dos coeficientes da etapa de pré-processamento especialista;
2. Avaliar a eficiência de discriminação de elétrons obtida com os métodos propostos para o treinamento de redes neurais;
3. Analisar os coeficientes obtidos da etapa de pré-processamento para as distintas faixas de energia e região de interação do calorímetro;
4. Reduzir a aceitação de falsos elétrons em diferentes pontos de operação.

1.3 Principais Contribuições desta Tese

Tendo em vista os planos para o *upgrade* do LHC, que progressivamente pretende aumentar o número médio de colisões por cruzamento do feixes de partículas, bem como o aumento do centro de massa de energia de colisão, os algoritmos de classificação incluídos na etapa de seleção *online* precisam ser melhorados.

Esta tese de doutorado investiga uma nova abordagem para a classificação de elétrons no segundo nível de seleção *online* do ATLAS. É proposta uma etapa de pré-processamento que permite destacar a informação lateral do chuveiro de partículas.

Esta estratégia explora o atributo lateral discriminante do chuveiro, haja vista que o conhecimento especialista da física teórica indica que elétrons apresentam perfil lateral de deposição de energia menos largo em relação a jatos hadrônicos.

Para ajustar os coeficientes da etapa de pré-processamento, que ponderam a energia medida em anéis e a distância dos anéis ao primeiro anel, métodos de treinamento de redes neurais artificiais são propostos. É apresentado um método baseado em algoritmos bio-inspirados que realiza o treinamento dos classificadores, utilizando como figura de mérito as restrições de taxas de probabilidade de detecção ou de falso alarme presentes no ambiente de operação do sistema de filtragem. Um método baseado em retropropagação foi também adaptado para ajustar simultaneamente os pesos da rede neural e os coeficientes da etapa de pré-processamento. Estas estratégias permitem que a etapa de pré-processamento especialista e os sistemas de classificação sejam conjuntamente ajustados, encontrando um melhor mapeamento para a classificação dos eventos. Os métodos foram avaliados em dados experimentais do ATLAS em colisões realizadas nos anos de 2017. Comparações foram realizadas entre os métodos propostos e o *NeuralRinger*, em relação às taxas de falso alarme encontradas para os distintos pontos de operação, bem como foram realizadas verificações no impacto dos métodos sobre a eficiência geral do *trigger*.

Durante o desenvolvimento desta tese, o autor foi responsável pelo desenvolvimento, manutenção e implementação de módulos dos sistemas computacionais que realizam a monitoração dos algoritmos do sistema *online* de filtragem do ATLAS, para os canais de elétrons e fótons. Os desenvolvimentos foram realizados ao longo do processo de qualificação¹ como autor do experimento ATLAS, onde foram conduzidas implementações para monitoração *online* e *offline* de grandezas que descrevem o chuveiro de partículas nas cadeias de filtragem e algoritmos de reconstrução, bem como a avaliação de qualidade dos dados experimentais a serem utilizados como referência durante as colisões. As contribuições do autor permitiram melhorias e novas funcionalidades ao *software* principal de análise do experimento, no contexto da preparação para o período de tomada de dados agendado para 2022-2025.

¹Para ser enquadrado como autor do experimento ATLAS, os pesquisadores precisam cumprir uma tarefa de qualificação, que em geral compreende o desenvolvimento, validação e análises de atividades urgentes/importantes para o avanço do experimento. O processo é realizado em um estágio de duração mínima de um ano, sob a orientação de membros do experimento para acompanhamento e discussão dos avanços do desenvolvimento, sendo apresentados em reuniões periódicas à grupos do experimento com responsabilidades específicas (no caso deste trabalho, o trabalho esteve ligado ao grupo *TrigEgamma Signature*, grupo responsável pelo sistema de filtragem *online* de elétrons e fótons do ATLAS). Tradicionalmente, após o período de qualificação o autor mantém contribuições, adicionalmente com outras responsabilidades para manutenção, ajustes e melhorias do experimento.

1.4 Organização do Documento

Este documento está organizado do seguinte modo. No Capítulo 2 é apresentado ao leitor o ambiente de física de altas energias, dando enfoque ao acelerador LHC e ao experimento ATLAS. São apresentados os conceitos que envolvem a interação de partículas com calorímetros, os principais objetivos do experimento ATLAS bem como a descrição dos seus diversos subsistemas de detecção, detalhando o sistema de filtragem *online* de elétrons do ATLAS. Os principais algoritmos e técnicas empregadas para classificação e reconstrução de eventos no *trigger* são abordadas. O Algoritmo *NeuralRinger* é detalhado, apresentando também o histórico de seu desenvolvimento no contexto do ATLAS. Neste capítulo, também é realizada uma revisão bibliográfica dos trabalhos que foram desenvolvidos ao longo dos anos, com o objetivo do aprimoramento do *NeuralRinger* desde a sua concepção.

O Capítulo 03 é reservado para detalhamento de métodos de treinamento de redes neurais e algoritmos bio-inspirados. São apresentados também os principais conceitos de sistemas de classificação e aplicação de redes neurais em experimentos de física de partículas. O Capítulo 04 é dedicado para detalhar a metodologia proposta para treinamento e ajuste (pós-treinamento) dos modelos neurais para classificação de elétrons. Neste capítulo, é descrito o procedimento de avaliação de desempenho para comparação entre os métodos de treinamento de redes neurais avaliados neste trabalho. São também apresentadas as bases de dados experimentais utilizadas.

O Capítulo 05 é destinado a apresentar os resultados obtidos com os métodos de treinamento de redes neurais avaliados. Serão apresentados resultados encontrados utilizando dados experimentais provenientes de colisões realizadas em 2017 e 2018 no LHC. No Capítulo 06, as considerações finais e trabalhos futuros são discutidos, que é prosseguido por dois apêndices: A produção acadêmica realizada ao longo do desenvolvimento desta tese (Apêndice A). Alguns aspectos das contribuições do autor desta tese, no sistema de monitoração e qualidade de dados, dos algoritmos de hipóteses e reconstrução do *trigger* do ATLAS, durante seu estágio de qualificação no experimento são apresentados no Apêndice B.

Capítulo 2

Física de Partículas e o Experimento ATLAS

A física experimental de partículas elementares é um campo no qual se investigam, por meio de experimentos, as partículas e forças que constituem a matéria. Os resultados experimentais apoiam o desenvolvimento, ajustes e proposição de novas teorias que buscam explicar as interações fundamentais das diversas partículas e forças observadas até hoje [24].

A discussão formalizada sobre os constituintes básicos da matéria tem origem na Grécia Antiga no século VI a.C., onde alguns conceitos foram postos por meio de observações. Tales de Mileto (623-556 a.C.) elaborou observações sobre a relação da água com as sementes que originavam plantações, e por este motivo, associou a água como um elemento fundamental [25]. Ainda na Grécia Antiga, Demócrito (460-370 a.C.), por proposições teóricas, afirmou que toda a matéria poderia ser dividida até um determinado ponto, que seria sua parte mais fundamental e indivisível, o que ele denominou como átomo [26]. A teoria atômica foi revisitada, por meio de experimentos sobre a absorção de gases pela água, em 1803 pelo químico John Dalton (1766-1844), que publicou o trabalho “Absorção de Gases Pela Água e Outros Líquidos” [27]. Muitas investigações em física de partículas obtiveram destaque no fim do século XIX e início do século XX. Em 1898, os físicos Marie Skłodowska-Curie e Pierre Curie assinaram em conjunto um artigo relatando a descoberta experimental de um novo elemento químico radioativo, imediatamente denominado polônio. Em dezembro do mesmo ano, publicaram a descoberta de outro elemento químico radioativo, que chamaram de rádio¹. Em 1913 Niels Bohr (1885-1962) propôs um modelo inspirado nas investigações conduzidas por Rutherford, que afirmou que o

¹Tais contribuições conferiram o Prêmio Nobel de Física de 1903 para Marie Curie, Pierre Curie e Henry Becquerel pela descoberta da radioatividade espontânea e fenômenos de radiação [28]. Em 1911, Marie Curie foi contemplada mais uma vez com o Prêmio Nobel de Química, devido aos estudos do isolamento do rádio, descoberto por meio de procedimentos experimentais [29].

átomo teria um núcleo de carga elétrica positiva de tamanho muito pequeno em relação ao seu tamanho total, sendo que este núcleo, seria rodeado por elétrons de carga elétrica negativa [30]. O modelo atômico de Borh aperfeiçoou o modelo de Rutherford, propondo principalmente que os elétrons não se movem aleatoriamente ao redor do núcleo mas sim em órbitas circulares, onde cada órbita apresenta níveis de energia definidos².

O conhecimento dos processos atômicos e interações das partículas com a matéria permitiram o desenvolvimento de experimentos para física de partículas com elevada infraestrutura tecnológica. O CERN (*Centre Européene pour la Recherche Nucléaire*) é o maior laboratório de Física de Partículas do mundo [31], situado na região de fronteira da Suíça e a França, sendo composto por 23 estados-membros, que conta com a colaboração técnicos, físicos e engenheiros, provenientes de mais de 600 instituições ou universidades. No CERN, o primeiro acelerador de partículas foi construído em 1971, e desde então muitos experimentos foram desenvolvidos, o que resultou em descobertas importantes para a constituição do Modelo Padrão da Física de Partículas [32].

Experimentos modernos de física de partículas comumente utilizam como fontes de observação raios cósmicos, ou feixes de partículas produzidos em laboratórios [33]. Estes experimentos destacam-se por suas proporções, complexidade e grandes colaborações de instituições científicas. No caso da utilização de feixes, diferentes elementos químicos são utilizados na sua composição (hidrogênio, alumínio, entre outros) que podem ser colididos entre si, por meio de dispositivos que aceleram esses feixes. As colisões são realizadas em pontos onde se localizam os experimentos detectores que observam o subproduto das partículas que colidem [34]. Para permitir o estudo não somente da estrutura nuclear, mas também a estrutura dos prótons e nêutrons que compõem o núcleo atômico, a energia dos feixes produzidos em laboratório tem sido cada vez mais alta [35].

Outros experimentos para observação de raios cósmicos com elevada energia também destacam-se por suas proporções. Um exemplo da magnitude destes experimentos é o IceCube [36], localizado no Polo Sul. Este experimento compreende um quilômetro cúbico de gelo, dotado de instrumentos de detecção de partículas, no qual procura-se por neutrinos, que podem carregar informações de fontes astrofísicas como explosão de estrelas, buracos negros e estrelas de nêutrons. Outro exemplo notável em relação às suas proporções é o experimento CTA (*Cherenkov Telescope Array*), que consiste em um experimento para observação de astropartículas [37].

²As contribuições para o entendimento sobre a eletricidade tem se aprofundado principalmente entre o fim do século XIX e início do século XX, quando a ideia do átomo como elemento constituinte da matéria foi aceita e, com ela, a convicção de que a eletricidade é uma propriedade de partículas elementares que compõem o átomo (elétrons, prótons e nêutrons).

O experimento é dotado por um conjunto de telescópios localizados nos hemisférios sul e norte, nos quais foram projetados para observar raios-gamma com altas energias que podem carregar informações de novos corpos celestes. Estes experimentos apoiam a compreensão das partículas e interações fundamentais, bem como dão suporte a estudos sobre a origem do universo e sua composição [38,39].

O conhecimento desenvolvido por meio de proposições teóricas e confirmações (ou rejeições) de hipóteses previstas no campo da física de partículas elementares deu origem ao Modelo Padrão da Física de Partículas (SM - do inglês: *Standard Model*) [32]. O SM é composto por um conjunto de teorias que descreve as propriedades das partículas fundamentais e suas interações entre si, e tem sido constantemente ajustada por meio de resultados obtidos dos modernos experimentos de física de partículas [21,24,32].

2.1 O Modelo Padrão da Física de Partículas

O modelo padrão da física de partículas é uma teoria que fornece uma descrição unificada de três das quatro forças (força forte, força eletromagnética (EM) e força fraca) que governam todas as partículas conhecidas até então [40,41]. A Figura 2.1 [42] apresenta as partículas do SM, que podem ser separadas em grupos: quarks, léptons, os portadores de força e o bóson de Higgs. Para cada partícula apresentada no modelo, são mostradas informações a respeito da sua massa, carga elétrica e o *spin*.

- Léptons carregados: elétron, múon, tau;
- Léptons neutros (neutrinos): ν_e , (neutrino do elétron) ν_μ (neutrino do múon) e ν_τ (neutrino do tau);
- Quarks do tipo *up* com *spin* $\frac{2}{3}$: *up*, *charm*, *top*;
- Quarks do tipo *down* com *spin* $-\frac{1}{3}$: *down*, *strange*, *bottom*;
- Portadores de força: fóton, bóson W, bóson Z, glúon;
- Bóson escalar: Higgs.

As partículas indicadas no SM, são classificadas em férmions e bósons. Os 12 férmions fundamentais e suas antipartículas³, são subdivididos em 6 quarks e 6 léptons, que são organizados em 3 famílias. Os férmions apresentam massa idêntica e números quânticos opostos [40]. Experimentalmente, a principal diferença entre léptons e quarks é que os léptons podem ser observados como partículas isoladas, enquanto os quarks estão sempre ligados para formar partículas incolores com carga inteira, chamadas de hádrons [24].

³De um modo geral, uma antipartícula tem a mesma massa e o mesmo *spin* da partícula em questão, porém cargas opostas.

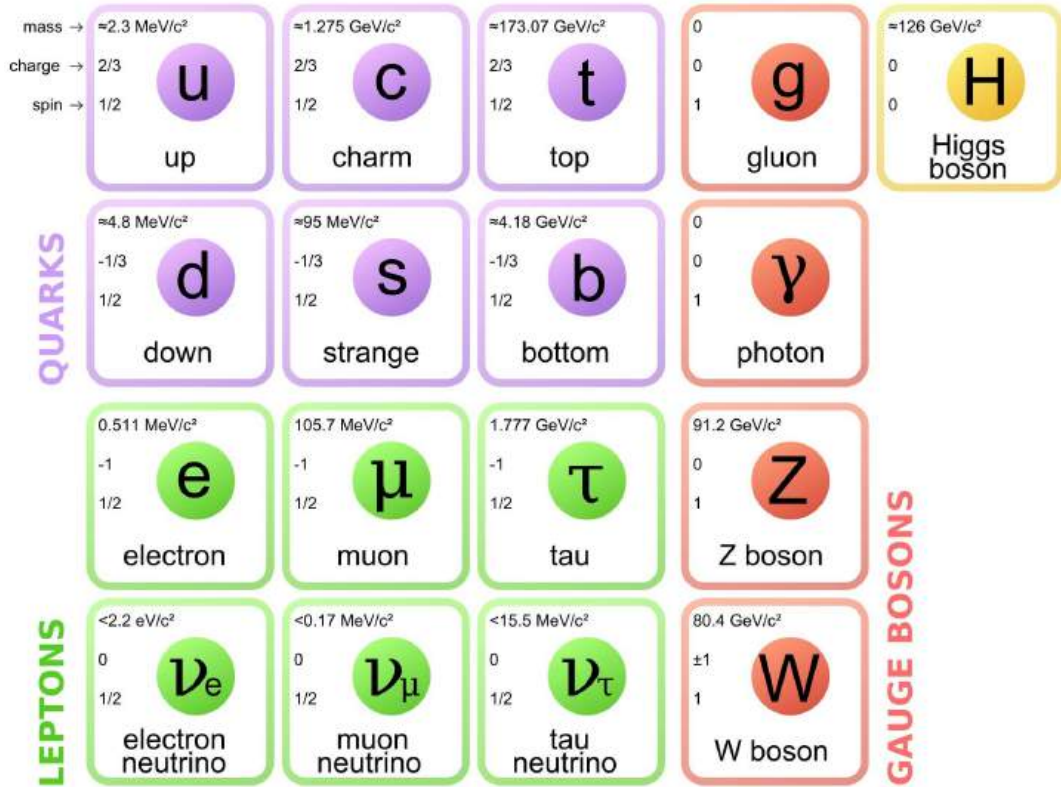


Figura 2.1: Representação gráfica do modelo padrão das partículas e interações fundamentais da matéria. Extraído de [42].

Os léptons interagem apenas por meio da força fraca e força eletromagnética, com os números quânticos correspondentes que são o isospin⁴ fraco e a carga elétrica, respectivamente [32]. Os léptons têm carga elétrica de valor inteiro, em que suas interações apresentam assinaturas destacadas que permitem identificá-los. A maioria dos experimentos possui detectores dedicados para registrar e identificar elétrons e múons, já os taus têm vida média de $2.9 \times 10^{-13} \text{s}$ e por isso, em geral são reconstruídos a partir de seus produtos de decaimento [32].

Quarks têm um número quântico adicional, a cor⁵, que é relacionada a força forte. Os quarks não são observados diretamente, eles são produzidos nas interações formando hádrons (partículas com estrutura interna) [32]. Hádrons podem ser observados pela sua interação com o material do detector ou por seus produtos de de-

⁴O isospin é um termo criado em 1961 que representa um número quântico relacionado às forças fortes no estudo das partículas elementares. Esta teoria surgiu com a constatação de que o próton e o nêutron possuem o mesmo *spin*, praticamente a mesma massa, mas possuem cargas elétricas diferentes [41].

⁵Trata-se de uma propriedade, não uma cor propriamente dita. Vermelho, verde e azul são apenas aspectos dessa propriedade. Assim como a carga elétrica, que é também uma propriedade de certas partículas, pode ser positiva ou negativa, a propriedade cor, apresenta três variedades que foram chamadas de vermelho, verde e azul.

caimento. Os hádrons são constituídos de bárions, quando formadas por três quarks ou três antiquarks, ou mésons, quando constituídos por um quark e um antiquark. A estrutura instável dos hádrons faz com que gerem uma sequência de decaimentos até atingirem o material dos experimentos de física de partículas. Assim, normalmente suas assinaturas são chamadas de Jatos Hadrônicos, que compreende uma associação do agrupamento de partículas com um jato de água nos detectores [32].

Outras partículas fundamentais destacadas no SM são os *bósons gauge*, ou bósons intermediários, que desempenham o papel de mediadores das forças entre os férmions elementares (excitadoras de campo). Cada uma das três forças representadas no SM é descrita por uma Teoria Quântica de Campos (QFT - *Quantum Field Theory*) correspondendo à troca de um bóson *gauge* de *spin* unitário. Todas as interações entre as partículas envolvem a troca de um bóson *gauge*, em que o fóton é o mediador da interação eletromagnética, descrito pela teoria Eletrodinâmica Quântica (QED - *Quantum Electrodynamics*), que explica a força eletromagnética [32]. Os bósons Z e W mediam a força fraca e são responsáveis pelo decaimento nuclear β , enquanto a força forte é transportada pelos glúons (que mantêm o núcleo atômico coeso), mantendo os quarks unidos dentro dos hádrons [35]. Os glúons são descritos pela Cromodinâmica Quântica (QCD - *Quantum Chromodynamics*). As forças com suas intensidades relativas e partículas relacionadas são sumarizadas na Tabela 2.1.

Tabela 2.1: Descrição da Intensidade Relativa, partícula mediadora, *spin* e massa relacionadas com as três forças fundamentais apresentadas no Modelo Padrão da Física de Partículas.

Força	Intensidade Relativa	Boson	<i>spin</i>	Massa(GeV/ c^2)
Forte	1	gluon (g)	1	0
Eletromagnética	10^{-3}	fóton (γ)	1	0
Frac	10^{-8}	W bóson ($W^\pm$)	1	80,4
		Z bóson (Z)	1	91,2

Finalmente, o bóson de Higgs é uma partícula fundamental neutra com *spin* nulo, introduzida no modelo padrão através do mecanismo de geração de massa dos *bóson gauge* e de todas as outras partículas elementares consideradas na teoria [32]. A detecção do bóson de Higgs foi um grande triunfo do Modelo Padrão, dando o Prêmio Nobel de Física de 2013 a dois dos físicos que propuseram o mecanismo de Higgs, Peter Higgs e François Englert. Os experimentos ATLAS e CMS encontraram evidências experimentais da nova partícula com massa em torno de 126 GeV⁶ em

⁶A unidade eV (elétron-volt) é normalmente utilizada nos estudos de física de partículas, em que 1 eV compreende a energia necessária para aumentar o potencial elétrico de um elétron em um volt, equivalente a $1,6 \times 10^{-19}$ J no Sistema Internacional de unidades e medidas (SI).

um dos possíveis canais de decaimento do bóson de Higgs [15].

Higgs e Englert teorizaram em 1964 a existência de um campo do qual as partículas massivas obtêm massa, e associado a esse campo haveria uma partícula mediadora dessa interação, o bóson de Higgs. De acordo com o chamado mecanismo de Higgs, as partículas W e Z se chocariam incessantemente com bósons de Higgs presentes em todo o espaço, explicando assim a origem das suas massas, não explicadas completamente pela teoria da força eletrofraca [32]. O campo de Higgs difere-se dos demais campos pois, segundo a teoria, o estado de energia mínima desse campo ocorreria não quando se anulasse (como é o caso, por exemplo, do campo eletromagnético) mas em um determinado valor específico distinto de zero. É previsto pelo Modelo Padrão que o bóson de Higgs possua um tempo de vida muito curto, decaindo imediatamente para estados finais bosônicos ou fermiônicos. Isso requer o desenvolvimento de experimentos capazes de detectar e medir a energia das partículas em seus estados finais, para então reconstruir o evento que deu origem a estes decaimentos.

O SM tem dado explicações consistentes em relação as partículas fundamentais e suas interações, entretanto, muitas questões permanecem em aberto, nas quais algumas delas são apresentadas a seguir. Uma delas se refere à gravidade, a quarta força fundamental que não é unificada com as outras interações no SM. A relatividade geral, que descreve a gravidade, não é uma teoria quântica, e não foi integrada ao SM. Em decorrência do efeito desprezível da gravidade em termos subatômicos (a massa das partículas é da ordem de 10^{-27} kg), ainda não foi possível fazer observações experimentais para quaisquer propostas de teoria quântica para a mesma [32].

2.2 Grande Colisor de Hádrons

O Grande Colisor de Hádrons (LHC do inglês - *Large Hadron Collider*) [8] é o maior acelerador de partículas construído até então, localizado no CERN entre a fronteira da Suíça e a França. Dotado de uma das mais complexas instalações científicas do mundo, o LHC foi projetado para colidir partículas com centro de massa ($\sqrt{s}$) de até 14 TeV, numa frequência de 40×10^6 vezes por segundo. Os feixes que percorrem o túnel do LHC em sentidos opostos, colidem em pontos específicos nos quais experimentos foram projetados para medir as características das partículas geradas nas colisões. Conforme ilustra a Figura 2.2 [43], o LHC foi instalado em um túnel circular com aproximadamente 100 metros de profundidade e uma extensão de 27 km.

O plano de operações do LHC foi projetado para que as tomadas de dados (período de colisões de partículas) fossem realizadas com níveis de energia e frequência

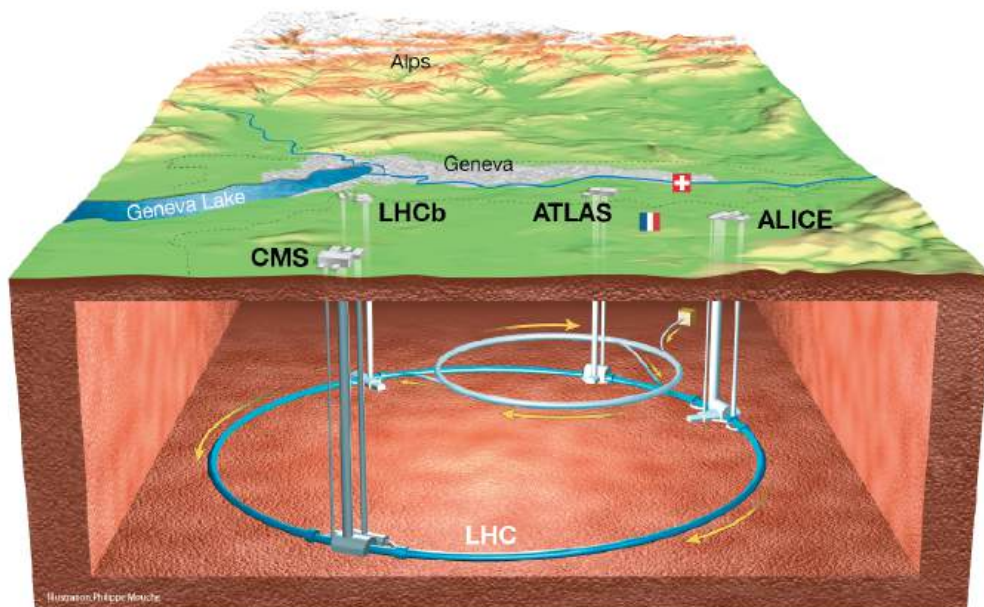


Figura 2.2: Representação gráfica da instalação subterrânea do LHC no CERN, e alguns de seus experimentos localizados em pontos de colisão de feixes. Extraído de [8].

de colisões crescentes, conforme indicado no planejamento em linha do tempo, mostrado pela Figura 2.3 [44]. Com o início da tomada de dados em 2010, o LHC operou com níveis de energia de colisão da ordem de 3,5 TeV por feixe, aumentando para 4 TeV por feixe em 2012. Após a tomada de dados em 2012 (*Run1*), o LHC passou pelo período programado de ajustes e atualização de seus experimentos. As colisões de partículas foram reestabelecidas entre os anos de 2015 e 2018 (*Run2*) com 6,5 TeV de energia por feixe. Para a *Run3*, período de tomada de dados previsto entre os anos de 2022 e 2025, o LHC pretende colidir partículas com 7 TeV para cada feixe, alcançando a máxima energia para o qual foi concebido. O LHC possui um plano de *upgrade* em que após a *Run3*, um período de atualização e ajustes dos experimentos é previsto para preparação das operações em alta luminosidade (HL - *High Luminosity*). No período HL (previsão a partir de 2027) do LHC serão geradas muito mais colisões, acumulando dez vezes mais dados do que o período predecessor. Espera-se que os experimentos instalados nos pontos de colisão sejam capazes de detectar fenômenos extremamente raros e melhorar a precisão das medições. A fim de explorar totalmente a quantidade crescente de dados, os experimentos foram submetidos a ambiciosos programas de atualização de seus detectores.

Como o LHC busca por eventos com rara ocorrência, sua operação demanda uma alta taxa de colisões. Deste modo, ele opera com uma frequência de colisão em 40

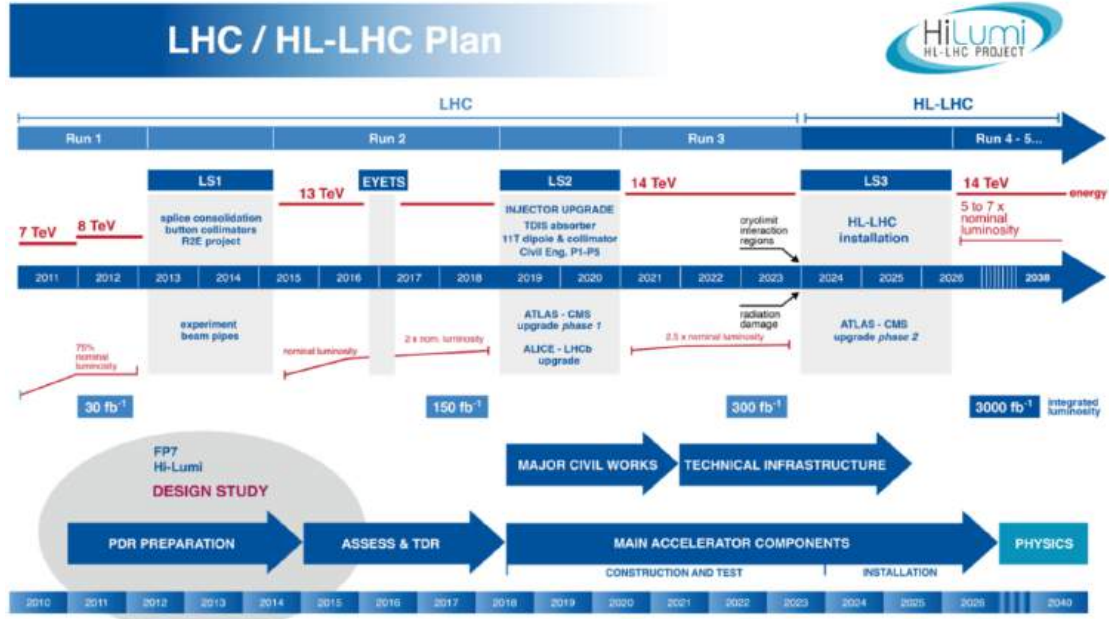


Figura 2.3: Planejamento em linha do tempo para a operação do LHC com seu aumento gradual da energia do centro de massa de suas colisões e os prazos reservados em janelas de parada para atualizações dos experimentos (*Long Shutdown*). Extraído de [44].

MHz, com as injeções e colisões dos pacotes de prótons a cada 25 ns. A medida do número de colisões por centímetro quadrado produzida a cada segundo é chamada de *Luminosidade* (L), descrita pela equação 2.1:

$$L = n \times \frac{N_1 \times N_2}{A} \times f, \quad (2.1)$$

no qual n é o número de feixe de partículas, N_i é o número de partículas em cada feixe, A é a área da seção transversal do feixe e f é a frequência de colisão. A Luminosidade é dada pelas unidades $cm^{-2}s^{-1}$ ou fb^{-1} .

Conforme indicado na Figura 2.4 [45], a aceleração dos feixes de partículas no LHC é realizada por meio de uma cadeia de aceleração, que inclui aceleradores utilizados no CERN em períodos anteriores. Os feixes de prótons são inicialmente acelerados pelo LINAC, que consiste em um acelerador linear utilizado para a extração de prótons de átomos de hidrogênio alcançando até 50 MeV. Em seguida, a aceleração é continuada por meio de uma cadeia de síncrotrons, que são aceleradores circulares que orientam as partículas utilizando magnetos.

O primeiro estágio de aceleração circular é realizado no *Proton Syncontron Booster* (BOOSTER) que alcança energia em torno de 1,4 GeV, em seguida os feixes são acelerados pelo *Proton Syncontron* (PSync) que eleva a energia dos feixes até 25 GeV. Após serem acelerados pelo PSync os feixes são encaminhados para o *Super*

Proton Síncrotron (SPS) que acelera circularmente os feixes até atingirem 450 GeV. Finalmente, após o estágio de aceleração no SPS, os feixes são encaminhados para o LHC que pode alcançar até 14 TeV [8].

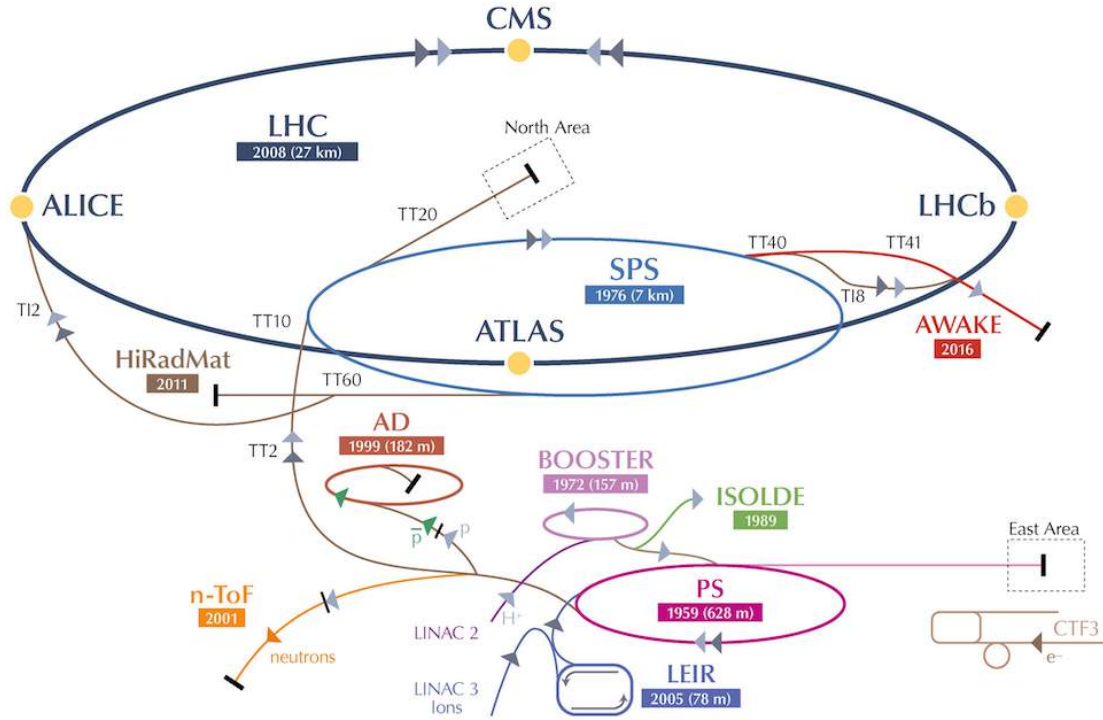


Figura 2.4: Complexo de etapas de aceleração de feixes de prótons no CERN, que compõem o acelerador LHC. Extraído de [45].

As colisões de partículas são realizadas em pontos específicos, nos quais experimentos com diferentes propósitos foram construídos. Nestes pontos, quatro experimentos são destacados em que os experimentos ATLAS (A Toroidal LHC ApparatuS) e CMS (Compact Muon Solenoid)⁷, são experimentos de Física de Partículas de propósito geral que operam em elevada luminosidade. Os experimentos ALICE (A Large Ion Collider Experiment)⁸ e LHCb⁹ (Large Hadron Collider beauty) apresentam propósitos específicos.

2.3 O Experimento ATLAS

O ATLAS [7] é o maior experimento do LHC, contando com uma colaboração (ATLAS Collaboration) composta por aproximadamente 3000 pesquisadores. O

⁷O ATLAS e o CMS são os maiores experimentos do LHC e apresentam propósitos múltiplos que investigam uma vasta gama de fenômenos previstos modelo padrão.

⁸O ALICE é um experimento especialista em estudos de íons pesados, destinado a explorar eventos físicos na interação núcleo-núcleo.

⁹O LHCb é especializado em investigar as pequenas diferenças entre matéria e antimatéria estudando a partícula b quark.

ATLAS possui o objetivo de investigar diferentes aspectos da física de partículas. Os desafios provenientes da operação em altas energias e elevada luminosidade exigem que o ATLAS apresente precisão na medição de energia durante a interação das partículas com seus subsistemas. Conforme é mostrado na Figura 2.5, o experimento ATLAS possui aproximadamente 25 m de altura e 44 m de comprimento, e com sua estrutura cilíndrica cobre aproximadamente 99% de todo o ângulo sólido em torno do ponto de colisão. Consiste sucessivamente em um detector interno rodeado por um solenoide supercondutor fino, calorímetros eletromagnético e hadrônico, e um espectrômetro de múons [7].

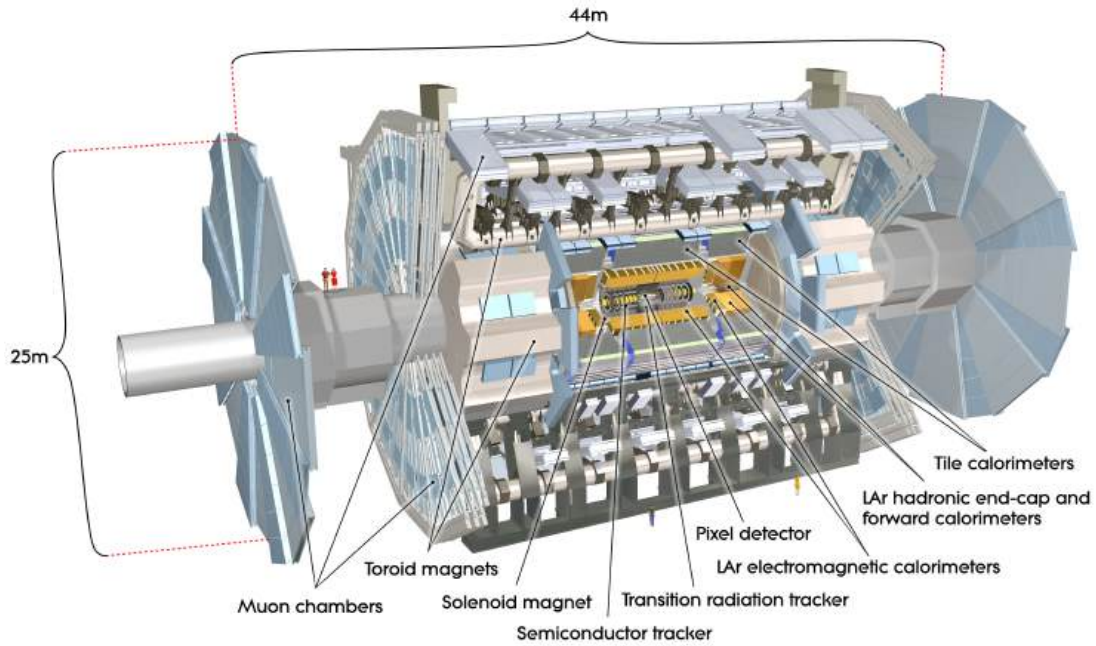


Figura 2.5: Representação gráfica do experimento ATLAS e seus subsistemas. Extraído de [7].

O ATLAS usa um sistema de coordenadas baseado na regra da mão direita, com sua origem no ponto de interação (PI) no centro do detector e no eixo z ao longo do feixe. A Figura 2.6 [46] apresenta uma descrição gráfica do sistema de coordenadas empregado pelo ATLAS. O eixo x aponta do PI para o centro do anel do LHC e o eixo y é perpendicular a este plano. Coordenadas cilíndricas (r, ϕ) são usadas no plano transversal, sendo ϕ o ângulo azimutal em torno do feixe, conforme a Equação 2.2:

$$\phi = \tan^{-1} \left(\frac{x}{y} \right). \quad (2.2)$$

A pseudo-rapidez (η) é definida em termos do ângulo polar θ (ângulo entre x e

y) como descrito pela Equação 2.3:

$$\eta = -\ln \left(\tan \left(\frac{\theta}{2} \right) \right). \quad (2.3)$$

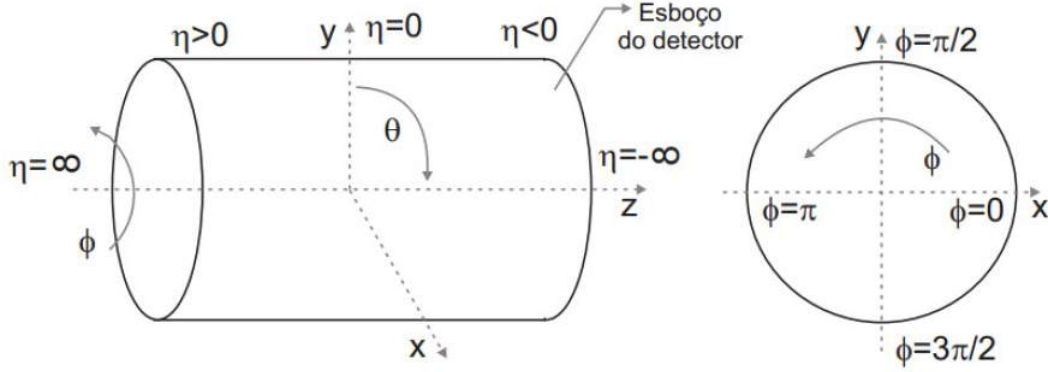


Figura 2.6: Sistema de coordenadas empregado pelo ATLAS. Extraído de [46].

Em experimentos que empregam colisão de feixes, sistemas de calorimetria em geral são utilizados para prover informação sobre a energia de partículas que interagem com o material do sistema, bem como descrever o perfil de deposição da energia em um tempo da ordem de nanossegundos [21]. A interação da partícula incidente com o calorímetro é caracterizada por um processo em cascata, em que comumente um número muito grande de partículas secundárias é produzido ao longo do detector.

2.3.1 Breve Introdução à Calorimetria

No campo da física de partículas, o termo calorimetria refere-se à detecção de partículas e medição de suas propriedades e interações através da absorção total destas, por um conjunto de materiais chamado calorímetro [21]. Em geral, o processo de medição empregado em um sistema de calorimetria é destrutivo e as partículas não estarão disponíveis após a sua passagem pelos calorímetros¹⁰.

Calorímetros podem ser classificados quanto a sua composição em homogêneos ou por amostragem [21]. Nos calorímetros homogêneos, todo o volume do seu material é sensível à partícula incidente. Os calorímetros por amostragem possuem um material passivo, responsável por absorver a energia da partícula incidente, e um material ativo, que produz o sinal que pode ser usado para a detecção. Em relação à precisão das medidas de energia efetuadas no calorímetros, a mesma aumenta com a energia

¹⁰Com exceção dos múons, que conseguem atravessar por uma elevada quantidade de matéria e para sua detecção é necessário um tipo de detector especial [47].

da partícula incidente, conforme a Equação 2.4:

$$\frac{\sigma_E}{E} \approx \frac{a}{\sqrt{E}} + b, \quad (2.4)$$

em que E é a energia da partícula incidente e σ_E representa o desvio padrão da medida de energia e a e b são constantes que dependem do tipo do material do detector e suas características geométricas das camadas ativas e passivas [21].

Como as características das interações dependem da energia e da natureza da partícula que incide no calorímetro, diferentes estratégias são desenvolvidas para medir a energia em função do tipo de interação. Um dos principais mecanismos de interação com a matéria é a interação eletromagnética por partículas carregadas que atravessam o material do calorímetro, e durante esta interação, uma cascata de novas partículas são geradas ao interagir com o material do calorímetro [21].

Em experimentos de física de altas energias, o *bremsstrahlung* é a principal forma de perda de energia de elétrons e pósitrons [21]. Um elétron com vários GeV de energia pode irradiar milhares de fótons ao cruzar pelo material passivo do calorímetro. A maioria destes fótons é absorvida por espalhamento *Compton* e efeito fotoelétrico. Fótons com energias maiores que 5-10 MeV criam pares elétrons-pósitrons [21], em que estes pares criados perdem sua energia liberando mais fótons, que podem produzir novos pares elétron-pósitrons. Esta sequência de produção de eventos é chamada de *chuveiro eletromagnético*, em que a quantidade de energia depositada no material absorvente aumenta à medida que o chuveiro se desenvolve [24]. No entanto, com o desenvolvimento do chuveiro, a energia média das partículas do chuveiro diminui, até o ponto em que nenhuma outra multiplicação venha a ocorrer. A partir dessa profundidade, os fótons do chuveiro estão, média, com maior probabilidade de produzir um elétron em suas interações, em vez de um par elétron-pósitron. Quanto aos elétrons e pósitrons, novamente em média, é mais provável que percam sua energia por meio da ionização do meio absorvente, em vez de produzir novos fótons (que por sua vez se converteriam em mais elétrons) através da radiação [21].

No caso da interação de partículas hadrônicas, devido à natureza da interação forte, cascatas produzidas por tais partículas são muito mais complicadas de descrever fisicamente do que os chuveiros eletromagnéticos. A variedade de processos que podem ocorrer é muito maior, ambos aqueles no nível da partícula e aqueles envolvendo o núcleo atingido. Quando um hádron de alta energia penetra um bloco de matéria, ele ioniza o átomos do meio atravessado, em um fluxo contínuo de eventos. Em algum ponto da interação o hádron pode encontrar um núcleo atômico com o qual interage. Nessa reação nuclear, o hádron pode, por exemplo, se transformar em muitos outros novos hádrons. Além disso, o núcleo atingido geralmente muda bastante em tais reações, podendo perder nêutrons e prótons no processo e acabam

em um estado altamente excitado, a partir do qual decai emitindo vários raios- γ . Píons e mésons podem decair através de interações eletromagnéticas ($\pi^0, \eta \rightarrow \gamma\gamma$), assim cascatas hadrônicas podem apresentar componente eletromagnética, que pode dificultar o processo de discriminação da partícula incidente [21].

As partículas produzidas na primeira reação nuclear (mésons, γ , etc) podem, por sua vez, perder sua energia ionizando o meio e/ou induzir novas reações, causando assim o surgimento de um *chuveiro hadrônico*. Inicialmente, o número de partículas de chuveiro aumenta como resultado de processos de multiplicação, assim como a energia depositada pelo chuveiro de partículas em um bloco de matéria de determinada espessura. No entanto, em alguma profundidade a multiplicação é equilibrada pela absorção das partículas de chuveiro. A partir desse chuveiro máximo, o número de partículas e a energia que elas depositam passa a diminuir gradualmente [21].

Apesar de algumas semelhanças, as diferenças do formato do chuveiro eletromagnético e hadrônico podem ajudar no processo de identificação da partícula que inicialmente incidiu no sistema de calorimetria. Chuveiros hadrônicos apresentam perfil lateral tipicamente mais largo do que aquele produzido por uma partícula eletromagnética. Outra característica do chuveiro hadrônico é que o mesmo se desenvolve muito mais profundamente do que um chuveiro eletromagnético [21]. As características laterais e longitudinais de chuveiros em sistemas de calorimetria, são geralmente exploradas para a identificação de partículas primárias.

2.3.2 Características Gerais do Experimento ATLAS

Em torno do Detector Interno está posicionado o Sistema de Calorimetria [7]. Este sistema é composto por calorímetros de amostragem baseados em argônio líquido (LAr - *Liquid Argon*) [21], com simetria e cobertura total em ângulo azimutal (ϕ), cobrindo a faixa de pseudo-rapidez $|\eta| < 4,9$. Ele está dividido em calorímetro eletromagnético (ECAL) e o calorímetro hadrônico (HCAL). A composição do sistema de calorimetria do ATLAS é apresentada na Figura 2.7 [7].

A região central do detector (barril) é dividida em duas metades ($-1,475 < \eta < 0$ e $0 < \eta < 1,475$) e dois componentes em sua extremidade (tampa) cobre a região de $1,375 < |\eta| < 3,2$. Na região de transição entre o barril e as tampas das extremidades ($1,37 < |\eta| < 1,54$) ocorre uma redução na quantidade de material ativo (que produz sinal), pois este espaço é reservado para passagem de cabos (*crack*), conexões e fibras ópticas por onde trafegam os sinais medidos nos sensores do detector. Na região $2,47 < |\eta| < 2,5$ existe uma junção dos dois blocos cilíndricos que compõem a tampa mais interna e externa do detector (barril estendido), compreendendo uma região com menor quantidade de sensores devido a presença de suportes mecânicos.

Na região dedicada às medições com elevado grau de precisão ($|\eta| < 2,5$) (ex-

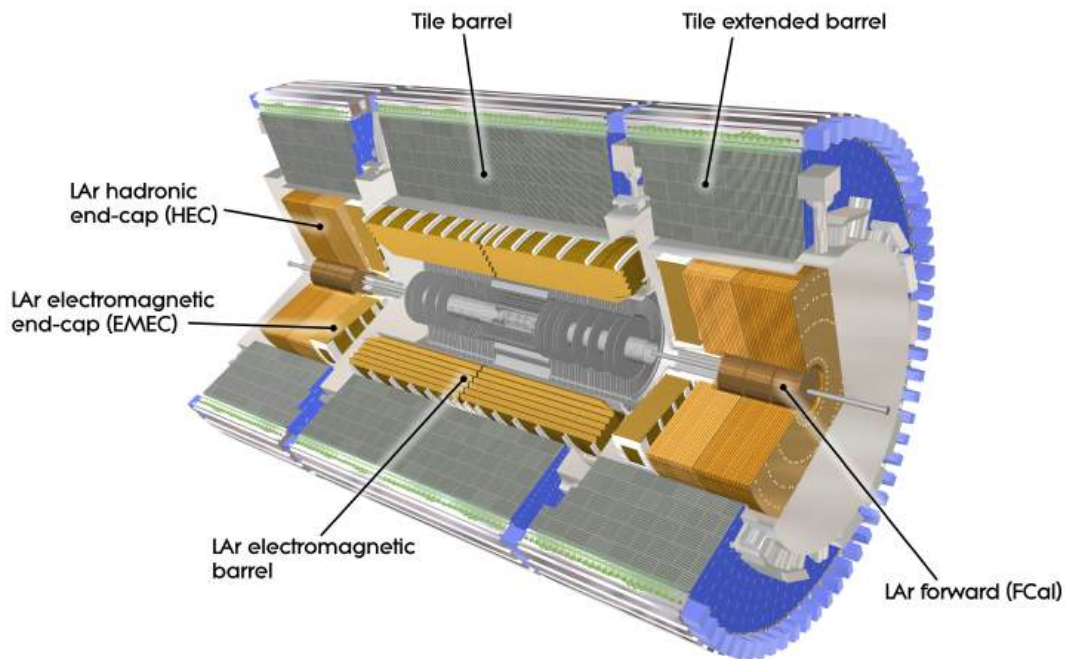


Figura 2.7: Esquemático da composição do sistema de calorimetria do ATLAS. Extraído de [7].

cluindo o *crack*), o ECAL é segmentado em três camadas longitudinais na direção da profundidade do chuveiro (EM1, EM2 e EM3). O ECAL é projetado de forma que a maior fração da energia seja coletada na camada EM2, enquanto a camada posterior coleta apenas a cauda do chuveiro eletromagnético. Três camadas de calorímetro hadrônico circundam o calorímetro ECAL, utilizadas para medir energia de partículas com natureza hadrônica (HAD0, HAD1 e HAD2). Uma camada fina de argônio líquido, chamada pré-amostrador (PS - *Pre Sampler*), cobre o intervalo de pseudo-rapidez $|\eta| < 1,8$, é usada para estimar a perda de energia das partículas no caminho até os calorímetros.

A Figura 2.8 [48] apresenta exemplos de interações ideais de partículas com os subsistemas do ATLAS. As partículas carregadas deixam traços no Detector Interno. Elétrons, pósitrons e fótons em uma interação ideal devem ser totalmente absorvidos pelo ECAL. O HCAL, é responsável por absorver partículas com componentes hadrônicas. Múons, mesmo contendo componentes eletromagnéticas, atravessam os calorímetros sem interagir com eles, assim o Espectrômetro de Múons é responsável por medir a energia destas partículas. Algumas partículas não sensibilizam o material ativo dos subsistemas (como por exemplo neutrinos) e atravessa o detector (representado pelas linhas tracejadas).

O ECAL é preenchido com meio ativo de LAr [49], com absorvedores compostos por eletrodos de cobre, enquanto o meio passivo é composto por placas de chumbo

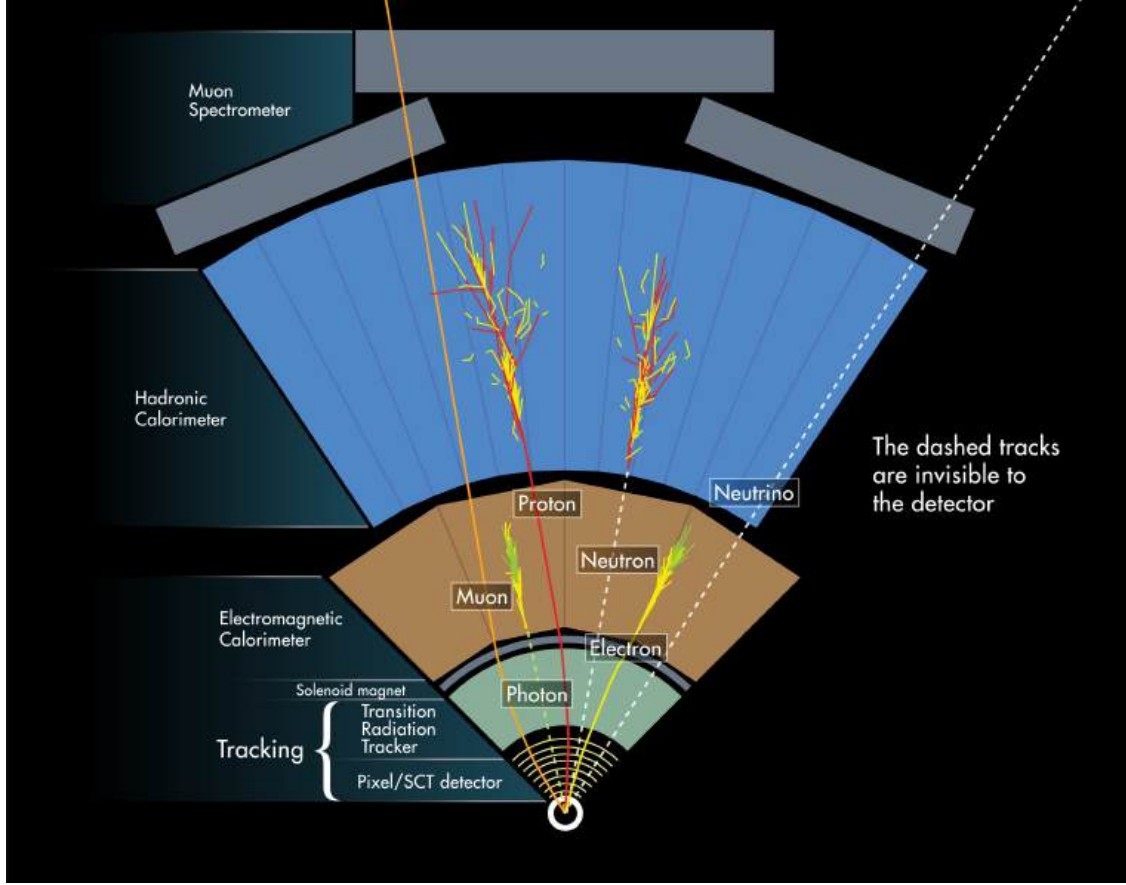


Figura 2.8: Corte transversal do ATLAS, mostrando os subsistemas e as interações ideais de algumas partículas. Extraído de [48].

usado como material absorvedor [7]. No ECAL, três camadas são utilizadas com distintas granularidades (concentração de células detectoras por unidade de área), que podem ser observadas na Figura 2.9 [49]. O ECAL apresenta uma incerteza na medição de energia, descrita pela Equação 2.5:

$$\frac{\delta E}{E} = \frac{10\%}{\sqrt{E}} \otimes 0,7\%, \quad (2.5)$$

ou seja, uma incerteza de 10% na medida quando a energia depositada no calorímetro for de 1 GeV, acrescido de um erro sistemático de 0,7%.

A primeira camada eletromagnética apresenta segmentação mais fina na direção η , possibilitando medição precisa do ponto de colisão ao longo desta coordenada. A fina segmentação desta camada é importante para a detecção de fótons, que não são medidos pelo detector de traços, e ao mesmo tempo para combinar os traços das partículas carregadas com seus respectivos chuveiros. A segunda camada apresenta células detectoras quadradas e uma maior profundidade, absorvendo maior parcela da energia. Na terceira camada, os detalhes finais do chuveiro eletromagnético são capturados.

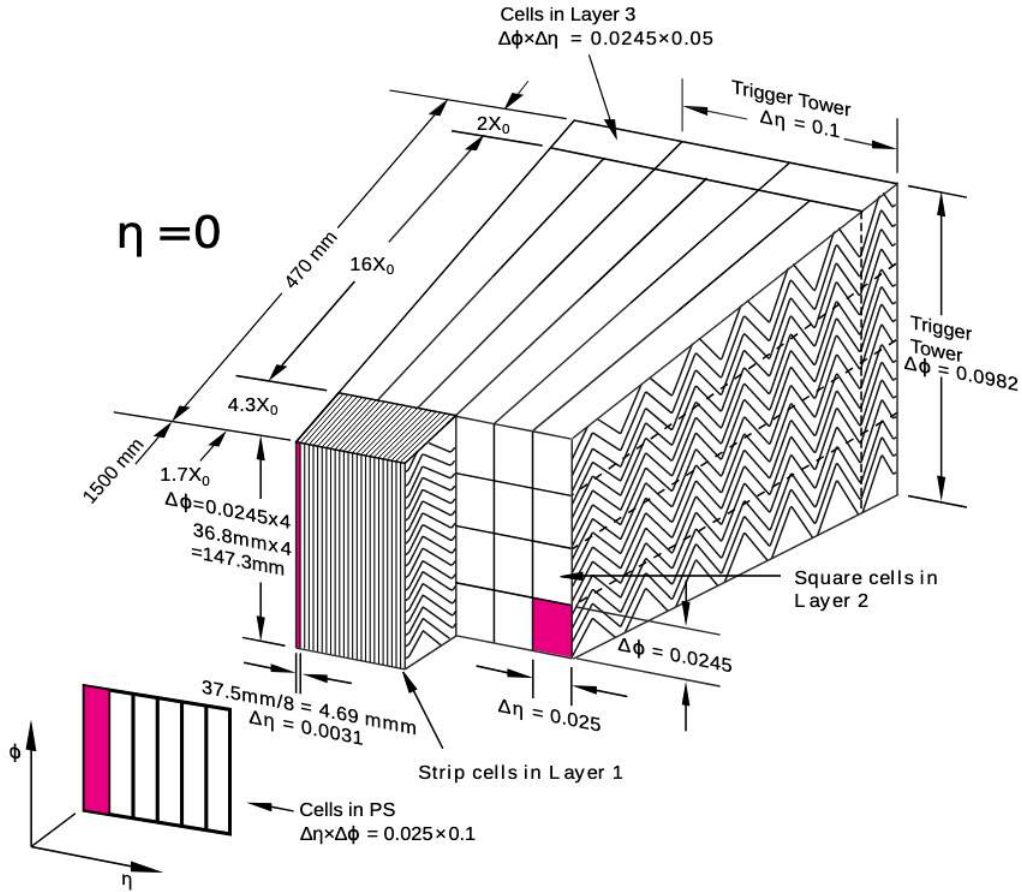


Figura 2.9: Representação gráfica da granularidade das camadas do barril do ECAL em $|\eta| = 0$ e um esboço da geometria de acordeão. Extraído de [49].

Os calorímetros eletromagnéticos são compostos por eletrodos de cobre-kapton em forma de acordeão, como indicado na Figura 2.10 [49], posicionados entre as placas absorvedoras de chumbo e mantidos em posição por espaçadores, enquanto o sistema está imerso em LAr. Usando a medição de energia e as posições de todas as células em todas as camadas do calorímetro contido no chuveiro, a energia da partícula incidente pode ser reconstruída e, aproveitando a segmentação fina da camada frontal, sua direção e características podem ser inferidas. Na Tabela 2.2 são apresentadas as informações da granularidade das células, em cada camada do sistema de calorimetria [7].

Na parte mais externa em relação ao eixo dos feixes do LHC, o Calorímetro Dianteiro (FCal - *Forward Calorimeter*) [50] fornece cobertura acima de 3, $1 < |\eta| < 4$, 9. Dada a proximidade com o feixe de partículas do LHC, esse calorímetro é especialmente projetado para suportar altas taxas de radiação e um elevado fluxo de partículas. Três módulos cilíndricos compõem o FCal, dispostos sequencialmente; o módulo mais próximo do ponto de interação (FCal1) é otimizado para medições eletromagnéticas e usa cobre como absorvedor, enquanto os dois módulos subse-

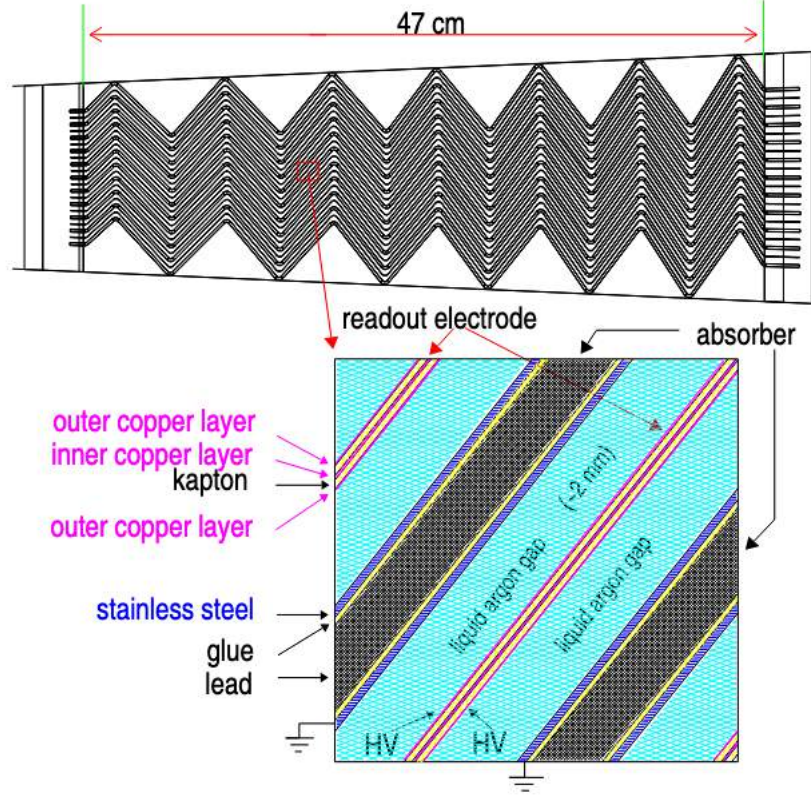


Figura 2.10: Estrutura de acordeão do barril eletromagnético. A figura na parte superior é uma vista de um pequeno setor do calorímetro em um plano transversal aos feixes do LHC. Extraído de [49].

quentes (FCal2 e FCal3) são feitos principalmente de tungstênio e são otimizados para medições hadrônicas.

Em torno do ECAL está posicionado o Calorímetro Hadrônico de Telhas (TileCal), responsável por medir a energia de hádrons, tais como píons, prótons, etc. Ele é dividido em duas partes, uma na região das tampas, e outra na região do barril central [51]. O TileCal é um calorímetro de amostragem que utiliza aço como material absorvedor e telhas cintiladoras como material ativo, cujos sinais são lidos e coletados por fibras óticas. Este calorímetro fornece dados essenciais para a identificação de jatos hadrônicos, por meio da medição de sua energia e direção.

Conforme visto na Figura 2.7, o TileCal é dividido em três segmentos ao longo do eixo dos feixes. Uma seção central (LB - *Long Barrel*) com 5,8 m de comprimento ($\eta < 1,0$), e duas seções estendidas (barril estendido) (EB - *Extend Barrel*) ($0,8 < |\eta| < 1,7$) em cada lado do barril, cada uma com 2,6 m de comprimento. Dois cilindros limitam o volume, em que o interno tem um raio igual a 2,28 m e o externo possui raio de 4,23 m. Longitudinalmente, tem três camadas radiais com profundidades de 1,5, 4,1 e $1,8\lambda$ ¹¹ respectivamente para o LB e 1,5, 2,6 e $3,3\lambda$ para o

¹¹O comprimento de interação nuclear (λ) é definido como o comprimento médio do caminho

Tabela 2.2: Região de cobertura em η , granularidade das células (ou *hits*) e camadas utilizadas no ATLAS. Extraído de [7].

Camada	Amostrador		Cobertura	Granularidade ($\Delta\eta \times \Delta\phi$)
Pré-amostrador	Barril	PSB	$0,00 < \eta < 1,58$	$0,025 \times 0,1$
	Tampa	PSE	$1,50 < \eta < 1,80$	$0,025 \times 0,1$
Calorímetro Eletromagnético				
Camada 1	Barril	EMB1	$0,00 < \eta < 1,55$	$0,003 \times 0,1$
	Tampa	EMEC1	$1,37 < \eta < 1,80$	$0,003 \times 0,1$
			$1,80 < \eta < 2,00$	$0,025 \times 0,1$
			$2,00 < \eta < 2,37$	$0,006 \times 0,1$
			$2,37 < \eta < 3,20$	$0,1 \times 0,1$
Camada 2	Barril	EMB2	$0,00 < \eta < 1,50$	$0,025 \times 0,025$
	Tampa	EMEC2	$1,35 < \eta < 2,50$	$0,025 \times 0,025$
			$2,50 < \eta < 3,20$	$0,1 \times 0,1$
Camada 3	Barril	EMB3	$0,00 < \eta < 1,58$	$0,05 \times 0,1$
	Tampa	EMEC3	$1,35 < \eta < 2,50$	$0,05 \times 0,025$
			$2,50 < \eta < 3,20$	$0,1 \times 0,1$
Calorímetro Hadrônico				
Camada 1	Barril	TileCal1	$0,00 < \eta < 1,09$	$0,1 \times 0,1$
	Barril Extendido	TileExt1	$0,94 < \eta < 1,77$	$0,1 \times 0,1$
	Tampa	HEC1	$1,50 < \eta < 2,50$	$0,1 \times 0,1$
			$2,50 < \eta < 3,20$	$0,2 \times 0,2$
Camada 2	Barril	TileCal2	$0,00 < \eta < 1,09$	$0,1 \times 0,1$
	Barril Extendido	TileExt2	$0,85 < \eta < 1,41$	$0,1 \times 0,1$
	Tampa	HEC2	$1,50 < \eta < 2,50$	$0,1 \times 0,1$
			$2,50 < \eta < 3,20$	$0,2 \times 0,2$
Camada 3	Barril	TileCal3	$0,85 < \eta < 0,72$	$0,2 \times 0,1$
	Barril Extendido	TileExt3	$0,85 < \eta < 1,41$	$0,2 \times 0,1$
	Tampa	HEC3	$1,50 < \eta < 2,50$	$0,1 \times 0,1$
			$2,50 < \eta < 3,20$	$0,2 \times 0,2$

EB. Estas três camadas de células usadas para amostrar o chuveiro em três diferentes profundidades, chamadas A, BC e D têm uma granularidade $\Delta\eta \times \Delta\phi = 0,1 \times 0,1$ nas duas camadas mais internas e $\Delta\eta \times \Delta\phi = 0,2 \times 0,1$ na camada mais externa. Cada célula representa um conjunto de telhas cintiladoras agrupadas de forma a para reduzir o fluxo de hádrons primários relativísticos a uma fração $1/E$ [21].

medir, em conjunto, a energia de uma partícula que passe por estas células [52]. O TileCal é projetado para ter uma resolução de energia para jatos hadrônicos, descrita pela Equação 2.6:

$$\frac{\delta E}{E} \approx \frac{50\%}{\sqrt{E}} \otimes 3\%. \quad (2.6)$$

O responsável por amostrar a energia das partículas hadrônicas que atravessam o TileCal é o cintilador. O cintilador é um material de poliestireno em formato de telha (*Tile*) dopado com material cintilante que, ao interagir com uma partícula, transforma a energia depositada pela partícula em luz produzida através da cintilação do seu material. A luz produzida nas telhas cintiladoras está na faixa do Ultra-Violeta (UV) e sua intensidade é proporcional à energia depositada pela partícula [52]. A disposição de um módulo do TileCal é exibida na Figura 2.11 [52].

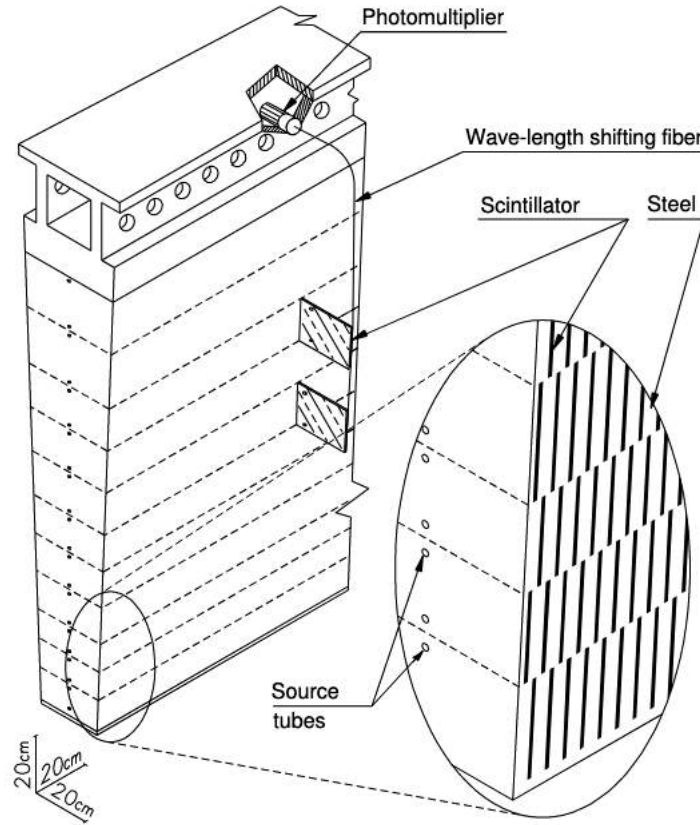


Figura 2.11: Representação gráfica de um módulo do TileCal. Extraído de [52].

A luz propagada por reflexões dentro da telha para suas bordas é então transmitida através de uma fibra óptica até um fotomultiplicador (*Photo Multiplier Tube* - PMT) que, então, transforma esse feixe de luz em um sinal elétrico. As fibras ópticas estão acopladas às duas laterais de cada telha cintiladora, e essas, por sua vez, são conectadas a dois fotomultiplicadores que leem a mesma célula, para obter redundância no sinal coletado.

Os sinais elétricos captados e transmitidos pelas conexões ópticas são enviados para o sistema de filtragem, onde sinais de interesse serão então separados dos sinais não relevantes para o experimento. O TileCal conta com um sistema eletrônico de leitura, responsável por formatar, amplificar e digitalizar os sinais provenientes do sistema óptico de leitura (telhas cintiladoras e fotomultiplicadores).

2.4 Seleção *Online* de Elétrons no ATLAS

Elétrons são partículas que estão relacionadas a diversos canais de interesse do experimento ATLAS, nos quais podem ser tão raros quanto o bóson de Higgs, que foi detectado experimentalmente em 2012 [15]. A detecção de elétrons também é importante para aspectos operacionais do experimento, como alinhamento e calibração dos sensores do detector, durante o início das operações de tomada de dados. Considerando padrões conhecidos de deposição de energia de eventos (como o decaimento $Z \rightarrow ee$), informações do comportamento do detector são monitoradas e caso algum problema seja detectado na qualidade dos dados obtidos, ações de reparo e ajustes podem ser realizadas.

Muitas das colisões ocorridas no LHC podem não ser capazes de gerar os eventos físicos de interesse pelo ATLAS, seja por limitação técnica ou pela raridade de ocorrência do fenômeno que se deseja observar. No LHC, para energia de 40 GeV, a relação de produção elétron/jato é aproximadamente 10^{-5} , o que torna difícil a detecção da física de interesse [7]. Neste contexto, o aumento da luminosidade e da energia envolvida na colisão pode então aumentar a probabilidade de ocorrência do evento de interesse, porém também aumenta a quantidade de processos físicos não relevantes para o experimento (ruído de fundo). A utilização de um sistema de filtragem *online* de eventos torna-se essencial para descartar processos que não agregam importância, que poderiam dificultar análises de eventos importantes em ambiente *offline* [12], e preservar os eventos que são de interesse para o experimento.

Em experimentos baseados em calorimetria, em geral, utiliza-se as diferenças entre o desenvolvimento do chuveiro eletromagnético e do chuveiro hadrônico observadas nos perfis laterais e longitudinais de deposição de energia para realizar a separação elétron/Jato [24]. O ATLAS é dotado de um sistema de filtragem e aquisição de dados (TDAQ - *Trigger and Data Acquisition*) [53], responsável por decidir se deve registrar dados de uma determinada colisão. Este sistema, que é dividido

¹²O ambiente *offline* está relacionado com a infraestrutura de análise, seleção e reconstrução de eventos do ATLAS que utilizam os dados gravados pelo experimento. Neste ambiente, as assinaturas dos eventos já estão armazenadas em mídia permanente, consequentemente não ocorrem as extremas limitações temporais de análise, se comparado ao ambiente *online* permitindo que algoritmos mais sofisticados de análise possam ser empregados.

em dois níveis, é encarregado pelo processamento *online*, seleção e armazenamento de eventos de interesse para análise *offline*. Para realizar a filtragem *online* de elétrons, é estimado um conjunto de grandezas que descrevem o perfil de deposição de energia da partícula nos calorímetros, como também são consideradas informações do detector interno (traços).

2.4.1 Primeiro Nível de Filtragem

O primeiro nível de filtragem (L1) é um sistema implementado em *hardware*, que usa sistemas eletrônicos dedicados para selecionar informações provenientes de regiões do calorímetro (*L1Calo*) e dos detectores de múons (*L1Muon*). O L1 reduz a taxa de eventos de 40 MHz para abaixo de 100 kHz, dentro de uma latência de $2,5 \mu\text{s}$. Além de realizar a primeira etapa de seleção, o L1 identifica as regiões de interesse (RoIs - *Region of Interest*) em coordenadas η e ϕ do calorímetro, que são então analisadas pelo segundo nível de filtragem caso o evento processado pelo L1 seja aprovado. O diagrama apresentado na Figura 2.12 [54] apresenta o Sistema TDAQ utilizado na Run2 pelo ATLAS.

No *L1Calo*, os sinais analógicos medidos no detector são digitalizados e calibrados pelo módulo pré-processador (*Preprocessor*) e enviados, em paralelo, para os módulos CP (*Cluster Processor*) e JEP (*Jet/Energy-sum*). O CP é dedicado para identificar candidatos a elétrons, fótons e τ -leptons acima de um determinado limite de energia, e o sistema JEP identifica candidatos a jatos e produz somas globais de energia transversa total (E_T) e energia transversa perdida (E_T^{miss}). A tomada de decisão do L1 é realizada pelo módulo CTP (*Central Trigger Processor*) e para cada evento aceito pelo L1, a eletrônica de *Front-End* realiza a leitura dos dados do evento em todos os detectores. Os dados são enviados primeiro para *ReadOut Drivers* (RODs), que realiza o processamento inicial e em seguida, encaminha as informações para o *ReadOut System* (ROS) para armazenar os dados em *buffer*. Os dados dos diferentes sub-detectores são enviados do ROS para o segundo nível de filtragem, o *trigger* de alto nível (HLT - *High-Level Trigger*).

O *L1Calo* executa um algoritmo de janela deslizante para buscar as células com maior nível de deposição de energia (RoI). Esta busca é realizada na região central do barril ($|\eta| < 2,47$), usando uma janela de 2×2 Torres de Filtragem (TT - do inglês *Trigger Tower*) no ECAL, onde cada TT mede uma região de $0,1 \times 0,1$ no plano $\eta \times \phi$. Após a definição da RoI, o *L1Calo* utiliza critérios de aceitação que podem ser baseados em limiares¹³ de E_T , variáveis em função de η , ou de isolamento

¹³O limiar de decisão em função de E_T em geral dá nome a cadeia de *trigger* (e.g. um limiar $E_T > 22 \text{ GeV}$ dá nome a cadeia EM22). Tipicamente, são aceitas variações do limiar de decisão entre -2 a 3 GeV relativa ao valor nominal do limiar. Esta tolerância é utilizada para compensação

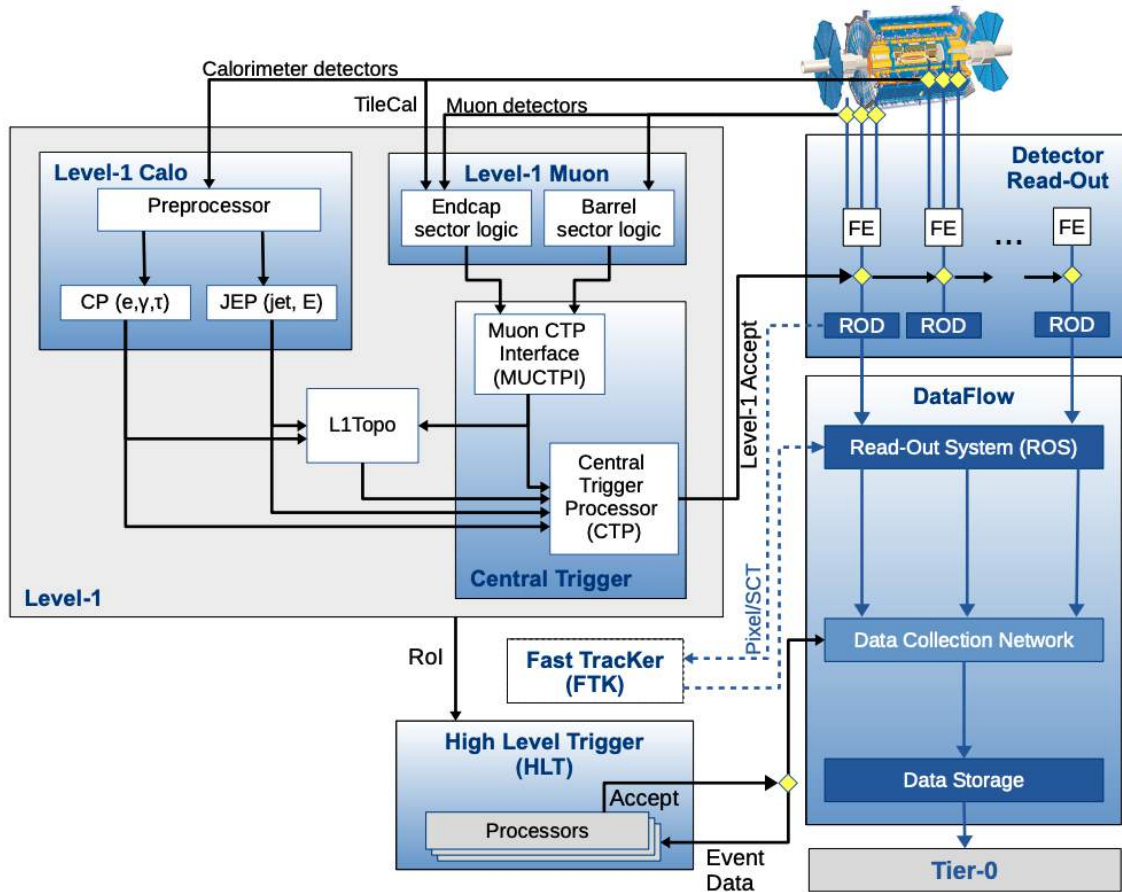


Figura 2.12: Sistema ATLAS TDAQ na *Run2* mostrando os componentes relevantes para o acionamento, bem como a leitura do detector e o fluxo de dados. Extraído de [54].

de energia.

A Figura 2.13 apresenta um esboço das TT utilizadas no algoritmo de janela deslizante. Candidatos a elétrons são rejeitados caso a soma da E_T em torres hadrônicas, combinadas com a região central 2×2 for de pelo menos 1 GeV. Dois critérios de isolamento são utilizados, o isolamento eletromagnético que compreende a energia depositada na região central eletromagnética, computada em 4×4 TT no ECAL. O outro critério é o isolamento hadrônico que é similar ao isolamento eletromagnético, porém utiliza as torres equivalentes no HCAL [55].

2.4.2 Segundo Nível de Filtragem

As informações das RoIs aceitas pelo *L1Calo* são encaminhadas para o segundo nível de filtragem, chamado de Filtragem de Alto Nível (HLT - do inglês *High Level Trigger*). O HLT é baseado em *software*, e é composto por algoritmos que avaliam a de variações de calibração da energia mediada nas células.

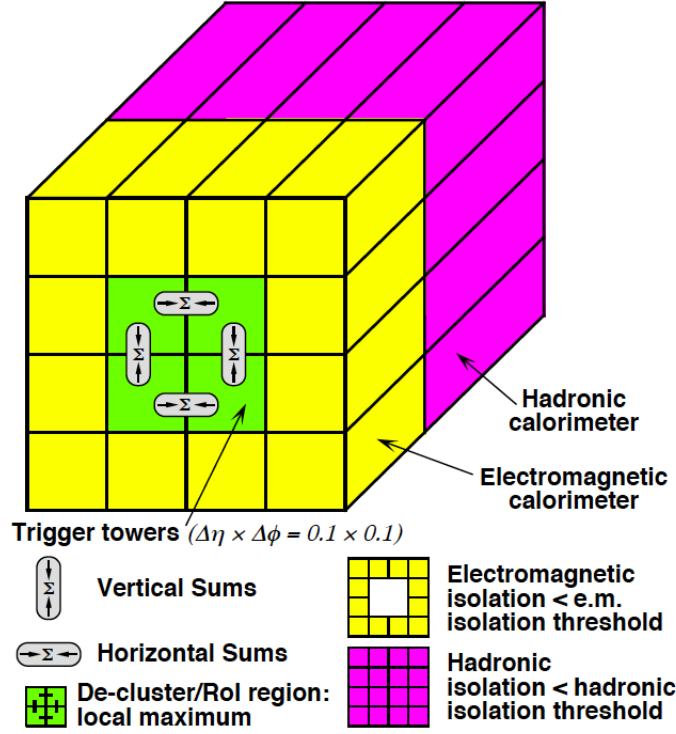


Figura 2.13: Esboço de um conjunto de células do calorímetro que o algoritmo de janela deslizante aplicado no *L1Calo* para realizar a seleção de elétrons. A região central em verde representa a *RoI*, a região de isolamento EM é marcada em amarelo e a região central e de isolamento HAD é marcada em roxo. O valor de energia das TT na região central é somada para fornecer a energia local. Extraído de [55].

RoI e também uma maior quantidade de informação do perfil de deposição de energia nos calorímetros. Neste nível, são disponibilizadas informações do calorímetro com fina granularidade, medições precisas do espectrômetro de múon e informações do ID, que não estão disponíveis no L1. Os algoritmos do HLT são executados em uma fazenda computacional dedicada, de aproximadamente 40.000 aplicativos distribuídos conhecidos como Unidades de Processamento (PU - *Processing Units*). A seleção das partículas candidatas pelo HLT é realizada sequencialmente, de modo que se o evento for reprovado em determinada etapa, as etapas subsequentes não são executadas.

Conforme o fluxo operacional apresentado na Figura 2.14, o HLT é subdividido em dois níveis de processamento sequenciais tendo a primeira etapa composta por algoritmos rápidos (*Fast*), e a segunda etapa formada por algoritmos de análises mais sofisticadas (*Precision*) [18]. Uma sequência de reconstrução típica (cadeia de filtragem)¹⁴ faz uso em um estágio inicial de algoritmos rápidos de *trigger* dedicados

¹⁴O termo *cadeia de filtragem* se refere a sequência de algoritmos de filtragem com configurações específicas relacionadas ao limiar de decisão de energia nos níveis de *trigger*, física de interesse, critérios de isolamento de energia, etc.

para fornecer uma primeira rejeição de ruído de fundo e em seguida, de acordo com a redução da taxa de eventos, algoritmos com maior complexidade são adotados para realizar a seleção final dos eventos que serão armazenados.

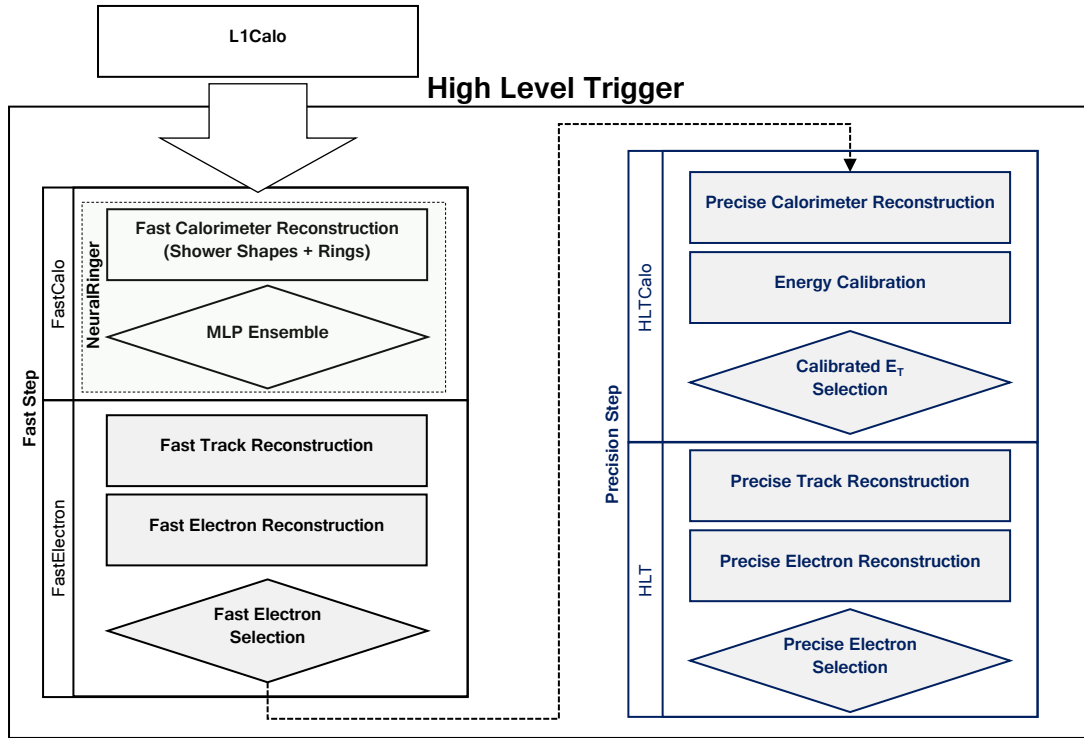


Figura 2.14: Sequências de *trigger* simplificadas de elétrons para obtenção de dados de colisão próton-próton durante a *Run2*. Adaptado de [18].

Na subetapa chamada etapa rápida de calorimetria (*FastCalo*), os algoritmos rápidos do HLT realizam a leitura de *clusters* de células do calorímetro com as RoIs identificadas pelo L1. Um *cluster* é construído a partir da identificação da célula com maior quantidade de energia depositada (região de interesse) e com a sua posição definida, são mapeadas as células ao seu redor que então formarão o *cluster*. É utilizada uma janela de 3×7 ($\Delta\eta \times \Delta\phi = 0,075 \times 0,175$), onde a posição final do *cluster* é obtida calculando a média ponderada da energia das células em função da sua posição na janela de 3×7 . Na região da tampa, o tamanho do *cluster* utilizado é de 5×5 . A informação do *cluster* é utilizada na etapa rápida pelos algoritmos sequenciais de extração de características e teste de hipóteses.

Durante a tomada de dados nos anos 2017 - 2018, dois métodos foram utilizados na etapa *FastCalo* para classificação de elétrons, no qual o *CutBased* foi o método adotado para RoIs com $E_T < 15 \text{ GeV}$. Para RoIs com $E_T \geq 15 \text{ GeV}$ o HLT emprega o algoritmo *NeuralRinger*, que utiliza a informação dos *clusters* para medir a energia nas células presentes em anéis concêntricos em torno da RoI. Ao todo, 100 anéis são gerados ao longo das camadas do ECAL e HCAL, onde as somas de energia medida

em cada anel são sinais de entrada para classificadores neurais executarem a tomada de decisão [15].

Após o aceite dos eventos pela sub-etapa *FastCalo*, os eventos são processados pela próxima sub-etapa, então chamada de etapa rápida para elétrons (*FastElectron*), na qual algoritmos para reconstrução rápida de traços e reconstrução rápida de variáveis discriminantes de calorimetria são executados. A reconstrução rápida de traço recupera as informações do ID para realizar o casamento entre o traço deixado pelo elétron com sua respectiva RoI. Na reconstrução rápida de elétrons, é realizada a extração de atributos discriminantes relacionados com o perfil lateral e longitudinal de deposição de energia (*shower shapes*). Nesta etapa uma seleção é baseada em limiares de E_T do *cluster* e nas variáveis R_{had} , R_η e E_{ratio} , que a partir da combinação de critérios fundamentados no conhecimento de física teórica, é realizada a decisão para aprovação do evento [16].

Os eventos aprovados pela etapa rápida são processados pela etapa precisa, que é dotada de algoritmos de extração de características, utilizados para a reconstrução das variáveis de calorimetria. Neste nível, uma maior quantidade de informação dos calorímetros é disponibilizada para realizar uma reconstrução mais precisa da energia dos *clusters* e das variáveis que descrevem o formato do chuveiro [18]. Essas informações são utilizadas para calibração de energia que recalcula com uma maior precisão a energia do *cluster* e em seguida, realiza uma seleção dos eventos baseada em limiares de energia. A seleção final de elétrons no HLT é realizada usando um discriminante multivariado de verossimilhança (LH - do inglês *Likelihood*) com quatro pontos de operação: "*lhvloose*", "*lhloose*", "*lhmedium*", e "*lhtight*". Estes pontos de operação consistem em limiares de decisão em que, no critério mais brando (*lhvloose*) é permitido a aprovação de uma quantidade elevada de eventos, porém ocorre uma maior aceitação de ruído de fundo. No critério mais restritivo (*lhtight*) uma menor quantidade de eventos é aprovada, porém há uma maior pureza [17] na amostra aceita, devido à uma quantidade menor de ruído de fundo ser aceita por este critério. Os algoritmos de reconstrução da etapa precisa são fundamentados no algoritmos de reconstrução e análise do *offline* e operam com algumas restrições para atendimento da latência exigida pelo processamento *online* [17].

A configuração do *trigger* é controlada pelo *menu de trigger*, que define uma lista

¹⁵Em 2017, o *NeuralRinger* substituiu o método *CutBased* para $E_T > 15$ GeV [18], pois com o aumento da luminosidade o algoritmo *NeuralRinger* apresentou melhor desempenho na classificação de elétrons e rejeição a falsos elétrons. Desde então, a colaboração ATLAS moveu esforço para também empregar o *NeuralRinger* em faixas de $E_T < 15$ GeV.

¹⁶ R_{had} é a razão entre a energia transversa do *cluster* no HCAL e no ECAL. R_η é definida como a razão da E_T em uma região central de 3×7 células em $\eta \times \phi$ em relação a energia medida na região 7×7 a partir do centro do *cluster*.

¹⁷Eventos com maior probabilidade de pertencerem à classe de sinal.

completa de cadeias L1 e do HLT, bem como as suas configurações. A composição do menu e os limiares de disparo são otimizados para as condições de operação do LHC (tipo de feixe, luminosidade, etc.) para se ajustar à taxa de aceitação do evento e às restrições de largura de banda do sistema de aquisição de dados do detector ATLAS, bem como as restrições de armazenamento *offline*. Uma vez que um evento é aceito por uma cadeia do HLT, os dados são enviados para armazenamento permanente para os algoritmos do ambiente *offline* reconstruir e exportar os dados para o Tier-0¹⁸ do CERN.

2.5 Algoritmo *NeuralRinger*

O *NeuralRinger* é um algoritmo que combina extração de características e teste de hipótese para a discriminação de elétrons, atualmente empregado na etapa rápida do HLT [18]. Este algoritmo explora a propriedade dos chuveiros de partículas desenvolver-se na direção lateral e em uma estrutura aproximadamente cônica longitudinalmente em torno da partícula inicial. Este recurso permite que as informações relevantes do chuveiro sejam codificadas em quantidades que descrevem somas de energia de todas as células pertencentes a um anel concêntrico, gerados em cada camada do calorímetro. Os anéis gerados são retangulares devido à estrutura das células do calorímetro, conforme ilustrado na Figura 2.15 [18].

Inicialmente, o algoritmo procura pela célula mais energética da segunda camada eletromagnética, compondo com esta célula o primeiro anel. A energia das células vizinhas ao primeiro anel são somadas, formando a energia do segundo anel. Esse procedimento é repetido até que todas as células da região de interesse sejam utilizadas (ou um número máximo de anéis seja alcançado). No ECAL, os anéis são centralizados em torno da célula mais energética de cada camada, enquanto no HCAL a posição da célula mais energética na segunda camada EM é utilizada como eixo. No total, cem anéis são definidos no total para cada uma RoI aprovada pelo *L1Calo*. Conforme indicado na Tabela 2.3, são gerados 8 anéis para as camadas PS, EM2 e EM3. Na camada EM1 64 anéis são gerados e 4 anéis são gerados em cada uma das três camadas do calorímetro hadrônico [18].

A soma da energia presente em cada anel (Er_k) é concatenada em um único

¹⁸O CERN possui um largo centro de computação para dar suporte aos seus experimentos. O WLCG (*Worldwide LHC Computing Grid*) possui cerca de 900.000 núcleos de computadores de mais de 170 locais em 42 países, provendo uma infraestrutura de computação distribuída que fornece às colaborações em todo o mundo acesso aos dados do LHC. O WLCG é composto por quatro níveis, ou “Tiers”, chamados 0, 1, 2 e 3. O Tier-0 é o *Data Center* do CERN, em que todos os dados do LHC passam por este *hub* central. O Tier-0 distribui os dados brutos e a saída reconstruída para o Tier-1, que reprocessa os dados quando o LHC não está em execução.

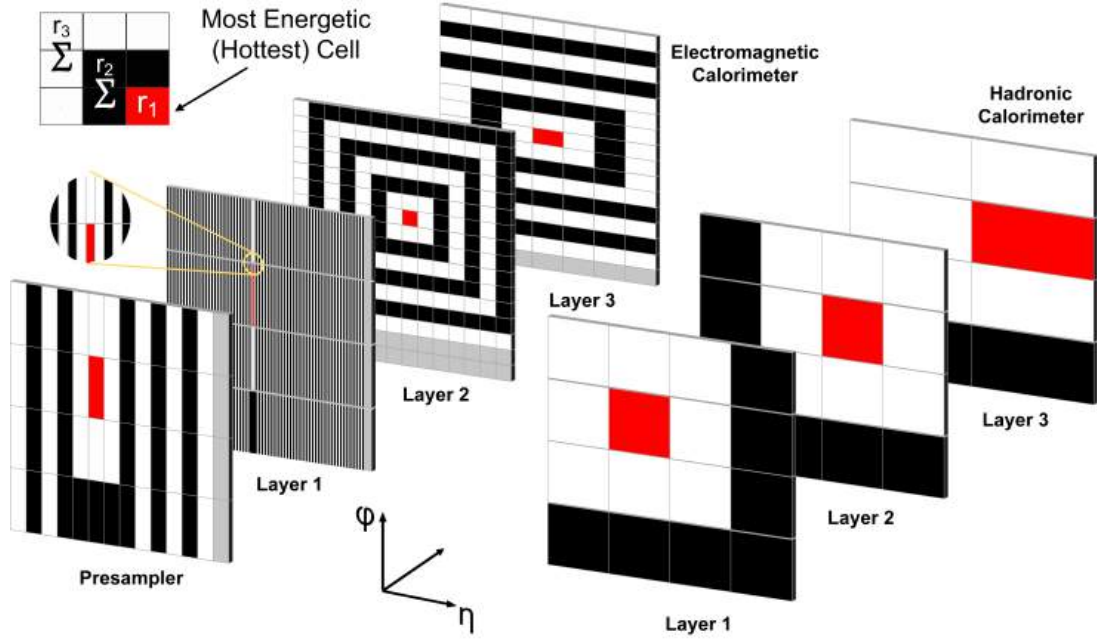


Figura 2.15: Ilustração de diferentes configurações de anéis para diversas camadas dos calorímetros do ATLAS. A célula mais energética (r_1) é mostrada em vermelho, enquanto os anéis ao seu redor alternam entre branco e preto. Extraído de [18].

Tabela 2.3: Número de anéis concêntricos gerados em cada camada do calorímetro.

Camadas	PS	EM1	EM2	EM3	HAD0	HAD1	HAD2
Anéis	8	64	8	8	4	4	4

vetor composto por 100 atributos. Uma normalização usando a energia total da RoI é aplicada para ajustar as variáveis, conforme indica a Equação 2.7. O valor absoluto é usado para evitar refletir os valores ao longo do eixo devido ao acúmulo de ruído negativo [22].

$$Er'_k = \frac{Er_k}{|\sum_{i=1}^{100} Er_i|}, \quad \forall k \in \{1, 2, \dots, 100\}. \quad (2.7)$$

A informação do chuveiro formatada em anéis é utilizada como entrada para um conjunto de redes neurais artificiais do tipo MLP (*Multilayer Perceptron*), treinadas por faixas definidas em $|\eta| \times E_T$ conforme indicadas na Tabela 2.4, para realizar a tomada de decisão se a RoI é ou não candidato a elétron. A principal motivação para o uso de um conjunto de redes neurais treinadas em distintas faixas de $|\eta|$ e E_T (espaços de fase) está relacionada com diferenças no sinal medido, em função da granularidade dos sensores e resolução de energia. Assim, a utilização de classificadores específicos por região de espaço de fase permite acomodar variações nos perfis de deposição variáveis devido à resposta do calorímetro [10].

Tabela 2.4: Faixas de E_T e η definidas para a operação do conjunto de redes neurais do *NeuralRinger* no HLT do ATLAS.

E_{T0}	E_{T1}	E_{T2}	E_{T3}	E_{T4}
$15 \leq E_T[\text{GeV}] < 20$	$20 \leq E_T[\text{GeV}] < 30$	$30 \leq E_T[\text{GeV}] < 40$	$40 \leq E_T[\text{GeV}] < 50$	$E_T \geq 50[\text{GeV}]$
η_0	η_1	η_2	η_3	η_4
$0 \leq \eta < 0,80$	$0,80 \leq \eta < 1,37$	$1,37 \leq \eta < 1,54$	$1,54 \leq \eta < 2,37$	$2,37 \leq \eta < 2,50$

O *NeuralRinger* aumentou o tempo gasto pela etapa rápida de calorimetria para ≈ 2 ms por evento, sendo aproximadamente 45% mais lento do que o *CutBased*. No entanto, ele possibilitou a redução por um fator de 1,5 a 6 vezes do número de candidatos de entrada para a etapa de rápida de reconstrução de traços, (*Fast Track Reconstruction*), que é mais exigente de CPU (que leva cerca de 64 ms por evento).

2.5.1 Breve Histórico de Contribuições ao *NeuralRinger*

Desde a concepção, o método *NeuralRinger* veio sendo aprimorado por pesquisadores, visando adequá-lo para as atualizações de condições de funcionamento do experimento ATLAS. O *NeuralRinger* explora o desenvolvimento do chuveiro de partículas em uma estrutura aproximadamente cônica, onde captura o espalhamento lateral e longitudinal do chuveiro, ao longo das camadas do sistema de calorimetria do ATLAS. Capturar a informação lateral do espalhamento do chuveiro tem motivação em estudos anteriores no LHC, que destacam a utilização de medidas da largura efetiva do chuveiro para fins de discriminação de partículas. No trabalho [20], a largura efetiva unidimensional do chuveiro (R_p) foi definida para medir a largura do chuveiro de partículas no experimento SPACAL, dada pela Eq. 2.8:

$$R_p = \frac{\sum_{i=1}^C E_i^\alpha r_i^\beta}{\sum_{i=1}^C E_i^\alpha}, \quad (2.8)$$

onde E_i representa a energia depositada em uma célula individual i , r_i é a distância euclidiana entre o centro desta célula e a célula de energia máxima depositada, α e β são parâmetros que podem ser ajustados e C o número total de células. Em [20], conforme indicado pela Figura 2.16, a variável R_p foi utilizada para a discriminação de elétrons e píons provendo boa separação entre as distribuições.

A proposta do mapeamento do perfil de deposição de energia em anéis concêntricos como sinais de entrada em redes neurais para classificação de elétrons foi inicialmente apresentada no trabalho [22]. Neste trabalho, a descrição do perfil de deposição de energia em anéis, alimentando uma rede neural completamente conectada encontrou 97% de eficiência para classificação de elétrons e 10,5% de erro de classificação de jatos hadrônicos em eventos sem *pile-up*. Também em [22], a

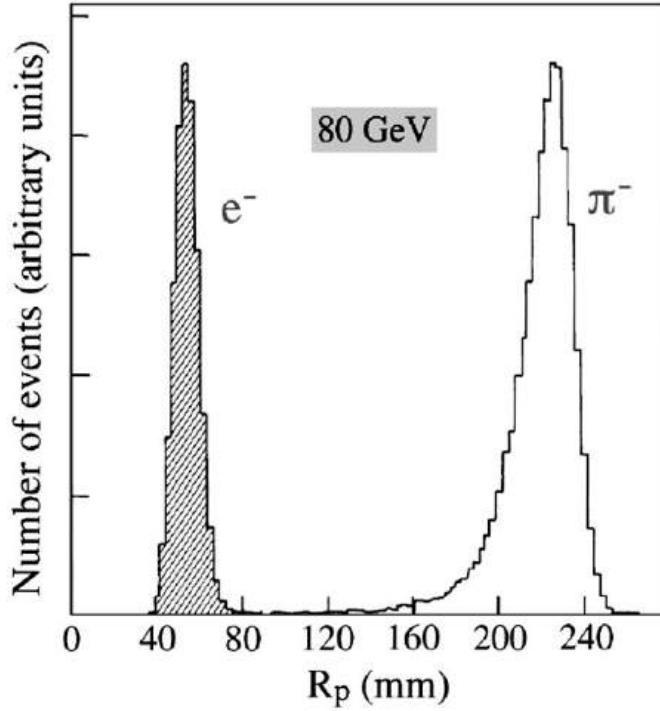


Figura 2.16: Distribuição da largura efetiva do chuveiro (R_p) unidimensional, para chuveiros produzidos por elétrons e píons, ambos em 80 GeV, avaliados no experimento SPACAL. Extraído de [20].

proposta do uso da largura do chuveiro [20] como variável discriminante combinada com anéis, para discriminação elétron/jato foi avaliada. Este método ofereceu a vantagem de combinar a compactação¹⁹ que a estrutura em anéis permite com o mapeamento discriminante R_p . Os anéis ponderados pela distância lateral ao centro foram configurados como sinais de entrada para redes neurais, alcançando 9% de erro de classificação de jatos, mantendo uma eficiência de detecção de elétrons em 97%.

No trabalho [56], o *NeuralRinger* foi implementado no ambiente computacional do sistema de filtragem *online*, e foi apresentado como um algoritmo alternativo à colaboração ATLAS para a classificação de elétrons. A versão implementada no ambiente computacional do ATLAS foi conectada às cadeias de filtragem, possibilitando a extração *online* dos anéis em simulações de Monte Carlo e em dados de raios cósmicos. Essa implementação permitiu comparações diretas de eficiência com o algoritmo *CutBased*, que era o método padrão de classificação da época. Nos trabalhos [57] e [46], estudos em paralelo foram conduzidos visando a exploração de características de treinamento das redes neurais, normalização dos anéis, bem como

¹⁹Aqui o termo compactação se refere a redução da quantidade de sinais a serem avaliados pelo discriminador ao usar anéis como entradas, ao invés de utilizar a informação individual da energia medida em cada célula do calorímetro diretamente como entrada.

foram investigados métodos de pré-processamento de sinais utilizando técnicas da teoria da informação como por exemplo Análise de Componentes Independentes e Análise de Componentes Principais, na tentativa de melhoria do desempenho do classificador. No trabalho [58] um estudo aprofundado foi realizado com dados experimentais, produzidos pelas primeiras colisões do LHC na *Run1*.

No trabalho [59], foi desenvolvida a versão do algoritmo *NeuralRinger* no ambiente *offline*. A implementação em ambiente *offline* permitiu a comparação direta dos resultados do *NeuralRinger* com os métodos de seleção de elétrons utilizados como referência no ambiente *offline*. Em paralelo, nos trabalhos [60, 61], a implementação do método no ambiente *online* foi realizada e adequada para as condições de filtragem exigidas para a *Run2*. No segundo período de tomada de dados da *Run2*, o *NeuralRinger* foi empregado como método principal para classificação de elétrons substituindo o *CutBased*, alcançando redução de falsos elétrons em um fator de 2, se comparado ao método baseado em cortes lineares em variáveis discriminantes. Um investigação para a utilização do *NeuralRinger* em faixas de $E_T < 15$ GeV foi conduzida nos trabalhos [62, 63], e em 2020 a implementação técnica do *NeuralRinger* foi concluída e integrada no ambiente computacional do ATLAS²⁰.

2.6 Monitoração e Qualidade de Dados no Sistema de Filtragem *Online*

Antes de serem certificados para análise física, os dados coletados devem ser examinados para garantir que estejam livres de qualquer problema proveniente de *hardware* ou *software*, que podem comprometer sua integridade. A identificação imediata de possíveis problemas na tomada de dados torna viável uma ação rápida para investigar, corrigir e potencialmente prevenir futuros problemas que poderiam tornar os dados inutilizáveis. Para isso, os dados coletados por cada subsistema de detecção são monitorados e suas condições registradas [64]. Para monitorar o funcionamento dos algoritmos de filtragem, bem como analisar a qualidade dos dados (DQ - *Data Quality*) coletados, o ATLAS possui um sistema de monitoração e avaliação de qualidade de dados dedicado, no qual o seu funcionamento é ilustrado na Figura 2.17 [65].

O sistema de qualidade de dados é dividido em estágios *online* e *offline*. No estágio *online*, histogramas de variáveis de calorimetria são exibidos por meio de

²⁰Diferentes processos físicos relacionados com elétrons dominam específicas faixas de energia. Para faixas de $E_T \geq 15$ GeV, as amostras de sinal utilizadas no treinamento das redes neurais são elétrons provenientes do decaimento $Z \rightarrow ee$ e para faixas de $E_T < 15$ GeV, os exemplos de sinal utilizados no treinamento são elétrons do decaimento do méson J/Ψ . Por este motivo, os procedimentos de treinamento dos classificadores são diferentes para o limiar de energia 15 GeV [10].

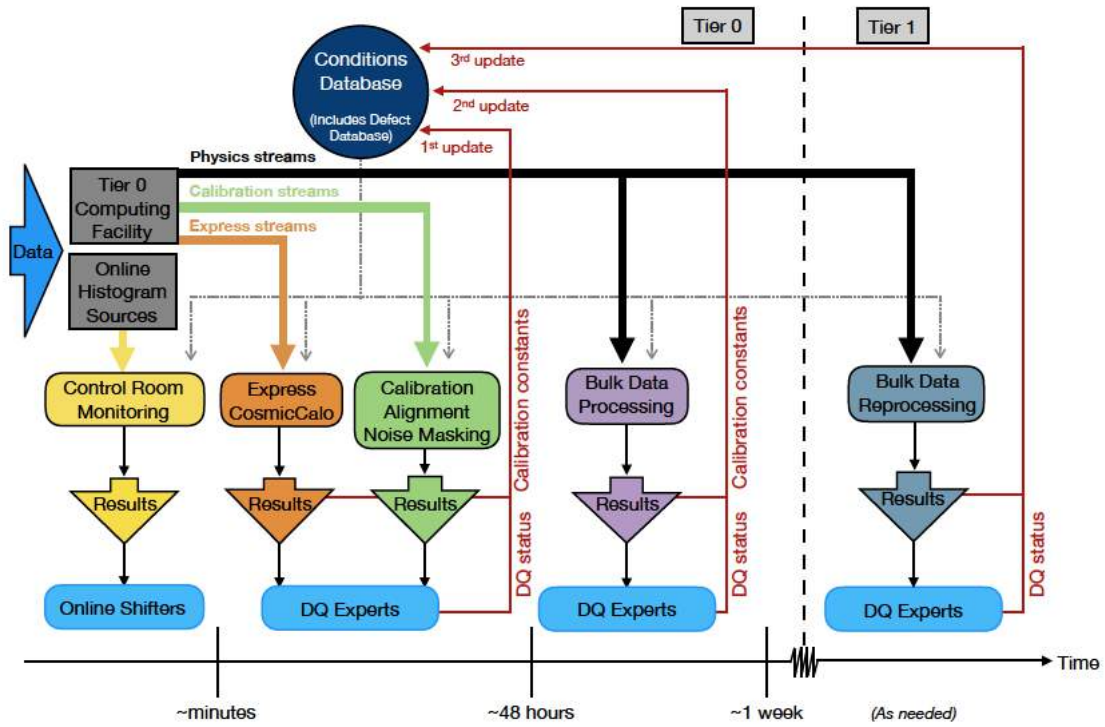


Figura 2.17: Diagrama esquemático do fluxo de trabalho na *Run2* para a avaliação da qualidade no ATLAS. Os histogramas fornecidos pelo sistema de DQ, apresentam distribuições de incluem cadeias de *trigger* previamente selecionadas para análises de qualidade de dados. Extraído de [65].

ferramentas interativas na sala de controle, para a equipe de turno (*online shifters*) acompanhar a coleta de dados e alertá-los sobre a presença de algum problema [66] [67]. Neste estágio informações do *hardware* do detector são analisadas, bem como são disponibilizadas informações dos dados selecionados pelo HLT usando algumas cadeias selecionadas para monitoramento [65]. As cadeias de filtragem reservadas para monitoramento são incluídas em um *menu de trigger*, que define uma lista completa de cadeias L1 e do HLT, bem como as suas configurações. No *menu*, as cadeias são divididas por grupos, permitindo que a equipe de turno possa selecionar quais cadeias devem ser exibidas. Os histogramas são disponibilizados para exibição após alguns minutos do início da tomada de dados e então são comparados²¹, nesta primeira verificação, com histogramas de referência.

Dados de referência são exaustivamente avaliados em rodadas de reprocessamento antes do período de tomada de dados. Estes reprocessamentos são realizados

²¹O teste estatístico de Kolmogorov-Smirnov é usualmente aplicado pelo sistema de avaliação e qualidade de dados, para indicar numericamente a proximidade entre os histogramas das grandezas de calorimetria dos dados que estão sendo coletados ao longo das colisões e histogramas de dados de referência coletados em tomadas de dados anteriores, que são extensivamente avaliados durante o período de *shutdown* do experimento [65].

com o objetivo de validar as novas versões do código fonte de análise do experimento e também identificar alguma falha nos dados que não foi observada com a versão do código fonte utilizado durante as colisões. Para avaliação dos reprocessamentos, um visualizador *web*, conforme apresentado na Figura 2.18 é utilizado, o qual possibilita a comparação da eficiência de filtragem e cada nível do HLT, permitindo também a comparação das distribuições de variáveis que descrevem o espalhamento longitudinal e lateral do chuveiro de partículas, conforme indicado pela Tabela 2.5, na qual algumas das principais grandezas são apresentadas ²².

Tabela 2.5: Lista de algumas das principais grandezas utilizadas pelo ATLAS, para a descrição do formato do chuveiro (*shower shapes*) calculadas pelos algoritmos de reconstrução.

Tipo	Nome	Descrição
Vazamento Hadrônico	R_{had1}	Razão em E_T da primeira camada do calorímetro hadrônico para a E_T do cluster EM (empregado para a região $ \eta < 0,8$ ou $ \eta > 1,37$.)
	R_{had}	Razão em E_T no calorímetro hadrônico para a E_T a E_T do cluster EM (empregado para a região $0,8 < \eta < 1,37$).
Região Traseira do Calorímetro EM	f_3	Razão de energia da camada traseira para a energia total do calorímetro de acordeão EM (aplicada somente para $E_T < 100$ GeV).
Camada Central do Calorímetro EM	$w_{\eta 2}$	Largura lateral do chuveiro, $\sqrt{(\sum E_i \eta_i)/(\sum E_i) - ((\sum E_i \eta_i)/(\sum E_i))^2}$ onde E_i é a energia e η_i a pseudorapidez da célula i .
	R_ϕ	Razão de energia em 3×3 células para a energia em 3×7 células centradas na posição do cluster do elétron.
	R_η	Razão de energia em 3×3 células para a energia em 7×7 células centradas na posição do cluster do elétron.
Camada de Tiras do Calorímetro EM	w_{stot}	Largura do chuveiro, $\sqrt{(\sum E_i (i - i_{max})^2)/(\sum E_i)}$ onde i percorre todas as tiras na janela $\Delta\eta \times \Delta\phi \approx 0,0625 \times 0,2$, correspondendo tipicamente a 20 tiras em η . Onde i_{max} é o índice da tira com a maior deposição de energia.
	E_{ratio}	Razão de energia da diferença entre os depósitos de maior e a de segunda maior energia no cluster para a soma dessas energias.
	f_1	Razão da energia na camada de tiras para o total de energia no acordeão EM do calorímetro.

²²Durante o desenvolvimento desta tese o autor deste trabalho foi responsável pelo desenvolvimento do *software* de monitoramento *online* para os algoritmos de tomada de decisão e reconstrução de elétrons e fótons no HLT do ATLAS. Além deste desenvolvimento o autor foi responsável pela manutenção dos exibidores web, algoritmos de monitoramento *offline* e pela avaliação da qualidade dos dados periodicamente reprocessados, para validação das atualizações do *software* central de análise do ATLAS. Detalhes da infraestrutura desenvolvida são apresentadas no Apêndice B desta tese.

levar aproximadamente uma semana. O produto final do monitoramento e avaliação dos dados é chamado de Catalogo de Dados de Referência (GRL - *Good Runs List*), que se trata de um conjunto de arquivos descritores, que contém a identificação dos dados e as condições de colisão durante a coleta. O dados registrados na GRL são utilizados também como referência para reprocessamentos posteriores.

Capítulo 3

Sistemas Inteligentes de Classificação

Este capítulo descreve de modo resumido os principais conceitos que dão suporte teórico aos Sistemas de Classificação. Também são apresentados brevemente métodos de treinamento de classificadores baseados em redes neurais artificiais utilizados neste trabalho. Algoritmos de computação bio-inspirada são apresentados no contexto do treinamento de redes neurais e uma revisão bibliográfica de aplicações de aprendizado de máquina em física experimental de altas energias é realizada.

3.1 Classificação de Sinais

Conceitualmente, sistemas de classificação são modelos que associam um sinal de entrada à uma classe de padrões previamente conhecida ou até mesmo a novas classes (também chamadas de *novidades*). Neste contexto, sistemas de classificação tem ganhado destaque em distintas áreas da ciência e indústria, como por exemplo em reconhecimento de faces [68], biometria, diagnóstico de doenças [69], direção autônoma veicular e detecção de falhas em equipamentos e peças industriais [70, 71]. Um sistema inteligente de classificação pode ser constituído por um modelo matemático (D), no qual seus parâmetros internos podem ser ajustados usando um conjunto de sinais de entrada-saída:

$$D : \mathbb{R}^N \rightarrow \Omega, \quad (3.1)$$

sendo $\Omega = \{\omega_1, \dots, \omega_c\}$ o espaço de classes do problema, com p possíveis classes, e N o número de atributos necessários para realizar a classificação. Um sistema de classificação canônico é considerado um conjunto de c funções discriminantes $G = \{g_1(\mathbf{x}), \dots, g_c(\mathbf{x})\}$, onde no mapeamento descrito pela Equação 3.2:

$$g_i : \mathbb{R}^N \rightarrow R, i = 1, \dots, c, \quad (3.2)$$

é produzida uma pontuação para a respectiva classe. Normalmente, o vetor de sinais de entrada $\mathbf{x} = \{x_1, x_2, \dots, x_N\}$ é rotulado na classe com a maior pontuação e a escolha de rotulagem é chamada de regra de associação máxima, de acordo com a Equação 3.3:

$$D(\mathbf{x}) = \omega_i \in \Omega \iff g_i(\mathbf{x}) = \max [g_i(\mathbf{x})], i = 1, \dots, c. \quad (3.3)$$

As funções discriminantes particionam o espaço de características $\mathbb{R}^N$ em c regiões de classificação denotadas por $R_1, \dots, R_c$, conforme Equação 3.4:

$$R_i = \{\mathbf{x} | \mathbf{x} \in \mathbb{R}^N, g_i(\mathbf{x}) = \max [g_k(\mathbf{x})], k = 1, \dots, c\}, i = 1, \dots, c. \quad (3.4)$$

A região de classificação para a classe ω_i é o conjunto de pontos para os quais a i -ésima função discriminante tem a maior pontuação. Tais regiões são especificadas pelo classificador D , ou de forma equivalente pelas funções discriminantes G .

3.1.1 Problemas de Detecção Binária

Um sistema de classificação binário corresponde a um modelo que realiza a escolha entre duas hipóteses: H_1 , normalmente associada à ocorrência do evento de interesse e H_0 , que representa a ocorrência de evento não relevante. Considerando que cada uma das hipóteses é associada a uma saída, mapeada numa região do espaço de observação de dimensão N , um ponto neste espaço pode ser representado pelo vetor dado pela Equação 3.5:

$$\mathbf{r} = [r_1, r_2, \dots, r_N]. \quad (3.5)$$

Então, um mecanismo de transição probabilística gera pontos de acordo com as densidades de probabilidade condicionais $P_{r/H_0}(R/H_0)$ e $P_{r/H_1}(R/H_1)$. Quando tais funções densidade de probabilidade (*pdf – probability density function*) são conhecidas (ou podem ser estimadas), critérios como Máximo a Posteriori (MAP), de Bayes e Neyman-Pearson, podem ser utilizados para a tomada de decisão [72] entre as hipóteses, sendo realizada por meio de uma ou mais observações do processo de interesse. Considerando que o processo de observação pode conter ruído, tem-se:

$$H_0 : \mathbf{r} = \mathbf{n} \quad (3.6)$$

$$H_1 : \mathbf{r} = \mathbf{n} + \mathbf{s}, \quad (3.7)$$

onde $\mathbf{r}$ é o sinal observado, $\mathbf{n}$ o ruído aditivo e $\mathbf{s}$ é o sinal de interesse. Um sistema de classificação deve operar sobre o sinal recebido ($\mathbf{r}$) e decidir por uma das duas hipóteses. Uma regra de decisão usualmente produz o particionamento do espaço

de observação em duas regiões Z_1 e Z_0 , para as quais são associadas as hipóteses H_1 e H_0 a depender da observação de $\mathbf{r}$, conforme ilustrado na Figura 3.1. Nesses casos, a razão de verossimilhança $L(\mathbf{R})$ entre as *pdfs* do sinal recebido em cada hipótese maximiza a discriminação do sinal de interesse é encontrada por meio da Equação 3.8:

$$L(\mathbf{R}) = \frac{f_{r/H_1}(R/H_1)}{f_{r/H_0}(R/H_0)}. \quad (3.8)$$

Considerando γ_l um patamar de decisão, definido pela Equação 3.9, em que P_0 e P_1 são as probabilidades a priori da fonte produzir H_0 ou H_1 respectivamente, e de acordo com a Equação 3.10, caso a razão de verossimilhança $L(\mathbf{R})$ tenha valor maior que o patamar (γ_l) decide-se por H_1 , caso contrário escolhe-se H_0 . No limite quando $L(\mathbf{R}) = \gamma_l$ pode-se decidir tanto por H_1 como por H_0 .

$$\lambda = \frac{P_0}{P_1}, \quad (3.9)$$

$$L(\mathbf{R}) \geq_{H_0}^{H_1} \gamma_l. \quad (3.10)$$

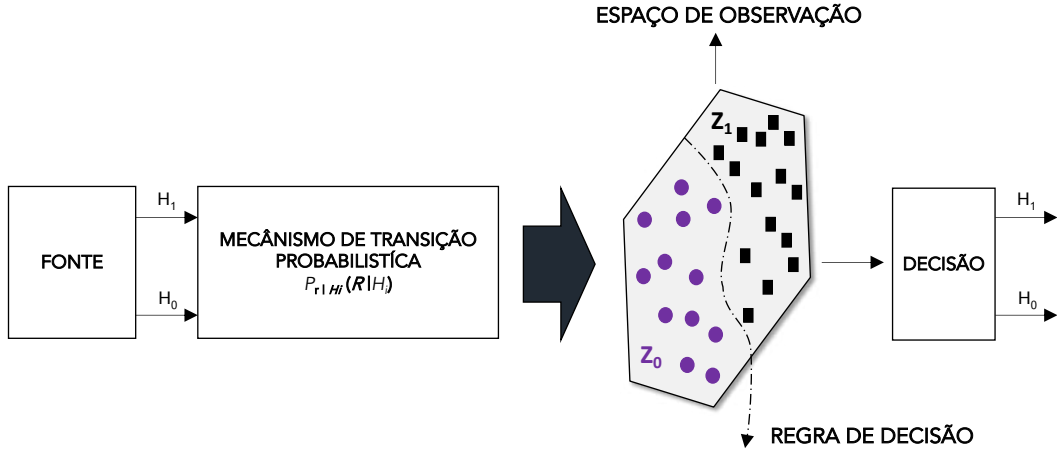


Figura 3.1: Ilustração de um problema de classificação entre duas hipóteses (decisão binária).

A avaliação de classificadores pode ser realizada por medidas de desempenho, que os qualificam em relação ao quanto de acerto e erros de classificação são encontrados na operação do modelo. A probabilidade de detecção (P_D) é relacionada com a medida de acerto de classificação da classe de interesse, podendo ser definida como a probabilidade do classificador de identificar a hipótese positiva (H_1) corretamente, conforme a Equação 3.11:

$$P_D = \int_{z_1} P_{r/H_1}(R/H_1) dR, \quad (3.11)$$

sendo $z_1 \subset \mathbb{R}^N$ a região definida pelo classificador para a classe de interesse (hipótese H_1). A probabilidade de Falso Alarme (P_F) está relacionada à aprovação do classificador de exemplos da hipótese negativa (H_0) como pertencentes à classe de hipóteses positivas (H_1), indicada pela Equação 3.12:

$$P_F = \int_{z_1} P_{r/H_0}(R/H_0) dR. \quad (3.12)$$

Nem sempre é possível conhecer as pdfs condicionais e por isso um procedimento de aprendizagem de máquina, baseado em dados do processo pode ser utilizado para construir um modelo com regras de decisão [72]. Nestes modelos, o patamar de decisão (valor para a determinação das classes) pode ser determinado a partir de realizações do processo e do conhecimento especialista usado para atribuir os rótulos das classes. O patamar de decisão pode ser ajustado para priorizar a detecção da classe de interesse, ou para impedir a aceitação da classe não relevante, ou ainda ser definido para o equilíbrio de desempenho no acerto das classes. A escolha do patamar depende então das características de operação do problema de classificação, no qual algumas classes podem possuir custos distintos, como pode ser observado por exemplo em sistemas de apoio ao diagnóstico de doenças graves. Nestes problemas, em geral, o principal interesse do sistema de auxílio ao diagnóstico é detectar se o indivíduo doente (hipótese H_1) tem a enfermidade, pois caso o diagnóstico informe que o indivíduo avaliado esteja saudável, um tratamento inapropriado pode ser realizado contribuindo então para a piora da sua saúde. Classificar indivíduos doentes como saudáveis apresenta um custo maior para o paciente, se comparado ao erro em classificar um indivíduo saudável (hipótese H_0) como doente, que mesmo também sendo um erro de diagnóstico apresenta menor risco sobre a vida do indivíduo examinado. Estas restrições de operação devem ser avaliadas no projeto de classificadores, para que sejam adequados na aplicação.

Conforme indicado pela Figura 3.2, o resultado da tarefa de classificação binária pode gerar quatro situações distintas: Verdadeiros Positivos (VP), Falsos Positivos (FP), Verdadeiros Negativos (VN) e Falsos Negativos (FN), conforme descritos a seguir:

- Verdadeiros Positivos (VP): Exemplos da classe positiva classificados corretamente. A Taxa de Verdadeiros Positivos pode ser calculada através da equação 3.13:

$$TVP = \frac{VP}{VP + FN}. \quad (3.13)$$

- Falsos Positivos (FP): Eventos pertencentes a classe negativa, classificados incorretamente como classe positiva. A Taxa de Falsos Positivos pode ser

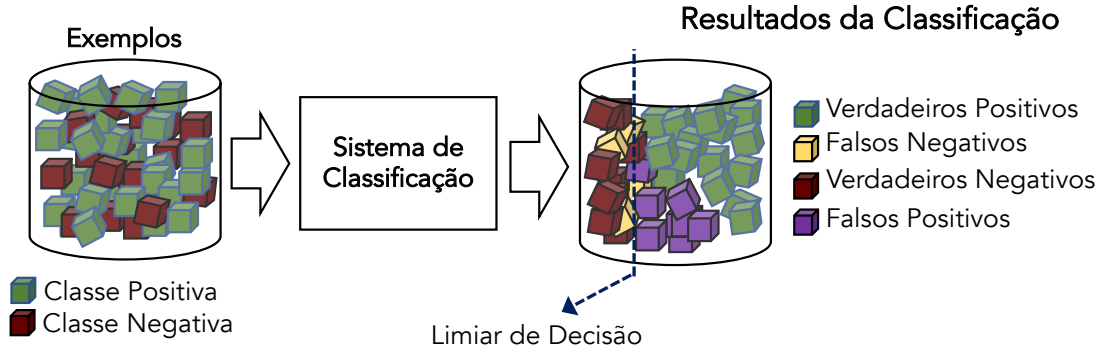


Figura 3.2: Ilustração que exemplifica os possíveis resultados de um sistema binário de classificação.

calculada através da equação 3.14:

$$TFP = \frac{FP}{FP + VN}. \quad (3.14)$$

- Verdadeiros Negativos (VN) (ou Especificidade): Exemplos da classe negativa classificados corretamente. A Taxa de Verdadeiros Negativos ($TVN = 1 - P_F$) pode ser calculada através da equação 3.15:

$$TVN = \frac{VN}{VN + FP}. \quad (3.15)$$

- Falsos Negativos (FN): Exemplos pertencentes a classe positiva, classificados incorretamente como classe negativa. A Taxa de Falsos Negativos ($TFN = 1 - P_D$) pode ser calculada através da equação 3.16:

$$TFN = \frac{FN}{VP + FN}. \quad (3.16)$$

Um modo de apresentar os resultados obtidos da operação de um sistema de classificação é por meio de uma tabulação cruzada entre a classe prevista pelo modelo e a classe real dos exemplos. Essa tabulação é conhecida como *matriz de confusão*, que armazena as frequências absolutas dos eventos classificados e suas relações com as classes conhecidas, como exemplifica a Tabela 3.1.

O patamar de decisão do sistema de classificação pode ser escolhido através da análise da curva ROC (*Receiver Operating Characteristics*) [73], que exhibe o desempenho do classificador binário considerando eventos classificados como VP e FP em função do patamar. A curva ROC pode ser obtida através dos pontos no qual os eixos das coordenadas representam a $P_F(y, \gamma_l)$ e $P_D(y, \gamma_l)$, respectivamente, sendo

Tabela 3.1: Matriz de confusão para um prolema de classificação binária.

Classe Observada	Classe Prevista	
	Condição Positiva	Condição Negativa
Observação Positiva	VP	FP
Observação Negativa	FN	VN

y o sinal de saída do classificador e γ_l o patamar de decisão, definido entre o valor mínimo e máximo possível de saída do classificador $[t_1, t_2]$, conforme a Equação 3.17:

$$\text{ROC} = (P_F(y, \gamma_l), P_D(y, \gamma_l)), \gamma_l \in [t_1, t_2]. \quad (3.17)$$

Observa-se na Figura 3.3 o desempenho de dois classificadores hipotéticos, onde o Classificador A apresenta pontos da curva ROC mais próximos do ponto (0, 100). Esta característica indica um melhor desempenho de classificação em relação ao Classificador B, considerando uma mesma taxa de aceitação de falsos positivos (por exemplo para $P_F = 20\%$). Observa-se que para a faixa de operação no intervalo de $80 \leq P_F \leq 100\%$ os classificadores não apresentam diferença significativa de desempenho.

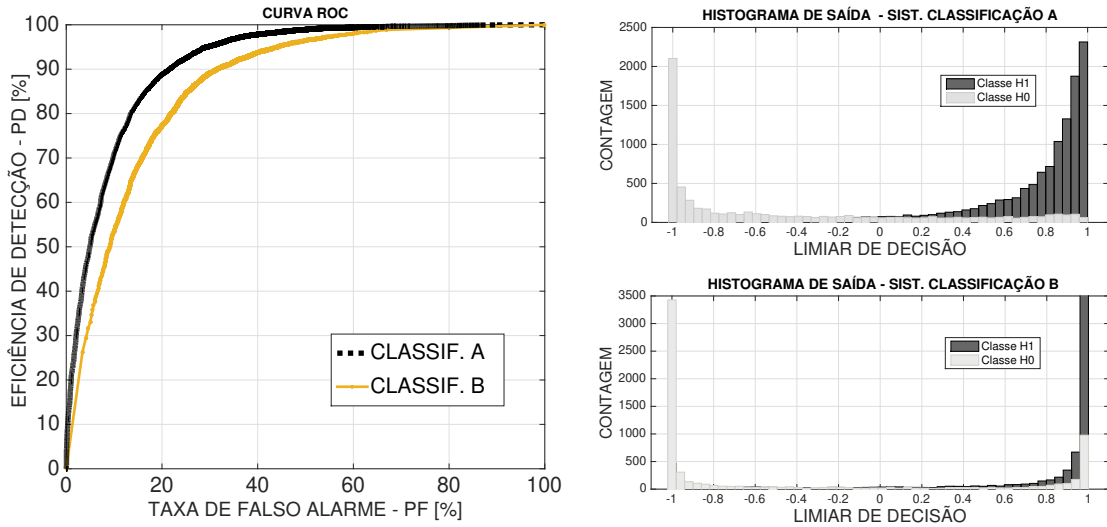


Figura 3.3: Curvas ROC de dois classificadores binários hipotéticos distintos e os respectivos histogramas de saídas.

O projeto de classificadores é tradicionalmente dividido em duas fases: a primeira relacionada a uma etapa de aprendizagem (treinamento) e a outra responsável por testar o modelo treinado, averiguando a capacidade do mesmo em identificar corretamente novos exemplos (etapa de testes). O principal objetivo do treinamento é ajustar uma função discriminante que seja capaz de classificar os eventos de modo que haja ocorrência reduzida de erros. Neste contexto, máquinas de aprendizagem

podem obter conhecimento a partir de um conjunto de casos com categorias conhecidas (aprendizado supervisionado) [74], e a partir de análises de agrupamento e correlações dos exemplos, cujas classes de padrões são desconhecidas (aprendizado não supervisionado) [75]. Na aprendizagem supervisionada, conforme indicado pelo diagrama na Figura 3.4, o modelo codifica em suas variáveis internas o conhecimento do padrão que melhor descreve as classes com base em um conjunto estatisticamente representativo com amostras rotuladas. Um evento (amostra) do conjunto utilizado para o treinamento, possui formato do tipo $\mathbf{x}_c = [x_1, x_2, \dots, x_N, C]$, em que N representa a dimensionalidade do espaço original de representação das classes e C é o índice que associa o exemplo a classe a qual ele pertence (rótulo). O modelo utiliza o conhecimento adquirido para então rotular um novo evento a ser apresentado ao classificador.

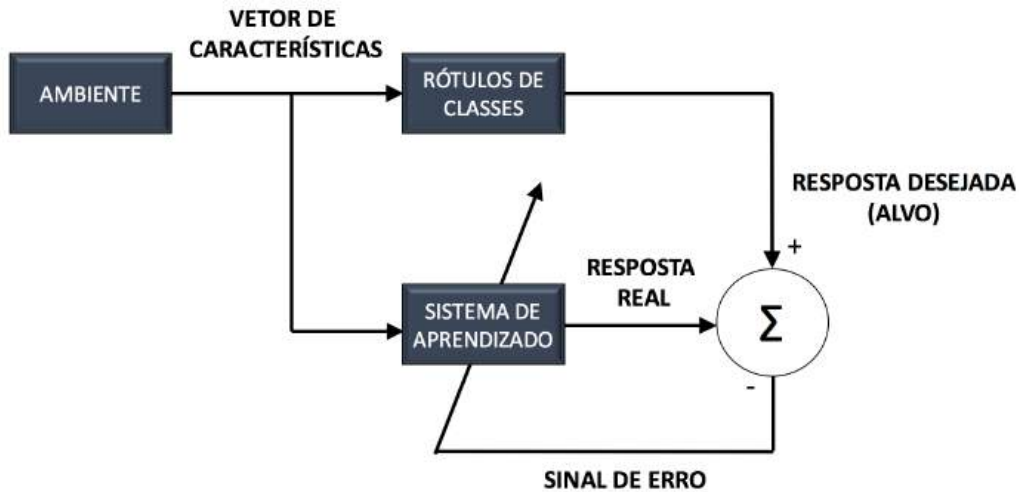


Figura 3.4: Diagrama em blocos do método de aprendizagem supervisionada para sistemas de classificação.

3.2 Redes Neurais Artificiais

Uma Rede Neural Artificial (RNA) [75] consiste em uma estrutura computacional com unidades de processamento inspiradas no modo como o neurônio biológico¹ processa uma informação em particular. As RNAs empregam uma conexão maciça de unidades de processamento distribuídas denominadas neurônios artificiais, que por meio de um processo iterativo (algoritmo de aprendizagem) a RNA ajusta os parâmetros livres dos neurônios (pesos) para atuação em problemas de reconhecimento de padrões, previsão de séries temporais, identificação de sistemas, etc. O diagrama

¹O neurônio biológico é um tipo de célula que compõe o sistema nervoso, sendo responsável pela recepção, condução e transmissão dos impulsos nervosos [76].

da Figura 3.5 [75] apresenta um modelo para o neurônio artificial (*perceptron*), onde podem ser observados três elementos principais: um conjunto de sinapses ou elos de conexão, cada um caracterizado por um peso próprio, o somador, que irá efetuar a soma dos sinais de entrada ponderados pelas respectivas sinapses (combinação linear) e uma função de ativação, que é aplicada para restringir a amplitude de saída de um neurônio e para realização de mapeamentos não-lineares.

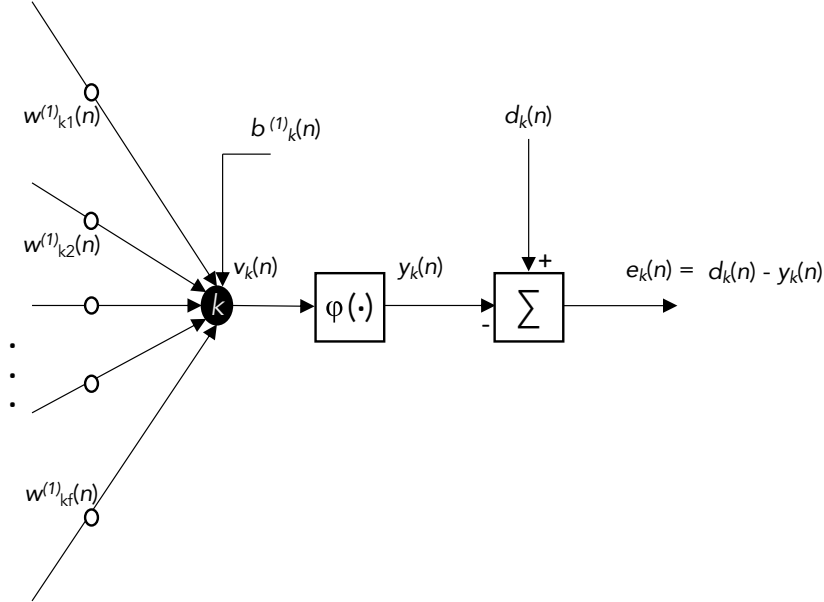


Figura 3.5: Grafo de fluxo de sinal ressaltando os detalhes do neurônio de saída k . Adaptado de [75].

Considerando o neurônio k de saída ilustrado na Figura 3.5, o sinal do erro $e_k(n)$ produzido por um sinal de entrada, é calculado através da comparação entre o alvo $d_k(n)$ e a saída real do neurônio $y_k(n)$, conforme Equação 3.18:

$$e_k(n) = d_k(n) - y_k(n), \quad (3.18)$$

onde n é o índice da iteração do algoritmo de aprendizagem. O valor instantâneo da energia total do erro ($\xi(n)$) é obtido somando-se os termos $\frac{1}{2}e_k^2(n)$, calculado por:

$$\xi(n) = \frac{1}{2} \sum_{k \in S} e_k^2(n), \quad (3.19)$$

sendo S o conjunto de todos os possíveis neurônios da camada de saída. Ao ser alimentado por um conjunto de sinais funcionais produzidos por uma camada de neurônios a sua esquerda, o neurônio k produz um sinal local ($v_k(n)$) antes da aplicação da função de ativação ($\varphi(\cdot)$) dado pela Equação 3.20:

$$v_k(n) = \sum_{f=0}^q w_{kf}(n) \times y_f(n), \quad (3.20)$$

sendo q o número total de exemplos aplicados ao neurônio k . Considerando o bias (b_k) aplicado ao neurônio k representado como peso sináptico w_{k0} , o sinal de saída $y_k(n)$ na iteração n resultante da aplicação da função de ativação $\varphi(\cdot)$ pode ser escrito conforme Equação 3.21:

$$y_k(n) = \varphi_k(v_k(n)). \quad (3.21)$$

Os pesos sinápticos $w_{kf}(n)$ (que conectam uma entrada f ao neurônio k) são então corrigidos por $\Delta_{kf}(n)$, que é proporcional à derivada parcial definida pela Equação 3.22:

$$\frac{\partial \xi(n)}{\partial w_{kf}} = \frac{\partial \xi(n)}{\partial e_k(n)} \frac{\partial e_k(n)}{\partial y_k(n)} \frac{\partial y_k(n)}{\partial v_k(n)} \frac{\partial v_k(n)}{\partial w_{kf}(n)}, \quad (3.22)$$

que determina a direção de busca no espaço de pesos, para o neurônio sináptico w_{kf} . Diferenciando ambos os lados da Equação 3.19 em relação a $e_k(n)$, obtêm-se:

$$\frac{\partial \xi(n)}{\partial e_k(n)} = e_k(n), \quad (3.23)$$

e diferenciando ambos os lados da Equação 3.18 em relação a $y_k(n)$, encontra-se:

$$\frac{\partial e_k(n)}{\partial y_k(n)} = -1. \quad (3.24)$$

A seguir, diferenciando a Equação 3.21 em relação a $v_k(n)$ e a Equação 3.20 em relação a $v_{kf}(n)$ obtêm-se as Equações 3.25 e 3.26 respectivamente:

$$\frac{\partial y_k(n)}{\partial v_k(n)} = \varphi'_k(v_k(n)), \quad (3.25)$$

$$\frac{\partial v_k(n)}{\partial w_{kf}(n)} = y_f(n). \quad (3.26)$$

Assim a Equação 3.22 pode ser rescrita como:

$$\frac{\partial v_k(n)}{\partial w_{kf}(n)} = -e_k(n) \varphi'_k(v_k(n)) y_f(n). \quad (3.27)$$

Utilizando a regra *delta*, a correção $\Delta_{kf}(n)$ aplicada ao peso sináptico $w_{kf}(n)$ pode ser definida pela Equação 3.28:

$$\begin{aligned} \Delta w_{kf}(n) &= -\eta \frac{\partial \xi(n)}{\partial w_{kf}(n)} \\ &= \eta \delta_k(n) y_f(n), \end{aligned} \quad (3.28)$$

sendo η a taxa de aprendizagem do algoritmo de retropropagação. O sinal negativo na Equação 3.28 indica que o procedimento busca uma mudança no espaço de pesos

que reduza o valor de $\xi(n)$. O gradiente local ($\delta_k(n)$) que aponta para as modificações necessárias nos pesos sinápticos é computado conforme Equação 3.29:

$$\begin{aligned}\delta_k(n) &= -\frac{\partial \xi(n)}{\partial v_k(n)} \\ &= e_k(n) \varphi'_k(v_k(n))\end{aligned}\tag{3.29}$$

Uma rede neural *perceptron* com múltiplas camadas (MLP - *Multilayer perceptron*) consiste em um conjunto de neurônios conectados entre si com nós sensoriais, que constituem uma camada de entrada, uma ou mais camadas ocultas de nós computacionais e uma camada de saída. A interconexão das entradas à camada de saída por pelo menos uma camada de neurônios intermediária (camada oculta) remove a limitação do modelo *perceptron* de solucionar somente problemas linearmente separáveis [77].

O algoritmo de retropropagação do erro (*backpropagation*) [75, 78] é tradicionalmente utilizado para o treinamento de redes neurais com múltiplas camadas. Algumas limitações dos algoritmos de retropropagação baseados em gradiente tem motivado o desenvolvimento de novas estratégias de treinamento de redes neurais MLP. Como o método de retropropagação é baseado em uma regra de descida sobre a superfície do erro, o algoritmo pode convergir para um mínimo local, que pode provocar baixa capacidade de generalização da rede, quando são apresentados exemplos que não foram incluídos na fase de treinamento [79]. A possível presença de mínimos locais na superfície do sinal de erro torna o algoritmo de otimização sensível ao procedimento de inicialização dos pesos sinápticos, sendo requeridas em geral, uma elevada quantidade de inicializações do algoritmo de otimização para diminuir a chance em escolher modelos nos quais a otimização dos pesos foi paralisada em um ponto de mínimo local, que eventualmente pode causar baixa generalização da rede [80]. O treinamento com retropropagação pode sofrer com a lentidão para convergência, aspecto que requer cuidadoso ajuste da taxa de aprendizagem para controlar os passos dados pelo gradiente sobre a função custo.

Os algoritmos de retropropagação com gradiente descendente também não permitem que classificadores neurais sejam otimizados para pontos de operação específicos (probabilidade de detecção e falso alarme). Usando o erro quadrático médio como função objetivo não é possível diretamente treinar RNAs com retropropagação para que se considere custos distintos entre às classes. Este recurso é importante principalmente em problemas cujas classes tenham diferentes graus de importância para o problema de classificação, como por exemplo no diagnóstico de doenças graves [81], identificação de eventos raros em HEP [53, 82, 83] e detecção de defeitos em instalações industriais [84]. Considerando estes aspectos, algoritmos de busca global,

como os algoritmos bio-inspirados têm sido projetados para o treinamento de redes neurais na tentativa de contornar problemas relacionados a sensibilidade a mínimos locais, bem como a inflexibilidade de modificações na função custo. Algoritmos baseados no modo como os seres vivos interagem e se adaptam às condições do meio em que vivem (otimização bio-inspirada) [85] têm sido crescentemente explorados na literatura no contexto do treinamento de redes neurais com distintas funções objetivo [84, 86], ajuste de fino de parâmetros da rede [87] e seleção de atributos relevantes em conjuntos de treinamento [88, 89].

3.2.1 Treinamento de Redes Neurais com Algoritmos Bio-Inspirados

Algoritmos bio-inspirados consistem em métodos computacionais inspirados por conceitos da teoria da evolução natural ou do comportamento dos seres vivos. Os Algoritmos Evolutivos (AE) [90] e a Inteligência de Enxame (IE) [91] são métodos que se destacam na literatura, sendo aplicados em muitos problemas de otimização. Conceitualmente, estes métodos adaptam uma população de indivíduos iterativamente, sendo os AEs inspirados no mecanismo de evolução natural e a IE inspirada no modelo de interação e convivência em grupos de seres vivos [92]. No treinamento de redes neurais artificiais, as Estratégias Evolutivas [93, 94] e Algoritmos Genéticos [95, 96] são os principais grupos de algoritmos que se destacam no campo dos Algoritmos Evolutivos. Na área de Inteligência de Enxames, o método Otimização por Enxame de Partículas (PSO - *Particle Swarm Optimization*) é explorado em distintos problemas de engenharia no contexto do ajuste de RNAs [97, 98]. Uma breve descrição dos métodos de otimização bio-inspirados utilizados nesse trabalho é apresentada a seguir.

Algoritmos Genéticos

Os Algoritmos Genéticos [90] (AG) são definidos como estratégias de busca probabilística inspirada no processo biológico da evolução natural, proposto inicialmente pelo cientista Charles Darwin no trabalho *A Origem das Espécies* [99] como também em aspectos da genética [85]. É um método de otimização global, que combina as características dos indivíduos da população através de uma troca estruturada e iterativa de informação genética.

O diagrama mostrado na Figura 3.6 mostra, de modo genérico, o procedimento iterativo da execução de um AG. Inicialmente uma população de soluções factíveis para o problema de otimização é gerada aleatoriamente, sendo cada indivíduo (cromossomo) composto pelas variáveis a serem otimizadas (genes). Os indivíduos da população são avaliados por meio de uma função custo, e o valor obtido é utilizado

para eleger, em termos probabilísticos, aqueles com maior potencial para serem preservados na população. Os indivíduos com melhor valor da função custo em uma determinada geração, são selecionados para serem os progenitores das soluções que irão fazer parte da geração seguinte.

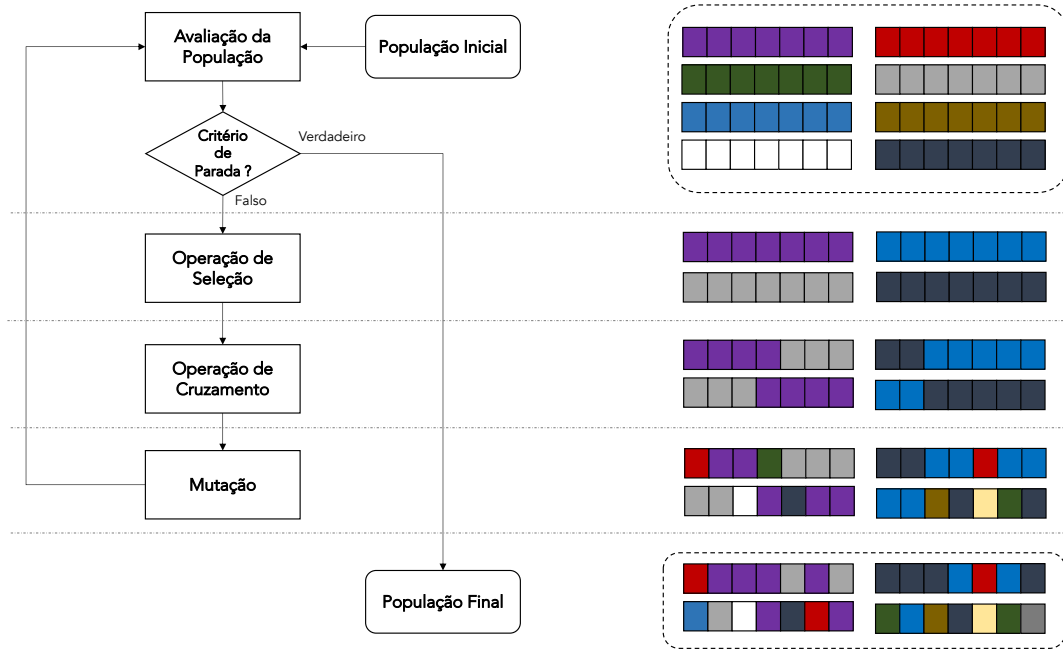


Figura 3.6: Fluxograma de um GA genérico com as principais operações.

Um processo probabilístico é utilizado para efetuar a seleção dos progenitores, sendo os principais métodos de seleção: roleta, torneio, truncamento e aleatória [85], permitindo que o indivíduo com maior valor de aptidão tenha maiores chances de participar da operação de reprodução. Como resultado, indivíduos com melhor pontuação na função custo têm maior chance de perpetuar suas características e os menos aptos tendem a desaparecer. Na etapa de reprodução, operadores genéticos são utilizados para recombinar indivíduos, gerando uma nova população a partir dos indivíduos selecionados. Um operador de mutação pode também atuar sobre as soluções resultantes do processo de recombinação, alterando aleatoriamente os genes dos indivíduos, com o objetivo de manter um nível de diversidade na população. A população final obtida pelo GA é encontrada caso a seja alcançado o critério de parada, que pode estar associado, por exemplo, a um valor pré-estabelecido da função custo, ou até mesmo uma quantidade sucessiva de iterações sem que haja melhoria na função custo. O GA não garante que a solução ótima do problema seja encontrada, mas fornece um conjunto final de soluções que podem ter valor da função custo próxima ao ótimo global da superfície de otimização.

A facilidade de implementação dos GAs e sua resiliência a mínimos locais torna atraente sua aplicação, e para o treinamento de redes neurais, em muitos problemas,

apresentam desempenho satisfatório. Entretanto, esses algoritmos podem apresentar elevado custo computacional à medida que a dimensão dos sinais de entrada e o número de camadas da rede aumenta (consequentemente o número de pesos a serem otimizados). Um outro problema enfrentado pelos AGs no contexto do treinamento de redes neurais é que ao utilizar operações de recombinação entre membros da população, não são consideradas as possíveis relações entre variáveis vizinhas (blocos construtivos) [85]. A quebra dos blocos construtivos pode ocasionar lentidão de convergência e número excessivamente grande de iterações do algoritmo.

Otimização por Enxame de Partículas

A Otimização por Enxame de Partículas (PSO do inglês *Particle Swarm Optimization*) é um método de otimização baseado em enxame, inspirado no comportamento de interação social e cooperativa de espécies [100]. O algoritmo PSO manipula uma população de l soluções candidatas (partículas), que possuem três atributos: posição ($\mathbf{p}^{(l)}$ - representação da solução candidata do problema de otimização), velocidade ($\mathbf{v}^{(l)}$) e experiência individual ($\mathbf{p}_{best}^{(l)}$). As partículas são movidas pelo espaço de busca $\mathbb{R}^S$ de acordo com operações matemáticas na posição e velocidade, considerando sua experiência individual e o melhor candidato do enxame para solucionar o problema de otimização na atual iteração ($\mathbf{g}_{best}$). A Figura 3.7 [101] ilustra o modo como uma partícula é movida no espaço de busca, e como ela é influenciada pelos agentes do enxame.

O movimento das partículas é realizado através da combinação da sua velocidade instantânea com a melhor posição na qual a partícula já ocupou no espaço de busca ($\mathbf{p}_{best}^{(l)}$) e a posição da partícula com melhor desempenho no espaço de busca até o momento ($\mathbf{g}_{best}$). A atualização da velocidade da l -ésima partícula na iteração n pode ser computada pela Equação 3.30:

$$\mathbf{v}^{(l)}(n+1) = w(n) \times \mathbf{v}^{(l)}(n) + c_1 \times r_1 \times (\mathbf{p}_{best}(n) - \mathbf{p}^{(l)}(n)) + c_2 \times r_2 \times (\mathbf{g}_{best}(n) - \mathbf{p}^{(l)}(n)), \quad (3.30)$$

sendo $w(n)$ o coeficiente de inércia, c_1 e c_2 são os coeficientes de interação social e cognitiva respectivamente, e r_1 e r_2 são parâmetros aleatórios definidos entre 0 e 1. A l -ésima partícula do enxame tem a sua posição atualizada pela Equação 3.31:

$$\mathbf{p}^{(l)}(n+1) = \mathbf{p}^{(l)}(n) + \mathbf{v}^{(l)}(n+1), \quad (3.31)$$

O vetor $\mathbf{p}_{best}^{(l)}$ é atualizado se a partícula apresentar melhor desempenho quando comparada à iteração anterior. Caso alguma partícula obtiver melhora no desempenho da função custo, se comparada a todo o enxame, este vetor assume a posição

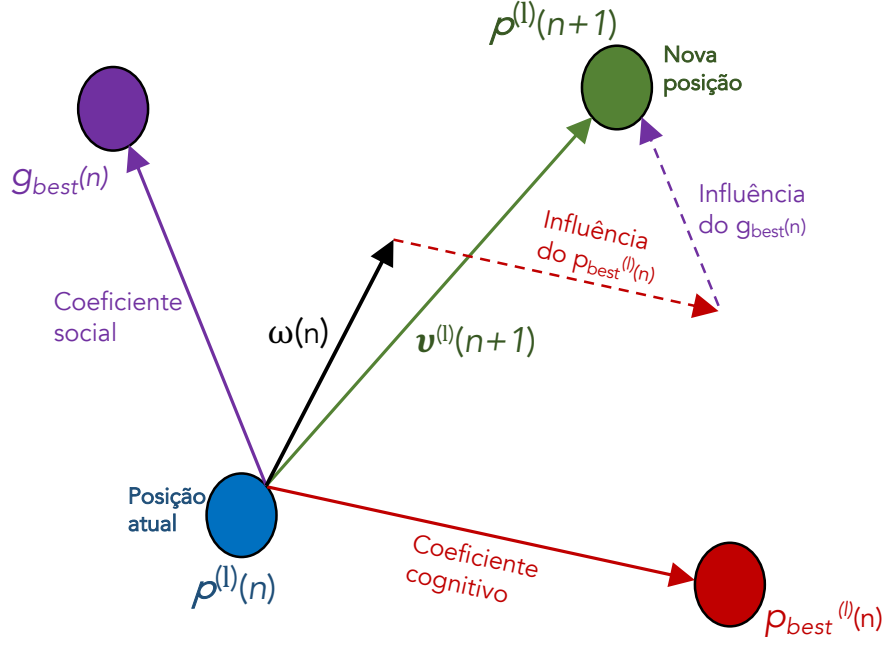


Figura 3.7: Ilustração de uma partícula com posição $\mathbf{p}^{(l)}$ que se move no espaço de busca, sendo influenciada pela partícula do enxame com melhor função custo ($\mathbf{g}_{best}(n)$) e da melhor posição ocupada pela partícula l na iteração n ($\mathbf{p}_{best}^{(l)}$). Os coeficientes cognitivo (c_1) e social (c_2) ponderam a influência de $\mathbf{p}_{best}^{(l)}$ e $\mathbf{g}_{best}(n)$ respectivamente na definição da velocidade. O coeficiente de inércia $w(n)$ ajusta o passo de atualização da velocidade $\mathbf{v}^{(l)}(n+1)$. Adaptado de [101].

da melhor partícula ($\mathbf{g}_{best}$). O coeficiente de inércia w é atualizado a cada iteração n , expressado pela Equação 3.32:

$$w(n) = (w_1 - w_2) \times \frac{(i_t - n)}{i_t} + w_2, \quad (3.32)$$

onde i_t é o número máximo de interações, w_1 and w_2 são os valores arbitrários iniciais e finais respectivamente do coeficiente de inércia w . As iterações do algoritmo PSO continuam até que um critério de parada seja encontrado ou sejam obtidas baixas variações nas velocidades das partículas.

Estratégia Evolutiva com Adaptação da Matriz de Covariância

As estratégias evolutivas trabalham com um conjunto de prováveis soluções do problema apresentando populações de tamanho variável e populações intermediárias. Numa EE do tipo $(\mu + \lambda)$ uma quantidade μ de indivíduos gera uma quantidade λ de descendentes, e dentre os $\mu + \lambda$ indivíduos os μ melhores são escolhidos para a próxima geração. Já em uma EE do tipo (μ, λ) os μ melhores são escolhidos entre

os λ componentes da geração intermediária. Algumas estratégias evolutivas tentam a cada geração estimar uma distribuição a ser utilizada para amostrar novos indivíduos, sendo que a distribuição é adaptada a cada geração com a informação dos melhores indivíduos da geração anterior [102].

A Estratégia de Evolução com Adaptação de Matriz de Covariância (*Covariance Matrix Adaptation - Evolution Strategy*) CMA-ES, é um tipo de Estratégia Evolutiva (EE) que reúne informações sobre etapas de busca bem-sucedidas e as utiliza para modificar a matriz de covariância que amostra indivíduos de uma população [103]. As mudanças na matriz de covariância são realizadas de tal forma que as variações nas direções do espaço de busca que foram previamente bem sucedidas sejam potencializadas. Para isso são utilizadas informações acumuladas por várias etapas da busca, o que permite adaptar de forma confiável a matriz de covariância. A Figura 3.8 [104] ilustra uma evolução da população realizada pelo CMA-ES.

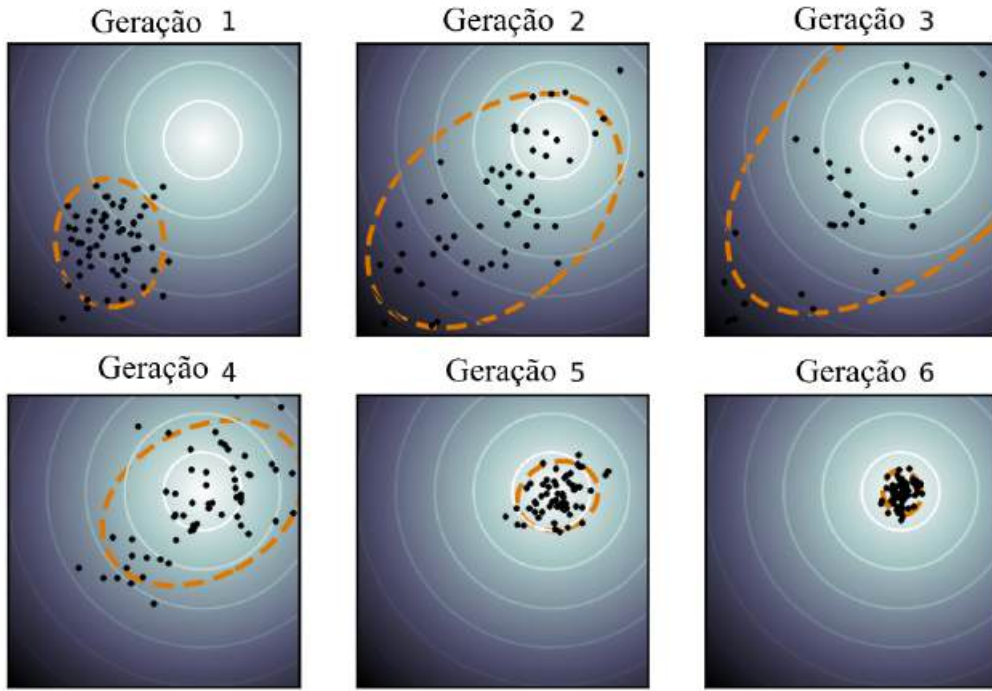


Figura 3.8: Ilustração do movimento provocado a cada geração pelo método CMA-ES aos indivíduos de uma população em um espaço de busca hipotético. Adaptado de [104].

A ideia principal é que o melhor indivíduo de uma geração g com λ indivíduos denotado por $\nu_{1:\lambda}^{(g)}$, seja selecionado para a geração $g + 1$ e dê origem a λ novos indivíduos a partir de perturbações aplicadas a ele. Para o CMA-ES, o indivíduo que dará origem a nova geração será obtido através da média entre os μ melhores da geração atual. O CMA-ES gera o k -ésimo indivíduo da próxima geração, denotado

por $\nu_k^{(g+1)}$, conforme modelo probabilístico definido pela Equação 3.33:

$$\nu_k^{(g+1)} \sim \mathbf{m}^{(g)} + \sigma^{(g)} N_k\left(0, \mathbf{C}^{(g)}\right) \quad \text{para } k = 1, 2, \dots, \lambda, \quad (3.33)$$

onde $\sigma^{(g)}$ é o passo de perturbação para a geração g ; $N_k(0, \mathbf{C}^{(g)})$ representa a distribuição normal multivariada com média zero e matriz de covariância $\mathbf{C}^{(g)}$, e $\mathbf{m}^{(g)}$ é computado por meio da média ponderada entre os μ indivíduos com melhor pontuação na função custo da geração g , de acordo com a Equação 3.34:

$$\mathbf{m}^{(g)} = \sum_{i=1}^{\mu} \omega_i \nu_{i:\lambda}^g, \quad (3.34)$$

sendo:

$$\sum_{i=1}^{\mu} \omega_i = 1, \quad \omega_1 \geq \omega_2 \geq \dots \geq \omega_{\mu}, \quad (3.35)$$

onde $\nu_{i:\lambda}^{(g)}$ é o i -ésimo melhor indivíduo da geração g . Os valores de ω_i podem ser obtidos conforme a Equação :

$$\omega_i = \frac{\log(\mu + 0,5) - \log(i)}{\sum_{i=1}^{\mu} \omega_i}. \quad (3.36)$$

Considerando o processo de perturbação, representado através da Equação 3.33, percebe-se que a geração de novos indivíduos está diretamente ligada à matriz de covariância ($\mathbf{C}^{(g)}$), portanto adaptá-la pode resultar na geração de indivíduos mais aptos. A matriz de covariância pode ser interpretada como um hiper-elipsoide que representa a densidade de distribuição da população, logo sua adaptação busca que essa variabilidade seja maximizada nos eixos que geraram melhores indivíduos, aumentando consequentemente as chances de serem gerados indivíduos mais aptos nas próximas gerações. Utilizando vetor médio apresentado na Equação 3.34, a matriz de covariância pode ser calculada pela Equação 3.37:

$$\mathbf{C}_{\lambda}^{(g+1)} = \frac{1}{\lambda} \sum_{i=1}^{\lambda} \left(\nu_{i:\lambda}^{(g+1)} - \mathbf{m}^{(g)} \right) \left(\nu_{i:\lambda}^{(g+1)} - \mathbf{m}^{(g)} \right)^T, \quad (3.37)$$

Para atualizar a matriz de covariância com ponderação pela matriz obtida para as gerações mais recentes a matriz de covariância pode ser computada pela Equação 3.38, chamada de *rank- μ -update*:

$$\mathbf{C}^{(g+1)} = (1 - c_{\mu}) \mathbf{C}^{(g)} + c_{\mu} \sum_{i=1}^{\mu} \omega_i (\mathbf{y}_{i:\lambda}^{(g+1)})(\mathbf{y}_{i:\lambda}^{(g+1)})^T, \quad (3.38)$$

sendo $\mathbf{y}_{i:\lambda}^{(g+1)}$ calculada pela Equação 3.39, o coeficiente $c_\mu < 1$ é a taxa de aprendizagem para atualização da matriz de covariância.

$$\mathbf{y}_{i:\lambda}^{(g+1)} = \frac{(\nu_{i:\lambda}^{(g+1)} - \mathbf{m}^{(g)})}{\sigma^{(g)}}, \quad (3.39)$$

Com este equacionamento, a matriz de covariância é atualizada usando como entrada os seus valores da iteração anterior, mais a informação sobre as mudanças observadas entre os μ melhores indivíduos de cada geração. A contribuição unicamente do melhor indivíduo encontrado na geração para a atualização da matriz de covariância pode ser definida pela Equação 3.40 (*rank-one update*):

$$\mathbf{C}^{(\mathbf{g}+1)} = (1 - c_1)\mathbf{C}^{(\mathbf{g})} + c_1(\mathbf{y}_{i:\lambda}^{(g+1)})(\mathbf{y}_{i:\lambda}^{(g+1)})^T, \quad (3.40)$$

onde c_1 tem a mesma função que c_μ , podendo ser calculado por meio da Equação:

$$c_1 = \frac{2}{(N + 1, 3)^2 + \mu_{eff}}, \quad (3.41)$$

sendo N a dimensão do espaço de busca e μ_{eff} definido pela Equação 3.42:

$$\mu_{eff} = \left(\sum_{i=1}^{\mu} w_i \right)^{-1}. \quad (3.42)$$

Um melhor ajuste da matriz de covariância pode ser obtido combinando os procedimentos *rank-one update* e *rank- μ update* [103].

3.3 Aplicações de Aprendizado de Máquina em Problemas de Física de Altas Energias

Redes Neurais Artificiais encontram diversas aplicações em experimentos de Física de Altas Energias [105, 106]. Devido à sua possibilidade de reconfiguração e realização mapeamentos não lineares, seu desenvolvimento tem sido crescente em diversos tipos de experimentos, como colisionadores de partículas, detectores de astro-partículas e detecção de corpos celestes usando ondas de rádio [107]. Considerando a rápida resposta, redes neurais com poucas camadas ocultas tem sido utilizadas em aplicações de seleção *online* de eventos, tendo resultados superiores a métodos baseados em cortes rígidos em variáveis discriminantes [18]. Em etapas de seleção *online* de eventos onde a maior parcela de ruído de fundo foi removida ou em ambientes *offline*, onde não há restrições de tempo para execução, redes profundas com muitas camadas ocultas, tem sido exploradas na literatura para classificação de sinais [108–110].

No trabalho [111] é apresentada uma investigação para a detecção de elétrons e fótons em um sistema de calorimetria com dupla leitura da colaboração RD52 [112]. Uma rede neural MLP foi utilizada para classificação elétron/píon, na faixa de 60 GeV. A rede foi alimentada com atributos que descrevem o perfil lateral do chuveiro de partículas, a relação entre a radiação de Cherenkov e o sinal de cintilação no calorímetro e o tempo de sensibilização dos sensores fotomultiplicadores. Os resultados encontrados indicaram que indica que 99,8% de todos os elétrons puderam ser identificados corretamente, admitindo 0,2% de erro de classificação de píons.

No trabalho [6] uma combinação de técnicas de aprendizado de máquina foi elaborada para classificação e estimar a distribuição angular de raios cósmicos, no experimento SciCRT [113, 114], no México. Um dos canais de interesse do SciCRT são nêutrons emitidos por tempestades solares, porém outras partículas vindo da interação dos raios cósmicos primários com núcleos da atmosfera terrestre entram no detector (raios gamma). A técnica kNN foi adotada para a etapa de separação raios cósmicos/raios gamma e em seguida, com os sinais classificados uma rede neural MLP com duas camadas ocultas foi utilizada para a estimar o desdobramento do espectro de energia de nêutrons e a sua distribuição angular. Utilizando sinais simulados por geradores de Monte Carlo de interação das partículas com o detector, a metodologia proposta encontrou 91% de acurácia na detecção de nêutrons com a utilização do algoritmo kNN. Os resultados obtidos com a rede neural para a estimação do espectro de energia apresentaram bom desempenho com o teste de Kolmogorov-Smirnov, quando comparado com distribuições de dados simulados, mostrando que a RNA foi capaz de reproduzir o espectro.

Algoritmos de otimização bio-inspirados também tem sido explorados em problemas de Física de Altas Energias. No artigo [115] os algoritmos GA e PSO foram utilizados para otimizar hiperparâmetros de um sistema de classificação baseado em árvores de decisão (BDT - *Boosted Decision Trees*). O sistema de classificação foi utilizado para separação sinal/ruído de fundo em sinais do experimento ATLAS. O principal objetivo foi detectar corretamente eventos do decaimento do bóson de Higgs ($H \rightarrow \tau\tau$) imersos em ruído de fundo [116], alcançando melhoria de 12% na significância mediana aproximada (AMS - *Approximate Median Significance*)^[2] se comparado a utilização de hiper-parâmetros padrões da técnica BDT.

O trabalho [110] apresentou o desenvolvimento de um sistema de classificação usando redes neurais profundas, para estimar a quantidade de múons presentes em chuveiros de raios cósmicos com ultra-energia^[3], que entram na atmosfera. O sistema

²O AMS é uma medida de desempenho que considerada o equilíbrio entre as quantidades de verdadeiros positivos e falsos positivos encontrados por um classificador, considerando um determinado nível de significância estatística [117].

³Partículas com ultra-alta energia apresentam energias acima de 10^{18} eV.

de classificação foi empregado em detectores de radiação de Cherenkov do Observatório Pierre Auger⁴, na Argentina [118]. A observação de múons contribui para o entendimento e caracterização de interações hadrônicas nos dados coletados pelo experimento, bem como provê suporte para a descoberta de fenômenos celestes geradores de raios cósmicos em faixas de ultra-alta energia. Um GA foi implementado para encontrar a melhor combinação entre números de neurônios ocultos, número de camadas e função de ativação de cada neurônio na rede neural profunda. Os resultados da metodologia proposta encontraram em média uma melhoria na precisão de 10%, para cada estação individual do experimento.

Uma investigação com métodos de aprendizagem de máquina foi realizada em [107] para a detecção de ondas de rádio originadas por pulsares⁵, que sofrem interferência por sinais de radiofrequência (RFI - *Radio Frequency Interference*). A RFI é um tipo de ruído que incide no processo de coleta de dados de radioastronomia que pode corromper os sinais cósmicos fracos e, potencialmente, levar a resultados enganosos. Ao serem treinadas com dados de longa gravação do experimento HTRU (*The High Time Resolution Universe*) [119] Redes MLP obtiveram taxas de falsos positivos de 0,059% em limiares de decisão que permitiram a detecção correta de 99,8% de sinais de pulsares.

A partir destes exemplos citados da literatura recente em diferentes tipos de experimentos em física de altas energias e áreas correlatas, nota-se que diversos problemas de classificação, ajuste de funções e otimização de parâmetros têm sido resolvidos com o projeto de classificadores neurais e algoritmos de otimização bio-inspirados.

⁴O projeto Auger apresenta uma série de 1660 estações de detecção de partículas Cherenkov de água espalhadas por 3.000 km² e 24 telescópios.

⁵Pulsares são estrelas de nêutrons que, em virtude de seu intenso campo magnético (da ordem de 10⁸ T), transformam a energia rotacional em energia eletromagnética. Estes corpos celestes permitem o estudo da relatividade geral em condições de intenso campo gravitacional.

Capítulo 4

Metodologia Proposta e Conjuntos de Dados

Neste capítulo, será apresentada a metodologia proposta para melhoria do sistema de classificação de elétrons baseado em redes neurais, atualmente empregado no segundo nível de seleção *online* de eventos do ATLAS.

O projeto dos classificadores neurais proposto neste trabalho é dividido em duas etapas distintas. A primeira etapa consiste em aplicar nos sinais em anéis, o pré-processamento R_p , conforme equação [4.1](#). O objetivo da etapa de pré-processamento é destacar nos sinais formatados em anéis, características discriminantes da informação lateral do chuveiro de partículas.

Em seguida, os sinais são utilizados como entradas para o treinamento do conjunto de redes neurais que realizará a classificação dos eventos. Para realizar o pré-processamento é necessário definir os coeficientes α e β que ponderam a energia e a distância dos anéis com relação a região de interesse respectivamente (ver equação [4.1](#)). Neste contexto, este trabalho propõe métodos de treinamento de redes neurais, projetados para ajustar os pesos do classificador neural simultaneamente com os coeficientes usados no pré-processamento.

Este capítulo está organizado do seguinte modo. Inicialmente, os métodos de treinamento de redes neurais propostos são apresentados, sendo que o procedimento para a extração da informação discriminante (informação lateral do chuveiro) dos sinais formatados em anéis também é descrito. Em seguida, as bases de dados utilizadas para o projeto dos classificadores neurais desenvolvidos neste trabalho são apresentadas. Uma seção é reservada para apresentar o modo como serão organizados os resultados obtidos com a metodologia proposta. Por fim, serão apresentados os parâmetros utilizados para avaliar o desempenho dos sistemas de classificação envolvidos.

4.1 Métodos de Treinamento Propostos

Conforme descrito na subseção 2.5.1 a informação lateral do chuveiro de partículas é uma grandeza comumente utilizada em experimentos de HEP, para apoiar o processo discriminativo entre partículas de natureza eletromagnética e hadrônica [20, 22]. O ATLAS utiliza diversas grandezas que descrevem o perfil lateral e longitudinal (ver Tabela 2.5), sendo estas grandezas usadas para fins de identificação de partículas.

No HLT (*High Level Trigger*), o algoritmo *NeuralRinger* captura as informações lateral e longitudinal do chuveiro desenvolvido no sistema de calorimetria, com uma geometria aproximadamente cônica que revela características do espalhamento e profundidade do chuveiro de partículas. Um modo de destacar a informação lateral (altamente discriminante) pode ser feito ponderando os anéis gerados em torno da célula mais energética pela sua energia e distância da região de interesse (RoI). Esse procedimento pode revelar características discriminativas entre o chuveiro de partículas produzido por elétrons ou jatos, apoiando-se no fato de que o espalhamento lateral da energia depositada por jatos, é maior do que o espalhamento lateral de elétrons [21]. Considerando estes aspectos, este trabalho propõe a adição de uma etapa de processamento no fluxo de execução do *NeuralRinger*, com o objetivo de ponderar os anéis por sua energia e distância ao primeiro anel, permitindo combinar a capacidade discriminativa dos anéis com a informação lateral do chuveiro. O diagrama simplificado do sistema proposto é indicado na Figura 4.1.

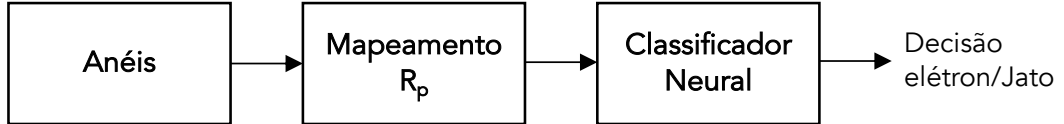


Figura 4.1: Fluxo de processamento do sistema de classificação proposto.

O mapeamento R_p multidimensional é realizado conforme descrito pela Equação 4.1):

$$R_p : X \rightarrow Y,$$

$$A_{Rpi}(E_i, r_i) = \frac{E_i^\alpha r_i^\beta}{\sum_{i=1}^{100} E_i^\alpha}, \quad (4.1)$$

em que, E_i é a energia do i -ésimo anel e r_i representa a distância unitária de cada anel em relação a região de interesse da camada do calorímetro, conforme indicado na Figura 4.2. Os coeficientes α e β são utilizados para ponderar a influência da energia¹ e do raio do anel no mapeamento. Para efetuar o treinamento das redes

¹Devido a desvios nos algoritmos de estimação de energia nas primeiras etapas de contabilização

neurais e otimizar simultaneamente² os parâmetros α e β , foram desenvolvidos métodos de treinamento utilizando algoritmos bio-inspirados e um método baseado em retropropagação.

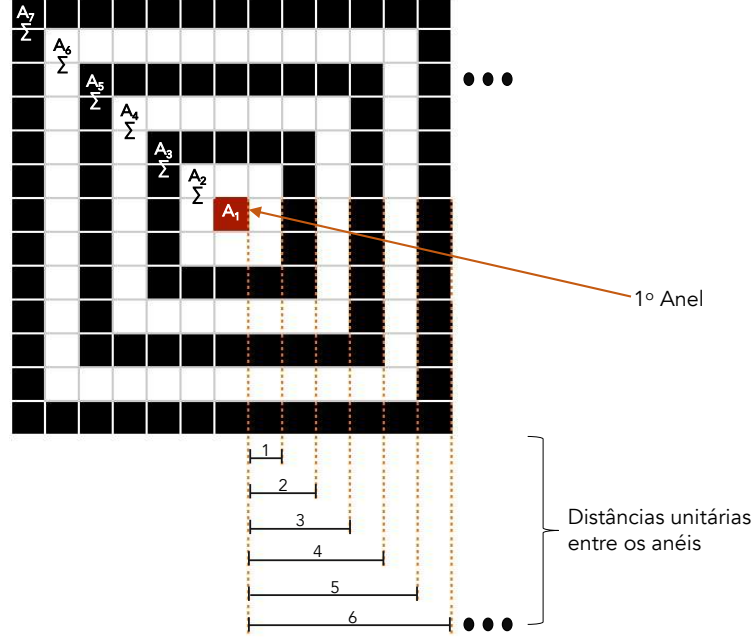


Figura 4.2: Ilustração dos anéis gerados em uma das camadas do calorímetro EM, com a indicação das distâncias unitárias entre os anéis concêntricos.

4.1.1 Treinamento usando Algoritmos Bio-Inspirados

Conforme ilustrado pela Figura 4.3, os sinais em anéis são processados pelo mapeamento R_p , que alimentam o classificador neural. Os pesos do classificador neural e os coeficientes α e β são otimizados a cada iteração por um algoritmo de otimização implementado.

O método proposto realiza o treinamento dos classificadores minimizando uma função objetivo projetada para otimizar o desempenho dos classificadores para pontos de operação específicos. Estas restrições de desempenho foram codificadas na da energia depositada nas células do calorímetro, algumas células podem apresentar valores negativos de energia [21]. Neste caso, para evitar que o mapeamento R_p possa eventualmente produzir números complexos, o valor 0 é atribuído ao anel que possivelmente apresente valores negativos de energia.

²O ajuste simultâneo dos parâmetros da etapa de pré-processamento com os pesos da rede neural evita que os coeficientes α e β sejam definidos por algum método de busca exaustiva, que exigiria maior tempo de processamento e custo computacional, pois seria necessário o treinamento dos classificadores para cada par de coeficientes em uma grade de busca desses parâmetros.

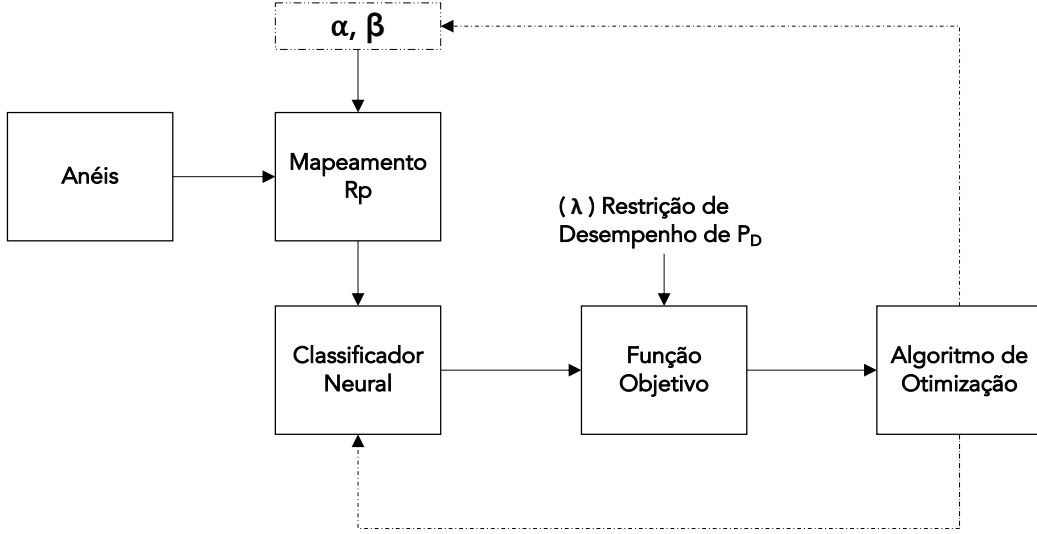


Figura 4.3: Ilustração do fluxo de processamento para computação da função objetivo e otimização dos parâmetros da etapa de pré-processamento.

função objetivo, descrita pela Equação 4.2:

$$\begin{aligned} \min f_t &= P_F(y, \gamma) + D_{KL}(y, T) + \alpha_r \Omega(\omega) \\ \text{Sujeito a } &\left\{ P_D(y, \gamma) \geq \lambda, \right. \end{aligned} \quad (4.2)$$

onde λ é o valor mínimo aceitável de P_D , usando o patamar de decisão (γ) ajustado para satisfazer o desempenho mínimo requerido. A divergência KL ($D_{KL}(y; T)$)³ é apresentada na Equação 4.3 para M exemplos de treinamento:

$$D_{KL}(y, T) = \sum_{m=1}^M y_m \times \ln \left(\frac{y_m}{T_m} \right). \quad (4.3)$$

A minimização da divergência KL simultaneamente com a taxa de falso alarme busca reduzir as diferenças estatísticas entre y e T , ponderando um melhor desempenho do classificador para operação em uma desejada taxa de P_D . O parâmetro de regularização (norma l_2) é adicionado na função objetivo visando a melhoria da generalização, sendo computado pela Equação 4.4:

$$\alpha_r \Omega(\omega) = \alpha_r \frac{1}{2} \|\mathbf{W}\|_2^2 = \alpha_r \sum_{v=1}^V \omega_v^2, \quad (4.4)$$

³A divergência de Kullback–Leibler de B dado A, indicada com $D_{KL}(A||B)$, é a medida da informação perdida quando B é usada para aproximar o valor de A. Uma divergência KL igual a 0 indica que as distribuições A e B são muito parecidas, enquanto uma divergência KL com valor igual a 1 indica que se comportam de maneira diferente [120].

onde ω é o vetor de pesos da rede neural e α_r é uma constante de regularização. O pseudo-código indicado no Algoritmo 1 descreve a função objetivo proposta para otimizar uma população de L classificadores neurais:

Algoritmo 1 Função Objetivo (f_t)

```

1: for  $l = 1$  to  $L$  do
2:   Computar as saídas ( $y_m$ ) com os sinais de treinamento  $\mathbf{x}_m$ ;
3:   Ajustar  $\gamma^{(l)}$  do  $l$ -ésimo classificador, para satisfazer  $P_D \geq \lambda$ ;
4:   Computar  $P_F^{(l)}$  com o patamar  $\gamma^{(l)}$  ajustado;
5:   Computar  $D_{KL}^{(l)}(y, T)$ ;
6:   Computar  $\alpha_r \Omega(\omega)^{(l)}$ ;
7:   Atribuir:  $f_T^{(l)} = P_F^{(l)}(y, \gamma) + D_{KL}^{(l)}(y, T) + \alpha \Omega^{(l)}(\omega)$ ;
8: end for

```

As especificações gerais de treinamento que se referem a critérios de parada, número de inicializações, funções de ativação, etc, foram definidas de modo heurístico, analisando combinações para a maximização do índice SP . Dois critérios de parada do treinamento foram adotados, um relacionado há um número máximo de épocas de treinamento da rede neural, e outro relacionado com um número máximo de iterações sem melhoria do índice SP . As redes neurais foram definidas com uma única camada oculta, função de ativação tangente hiperbólica, sendo que a quantidade de neurônios presentes na camada oculta foi obtida pelo SP máximo encontrado, variando o número de neurônios entre 2 e 10. Considerando a divisão dos intervalos de E_T e η apresentados na Tabela 2.4 (seção 2.5), um total de 25 classificadores neurais são treinados para cada intervalo.

Tabela 4.1: Resumo dos principais parâmetros utilizados no treinamento dos classificadores.

Descrição	Parâmetros
Validação Cruzada	<i>Stratified k-fold</i>
Alvos	+1/-1
Critério de Parada	25 sucessivas falhas de melhorias em SP
Número de Épocas	max. 1000
Número de Subconjuntos	10 (9 para treino, 1 para teste)
Número de Camadas Ocultas	1
Função de ativação	Tangente Hiperbólica
Arquitetura	Completamente conectada

Como propósito de comparação de desempenho, três diferentes algoritmos de

otimização. Algoritmo Genético (GA - *Genetic Algorithm*), Otimização Por Enxame de Partículas (PSO - *Particle Swarm Optimization*) e Estratégia Evolutiva com Adaptação da Matriz de Covariância (CMA-ES - *Covariance Matrix Adaptation - Evolution Strategy*) foram adaptados para a otimização da função objetivo descrita pela Equação 4.2. Uma solução do problema de otimização, que compõe a população a ser adaptada pelos algoritmos de otimização desenvolvidos é composta pelos pesos da camada oculta e de camada de saída, bem como os coeficientes α e β que ponderam a energia e a distância dos anéis ao primeiro anel, conforme indicado na Figura 4.4.

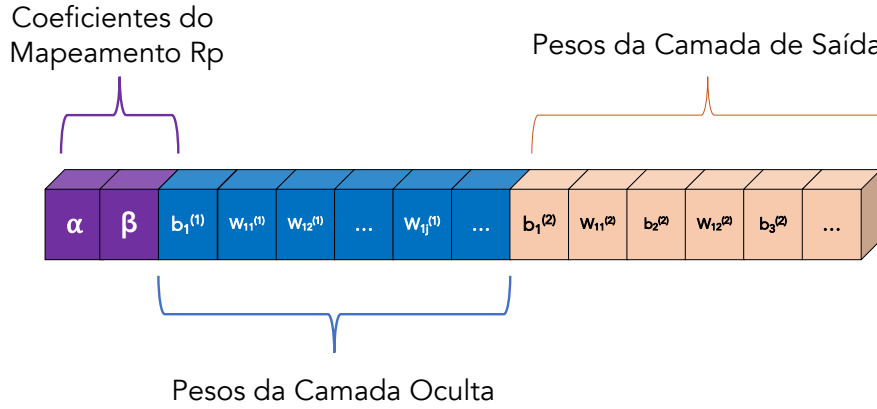


Figura 4.4: Composição das variáveis de um indivíduo (rede neural) a ser otimizado pelos algoritmos bio-inspirados.

Treinamento usando Algoritmo Genético

O GA implementado para treinamento dos classificadores considera o vetor de pesos da rede neural e os coeficientes do mapeamento R_p como o cromossomo a ser modificado durante o processo evolutivo, onde os genes são representados por cada variável do problema de otimização. O fluxo de execução do AG implementado pode ser visualizado na Figura 4.5.

Uma população de indivíduos é inicializada aleatoriamente, usando o método de inicialização proposto por Nyugen-Widrow [121]. Após a inicialização da população de RNAs, é calculado o valor da função objetivo para todos os indivíduos. De acordo com valor encontrado para a função objetivo os indivíduos são organizados de forma decrescente, para então serem submetidos aos processos de seleção e reprodução. Após este procedimento os 5% melhores cromossomos da população são reservados

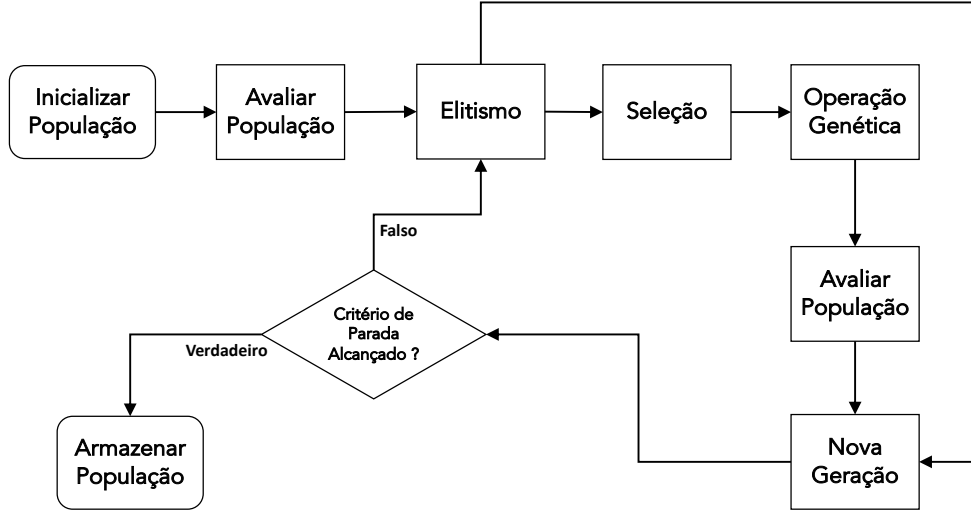


Figura 4.5: Ilustração do fluxo de execução método de treinamento proposto, utilizando o Algoritmo Genético desenvolvido.

para participar da próxima geração (elitismo). Baseado no valor encontrado para a função objetivo, são utilizadas 04 estratégias de seleção dos cromossomos que serão reservados como genitores nas operações genéticas. As estratégias utilizadas para seleção dos cromossomos genitores foram: torneio, aleatória, truncamento e roleta viciada. A escolha da estratégia de seleção é escolhida através de um sorteio aleatório a cada iteração, no qual cada mecanismo possui a mesma probabilidade de ser sorteado (probabilidade de $\frac{1}{4}$ para cada operador).

Os cromossomos selecionados são submetidos a operações genética de mutação⁴ e reprodução. Nesta etapa, os tipos de operadores de cruzamento são selecionados de modo aleatório, onde cada operador tem a mesma chance de ser sorteado (conforme indicado na Tabela 4.2⁵). Para verificar se a população apresenta convergência prematura um índice de variabilidade da população é calculado a cada geração (i_v), através da equação 4.5:

$$i_v(i) = \frac{\overline{f_T(y)}}{\max(f_T(y))} \quad (4.5)$$

em que $\overline{f_T(y)}$ é a média dos valores da função objetivo dos indivíduos da população, e $\max(f_T(y))$ seu valor máximo encontrado na atual iteração. Caso o índice i_v atinja o valor máximo ($i_v = 1$), 5% dos indivíduos são selecionados aleatoriamente,

⁴A operação de mutação consiste em somar um numero aleatório (entre -1 e 1) sorteado para cada gene a ser afetado, sendo este operador aplicado após a reprodução em cada iteração. A taxa de mutação (quantidade de genes afetados) é sorteada aleatoriamente entre 5% e 20%.

⁵Os hiperparâmetros adotados neste trabalho foram definidos através de experimentos, sendo que os mencionados aqui consistem na configuração que obteve os melhores resultados.

e novos cromossomos são inicializados e inseridos na população. Este procedimento foi implementado para permitir uma melhor exploração do espaço de busca das soluções, de modo à aumentar a quantidade de cromossomos distintos quando os valores da função objetivo de toda a população apresentem pouca diferença.

Os cromossomos resultantes das operações genéticas são avaliados pela função objetivo, utilizando o subconjunto de sinais reservados para treinamento. Caso o critério de parada seja alcançado, a última geração de cromossomos é armazenada e o indivíduo que apresenta melhor índice SP quando submetido aos dados de treinamento é então selecionado⁶. Caso contrário o algoritmo segue para uma próxima iteração. As características gerais do GA implementado são apresentadas na Tabela 4.2

Tabela 4.2: Características do GA implementado.

Descrição	Parâmetros
Codificação	Real
Taxa de Mutação	5%, 10%, 15%, 20%
Taxa de Elitismo	5%
Tamanho da População	100 indivíduos
Operador de Cruzamento	2pt, 4pt, 8pt, 16pt, uniforme
Operador de Seleção	Roleta, Torneio, Truncamento e Aleatório

Treinamento Usando Otimização por Enxame de Partículas

Um algoritmo de treinamento de redes neurais usando PSO foi implementado conforme ilustrado pela Figura 4.6. Cada partícula inicializada é composta por três atributos: posição, velocidade, $\mathbf{p}_{best}$. Os pesos do classificador neural e os coeficientes α e β compõem um vetor de posição (ver Figura 4.4) e são inicializados aleatoriamente [121]. Os vetores velocidade e $\mathbf{p}_{best}$ são inicializados com valor zero. Em seguida calcula-se o valor da função objetivo de cada partícula do enxame, usando o subconjunto dos sinais reservados para treinamento. Cada partícula tem sua velocidade computada (conforme Equação 3.30) e posição atualizada (conforme Equação 3.31).

Caso a partícula obtenha melhoria de desempenho em relação ao encontrado na iteração anterior o vetor $\mathbf{p}_{best}$ é atualizado, recebendo então os valores atuais do vetor de posição. Caso alguma partícula supere o melhor desempenho encontrado pelo enxame até então o vetor $\mathbf{g}_{best}$ é atualizado.

⁶O procedimento de separação dos conjuntos de treino e teste é apresentado na seção 4.1.3.

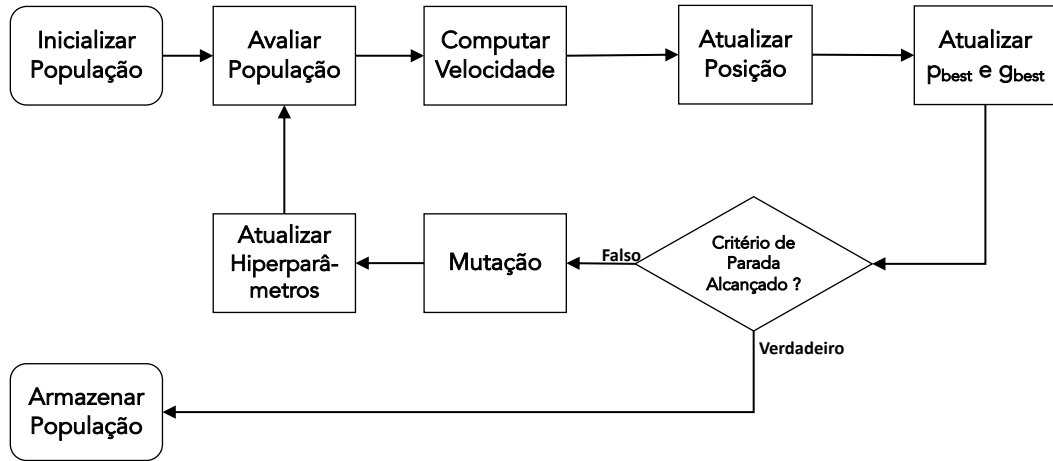


Figura 4.6: Ilustração do fluxo de execução método de treinamento proposto utilizando o algoritmo PSO implementado.

Caso o critério de parada não seja satisfeito, um operador de mutação é aplicado ao enxame com uma taxa de mutação variando entre 5% e 20% (taxa sorteada aleatoriamente). Para tentar explorar melhor o espaço de busca, durante as iterações o índice de variabilidade i_v (Equação: 4.5) é computado e caso alcance valor máximo, 5% dos indivíduos do enxame são mantidos e os demais substituídos por novos indivíduos, que são gerados aleatoriamente.

Em seguida para a próxima iteração, o coeficiente de inércia é atualizado (conforme equação 3.32) e as constantes de aceleração são sorteadas aleatoriamente, com valores no intervalo entre 0 e 1. As características do algoritmo PSO desenvolvido são mostradas na Tabela 4.3.

Tabela 4.3: Parâmetros do algoritmo PSO implementado.

Descrição	Parâmetros
Topologia de Conexão entre as Partículas	Global
Constante de Cognição (c_1)	2
Constante Social (c_2)	1
Tamanho do Enxame	100
Constantes de Aceleração (r_1, r_2)	Aleatória [0;1]
Coeficiente de inércia inicial (w_1)	0,9
Coeficiente de inércia final (w_2)	0,4
Taxa de Mutação	5%, 10%, 15%, 20%

Treinamento usando CMA-ES

Uma versão do método de treinamento proposto usando o algoritmo de otimização CMA-ES foi implementada, conforme indicado pela Figura 4.7. A principal motivação para utilização deste método é devido ao seu mecanismo otimizador preservar as relações entre as variáveis do problema de otimização (blocos construtivos) [85], que no caso de redes neurais tal dependência é presente devido a conexão massiva entre os neurônios.

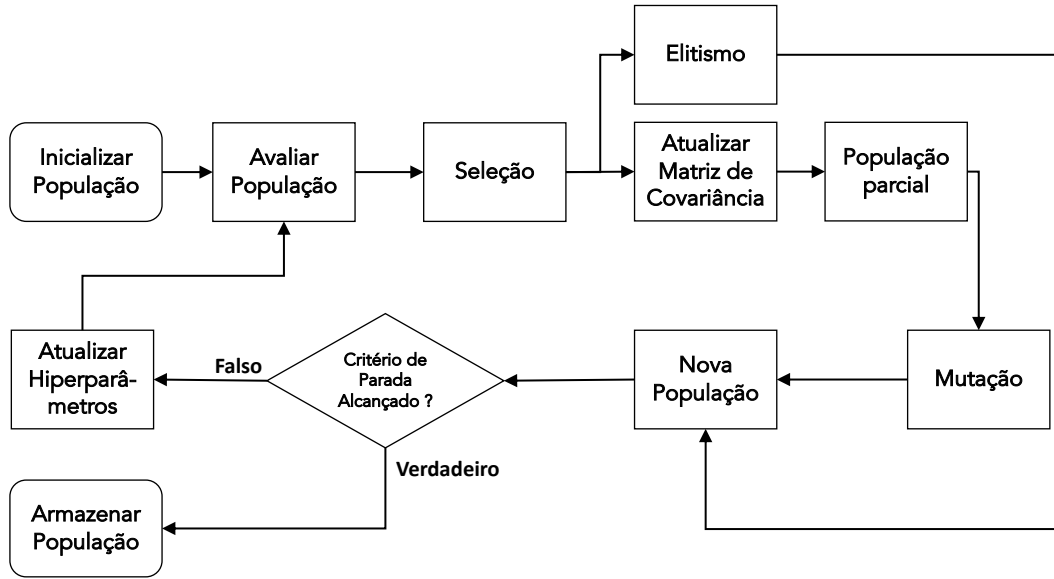


Figura 4.7: Ilustração do fluxo de execução do método de treinamento proposto com o algoritmo CMA-ES.

O procedimento inicializa uma população de λ candidatos, onde as variáveis do problema de otimização (pesos da rede neural e coeficientes α e β do mapeamento R_p) são inicializadas pelo método proposto em [121]. A estrutura dos candidatos é a mesma utilizada nas versões do método proposto usando os otimizadores AG e PSO, representada pela Figura 4.4.

Após a inicialização e avaliação da população pela função objetivo (Equação 4.2) usando o subconjunto de dados reservado para treinamento os indivíduos são organizados de modo crescente em função da sua pontuação na função objetivo, onde os μ melhores indivíduos são selecionados para participar do procedimento de atualização da matriz de covariância. Os 5% melhores indivíduos da população são reservados para participar da próxima geração (elitismo).

A matriz de covariância é então atualizada por meio do procedimento descrito na subseção 3.2.1, considerando os indivíduos selecionados. Com a matriz de cova-

riância atualizada, λ indivíduos são gerados, e em seguida submetidos ao operador de mutação, com a taxa de mutação é sorteada aleatoriamente a cada geração, conforme indicado na Tabela 4.4. Caso seja satisfeito o critério de parada o método é interrompido e a população de classificadores é então armazenada. Caso o critério de parada não seja satisfeito, os hiperparâmetros do algoritmo são atualizados e a função objetivo é avaliada novamente, seguindo para a próxima iteração do algoritmo. A Tabela 4.4 apresenta as principais características do algoritmo CMA-ES implementado.

Tabela 4.4: Parâmetros do algoritmo CMA-ES implementado.

Descrição	Parâmetros
Método de Inicialização da População	Nguyen-Widrow [121]
Tamanho da População (λ)	100
Quantidade de indivíduos selecionados (μ)	20
Máximo número de iterações	1000
Máximo número de sucessivas falhas de melhoria do SP	50
Intervalo das variáveis	[-4 ; 4]
Taxa de Mutação	5%, 10%, 15%, 20%

4.1.2 Treinamento usando *Backpropagation*

Foi desenvolvido um sistema integrado conforme indicado na Figura 4.8, usando o mapeamento R_p multidimensional com o classificador neural. Nesta estratégia o algoritmo de treinamento da rede (neste caso *backpropagation*) além de ajustar os pesos da rede, ajusta também os coeficientes α e β . Este conceito também apresenta a vantagem do próprio classificador neural encontrar o melhor perfil de ponderação da função $(E_i^\alpha r_i^\beta)$, desta vez usando um algoritmo de busca local [22].

Os pesos da rede e os coeficientes α e β do mapeamento R_p são inicializados usando o método Nguyen-Widrow [121]. Dois critérios de parada foram estabelecidos, um associado a um número máximo de épocas de treinamento (1000) e outro associado a um número máximo de falhas na melhoria do SP (25). A função de ativação tangente hiperbólica foi adotada em todos os neurônios da rede e a função objetivo utilizada foi o erro quadrático médio. A Tabela 4.5 apresenta o resumo dos principais parâmetros adotados nesta estratégia de treinamento.

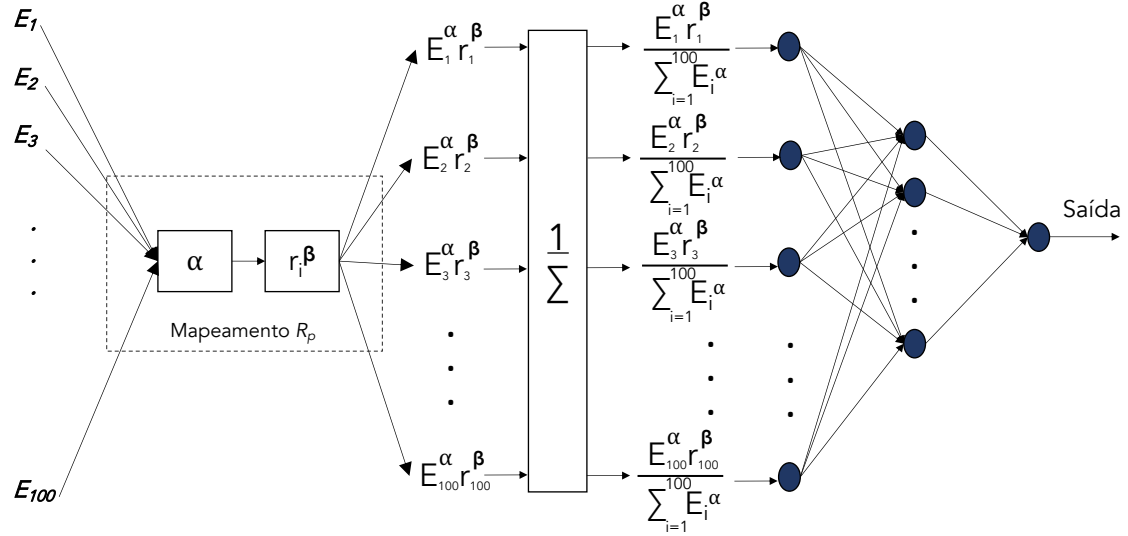


Figura 4.8: Ilustração do mapeamento R_p multidimensional aplicado aos sinais em anéis. A energia e a distância ao centro de cada anel é ponderada com os coeficientes α e β respectivamente. Estes coeficientes são ajustados juntamente com os pesos da rede neural, a cada iteração de treinamento com o algoritmo *backpropagation*.

Tabela 4.5: Resumo dos principais parâmetros utilizados para o treinamento dos classificadores neurais com *backpropagation*.

Descrição	Parâmetros
Algoritmo de Otimização	ADAM [122]
Taxa de Aprendizado	0.010
Critério de Parada	25 falhas sucessivas na melhoria do SP
Função Objetivo	Erro Quadrático Médio
Número Máximo de Épocas	1000
Tamanho da <i>Batch</i>	Número de exemplos da classe com menor quantidade
Inicialização dos Pesos	Nguyen-Widrow [121]
Número de Camadas Ocultas	1
Função de Ativação	Tangente Hiperbólica
Arquitetura	Completamente Conectada

4.1.3 Procedimento de Validação Cruzada

O modo como os conjuntos de dados são utilizados neste trabalho para o projeto dos classificadores neurais é apresentado na Figura 4.9.

O conjunto de dados experimentais composto por colisões ocorridas em todo o ano de 2017 é utilizado para as fases de treinamento e teste dos classificadores. O pro-

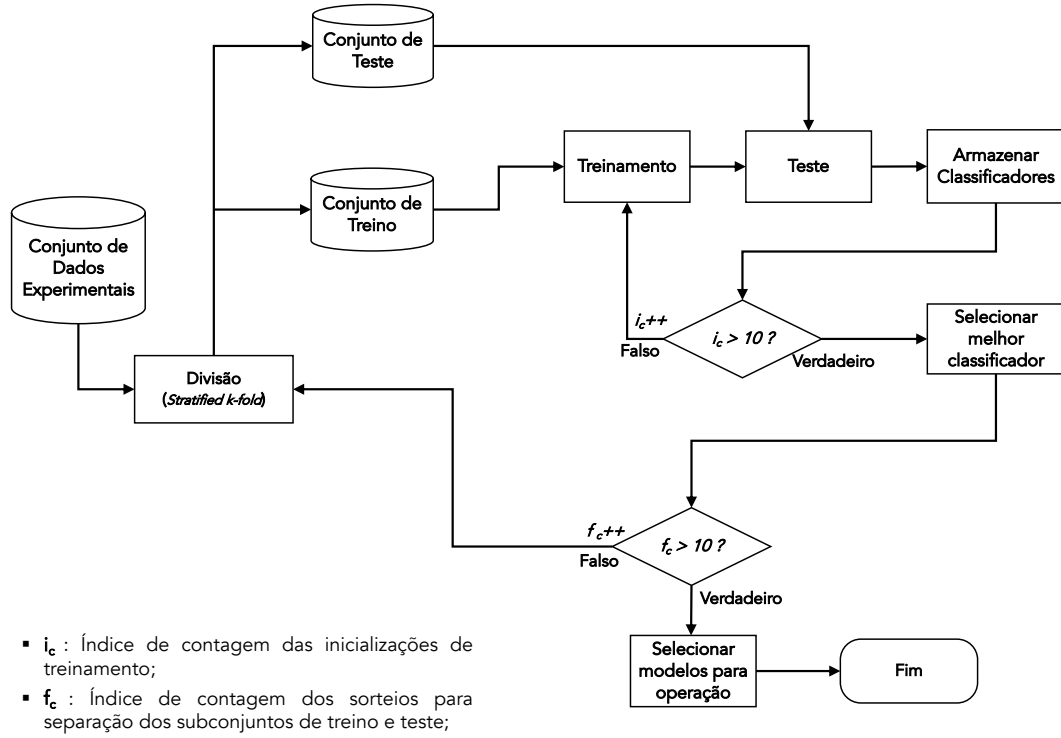


Figura 4.9: Fluxograma do procedimento de validação cruzada e utilização dos conjuntos de dados experimentais utilizados para o projeto dos classificadores neurais.

cedimento de validação cruzada adotado neste trabalho foi o *Stratified k-fold* [120], sendo aplicado para estimar flutuação estatística possivelmente presente nos conjuntos de dados, conforme descrito resumidamente a seguir:

1. Os sinais formatados em anéis, disponíveis para cada classe (assinaturas de elétrons/ruído de fundo) são divididos em 10 subgrupos (com aproximadamente a mesma quantidade de exemplos em cada subgrupo);
2. Sorteiam-se 9 subgrupos em cada classe para compor o conjunto de treino, e um subgrupo de cada classe para o conjunto de teste⁷;
3. Caso haja desbalanceamento do número de exemplos entre as classes, é realizado o procedimento de estratificação [120] (balanceamento do número de exemplos).
4. Com os subgrupos de treino e teste, o treinamento é realizado 10 vezes (para

⁷Neste trabalho, não foi utilizado um subconjunto de validação, de modo à oferecer uma maior quantidade de eventos para a fase de treinamento. Para contornar possíveis problemas de sobreajuste (*overfitting*), o conjunto de teste é avaliado ao longo do treinamento. Após o fim do treinamento, são analisadas as curvas de variação do erro médio quadrático e do índice SP, para verificar uma possível ocorrência do efeito de sobreajuste.

avaliar possível incidência de mínimos locais) e os resultados de cada treinamento e teste são armazenados;

5. Um próximo subgrupo (de cada classe) é selecionado para participar do conjunto de teste, e o subgrupo de teste selecionado anteriormente passa a integrar o conjunto de treinamento;
6. Repetem-se os itens 2, 3 e 4 até que todos os subgrupos sorteados de cada classe tenham participado do conjunto de teste;

Após finalização do procedimento de validação cruzada, 10 classificadores com melhor desempenho de índice SP (um modelo obtido a cada dez inicializações de treinamento) são então alimentados por todos os sinais disponíveis. O classificador que apresenta o maior índice SP quando submetido a todo o conjunto de dados é selecionado para a operação^[8].

4.2 Descrição dos Conjuntos de Dados Utilizados

Para validar a metodologia proposta neste trabalho, foram utilizados conjuntos de dados experimentais obtidos em colisões de prótons ocorridas no LHC no ano de 2017 (*Run2*). Dado que estes eventos são de natureza experimental, a determinação das classes a serem utilizadas no treinamento dos classificadores (elétrons ou ruído de fundo) é determinada pela decisão de critérios estabelecidos no algoritmo de identificação de partículas do ambiente *offline*^[9].

Estes critérios são especificados com a combinação de variáveis de calorimetria, permitindo diferentes graus de rejeição de ruído de fundo e eficiências de classificação de elétrons.

O grau de certeza da assinatura de elétrons aumenta progressivamente seguindo a ordem dos critérios: *VeryLoose*, *Loose*, *Medium*, *Tight*, onde por exemplo eventos aprovados pelo critério elétron-*Tight* tem maior probabilidade de serem realmente elétrons, enquanto aqueles aprovados pelo critério elétron-*VeryLoose* (se comparados aos aprovados pelo *Tight*) além de elétrons serem aprovados, uma maior quantidade de falsos elétrons são também admitidos.

⁸A seleção para operação se refere a escolha dos classificadores neurais que irão ser utilizados/integrados no ambiente computacional do HLT.

⁹O algoritmo de identificação de partículas do ambiente *offline* realiza a extração de estimativas das *pdfs* marginais para grandezas que descrevem o formato do chuveiro usando KDE Gaussiano [123]. Os *kernels* são aproximados por histogramas com fina granularidade que são interpolados para dar os valores de probabilidade do candidato pertencer a classe de sinal ou ruído de fundo para a i -ésima grandeza [10].

As aprovações pelo critério *Tight* apresentam maior exigência (análise em muitas variáveis de calorimetria), assim em geral, uma quantidade menor de eventos são coletados por este critério. Enquanto as aprovações pelo critério *VeryLoose* são realizadas com um menor grau exigência, consequentemente uma maior quantidade de eventos são aprovados por este critério^[10]. Estes critérios são utilizados em análises físicas, nos quais ora podem requisitar uma elevada quantidade de amostras ou podem requerer um grau elevado de certeza nas amostras coletadas.

No caso dos conjuntos de dados utilizados neste trabalho, os eventos que compõem a classe de sinal são elétrons provenientes do decaimento $Z \rightarrow ee$ ^[11] [18], que foram aprovados pelo critério de operação *Medium* do *offline*. Esta estratégia permite uma coleta maior de eventos de sinal (se comparado ao critério *Tight*) com pouca degradação na assinatura dos eventos classificados como elétrons.

Já os eventos utilizados para a classe de ruído de fundo, foram jatos hadrônicos reprovados pelo critério de operação *VeryLoose*. Esta estratégia permite coletar para o treinamento dos classificadores, jatos com menor probabilidade de serem elétrons.

A Figura 4.10 apresenta as distribuições em E_T (a) e pseudorapidez - η (b) dos anéis presentes no conjunto de dados. A Tabela 4.6 [18] apresenta dados das colisões do LHC, ocorridas em 2017.

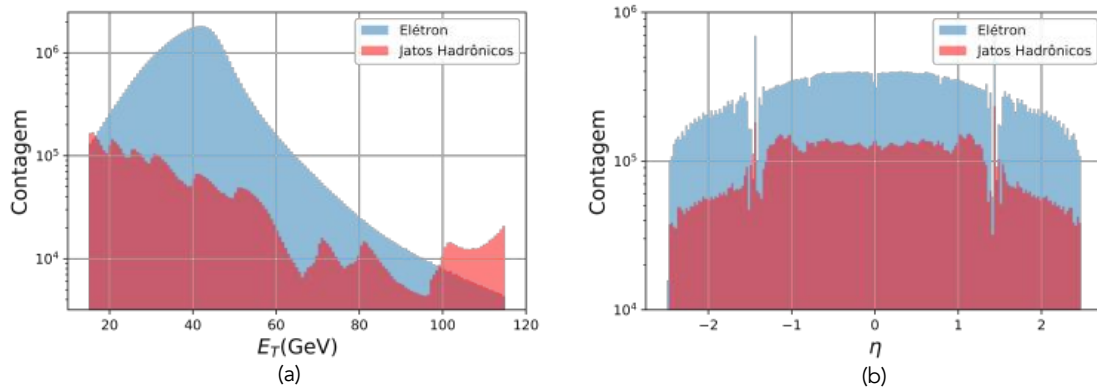


Figura 4.10: Distribuição dos conjuntos de dados experimentais de 2017, em função da energia (a) e da coordenada η do ATLAS.

Tabela 4.6: Dados da luminosidade de pico, luminosidade integrada, valores de pico e médios da taxa de empilhamento durante as colisões de prótons em 2017. Extraído de [18].

Ano	Lumi. de Pico ($cm^{-2}s^{-1}$)	Lumi. Integrada (fb^{-1})	$\langle\mu\rangle$ (pico)	$\langle\mu\rangle$ (médio)
2017	$2,1 \times 10^{34}$	43,9	80	37,8

¹⁰É esperado um aumento do índice de falso alarme com a utilização deste critério.

¹¹Decaimento dominante de elétrons para as faixas de energia analisadas.

Variações no perfil de deposição de energia podem ser observadas em diferentes intensidades de energia da partícula, e também com as regiões de interação do calorímetro (devido a mudança de granularidade das células sensoras ao longo da coordenada $|\eta|$). Na Figura 4.11 são apresentadas as distribuições da variável de calorimetria razão de energia hadrônica R_{HAD} (descrição detalhada na Tabela 2.5) para (a) elétrons e Jatos com $E_T \geq 50$ GeV e (b) $15 \leq E_T[\text{GeV}] < 20$ respectivamente. Conforme pode ser observado, elétrons com $E_T \geq 50$ GeV apresentam distribuição mais próxima de zero, se comparado aos elétrons com energia na faixa $15 \leq E_T[\text{GeV}] < 20$. Essa característica está relacionada com o aumento da absorção de energia pelo calorímetro eletromagnético, com o crescimento da energia de elétrons, o que provoca a diminuição do vazamento hadrônico para estas partículas. As diferentes flutuações dos perfis de deposição de energia em função de E_T e de $|\eta|$ motivaram o desenvolvimento do conjunto de classificadores especialistas em pares $E_T \times |\eta|$.

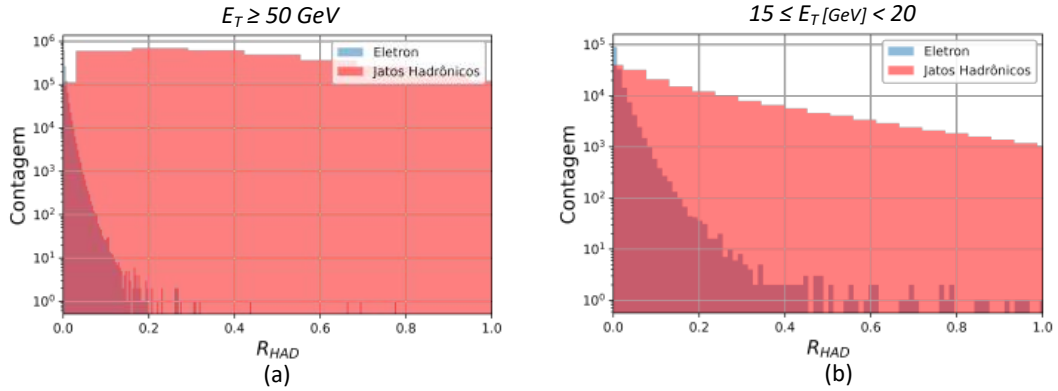


Figura 4.11: Razão de energia hadrônica (R_{HAD}) para (a) elétrons e Jatos com $E_T \geq 50$ GeV e (b) $15 \leq E_T[\text{GeV}] < 20$, do conjunto de dados experimentais de 2017.

Na Figura 4.12 é apresentado o perfil dos sinais em anéis (média e desvio padrão) para os conjuntos de dados de 2017 no intervalo de E_T em $[30 \leq E_T < 40]$ e de coordenada η em $[0, 00 \leq \eta < 0, 80]$ respectivamente. Nota-se que o perfil das assinaturas de elétrons apresenta deposição de energia principalmente nos primeiros anéis das camadas EM1 e EM2, enquanto a deposição de energia se mostra reduzida nas camadas hadrônicas. Nos perfis dos sinais de jatos, nota-se acentuada deposição de energia na camada EM2, assim como presença de energia nas camadas hadrônicas¹².

¹²O mapeamento topológico realizado pelo *NeuralRinger* permite que os classificadores neurais possam explorar as diferenças da estrutura cônica da deposição de energia de elétrons e jatos, ao longo das camadas do calorímetro. Entretanto, uma boa parte das assinaturas de jatos podem apresentar perfil de deposição de energia semelhante aos elétrons [21].

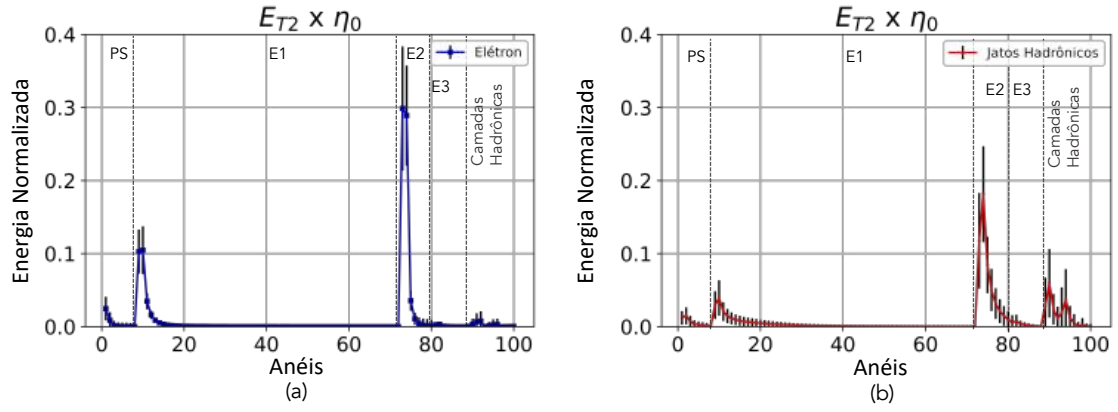


Figura 4.12: Sinais em anéis da faixa $E_{T2} \times \eta_0$ (média e desvio padrão) para elétrons (a) e jatos hadrônicos (b) do conjunto de dados experimentais de 2017.

A Tabela 4.7 apresenta as quantidades de exemplos para elétrons e jatos contidos nos conjuntos de dados de 2017, para cada faixa de E_T e $|\eta|$ utilizada para o projeto do conjunto de classificadores.

Tabela 4.7: Número de exemplos contidos no conjunto de dados experimentais de 2017 divididos por faixas de E_T e η (azul para elétrons e vermelho para jatos hadrônicos).

		E_{T0}	E_{T1}	E_{T2}	E_{T3}	E_{T4}
		$15 \leq E_T < 20$	$20 \leq E_T < 30$	$30 \leq E_T < 40$	$40 \leq E_T < 50$	$E_T \geq 50$
η_0	$0,00 \leq \eta < 0,80$	232.819	1.499.273	4.188.536	4.627.778	1.699.541
		187.182	315.773	223.347	162.034	3.758.845
η_1	$0,80 \leq \eta < 1,37$	141.000	896.615	2.539.137	2.837.736	1.059.922
		143.312	227.190	156.597	111.945	2.553.385
η_2	$1,37 \leq \eta < 1,54$	51.566	283.647	645.043	470.869	170.752
		29.985	47.427	35.127	25.389	559.409
η_3	$1,54 \leq \eta < 2,37$	227.345	1.132.030	2.377.013	2.429.079	901.075
		205.247	306.475	196.256	131.328	1.314.870
η_4	$2,37 \leq \eta < 2,50$	19.893	84.054	164.380	187.729	89.786
		15.474	36.432	24.917	16.056	93.532

4.3 Organização para Exibição dos Resultados

Os resultados obtidos com os métodos de classificação de elétrons avaliados neste trabalho, foram organizados para exibição conforme descrito a seguir. Inicialmente, os métodos com as notações listadas na Tabela 4.8 foram avaliados com o conjunto de dados da faixa $E_{T2} \times \eta_0$ ($[30 > E_T > 40] \times [0.00 < \eta < 0.80]$). Esta faixa foi

selecionada para as análises iniciais, pois em relação a energia, provê boa resolução para a resposta dos sensores, combinado com a coordenada η (barril central) que apresenta fina segmentação dos calorímetros.

Os métodos serão comparados com os resultados obtidos de dois métodos atualmente empregados pelo sistema de *trigger* do ATLAS, o *NeuralRinger* (método de classificação de referência) e o *CutBased* que foi o método padrão até 2017 para elétrons com $E_T > 15$ GeV. Os resultados do *CutBased* (método baseado em cortes lineares em variáveis de calorimetria) foram apresentados para manter-se a comparação dos principais métodos empregados pelo experimento.

Após os resultados avaliados na faixa $E_{T2} \times \eta_0$ os dois métodos propostos que obtiveram melhor desempenho quando comparados com os aqueles já empregados no ATLAS, foram analisados em todas as demais faixas de $E_T \times \eta$.

Tabela 4.8: Descrição da notação utilizada para indicar os métodos de classificação de elétrons com $E_T \geq 15$ GeV avaliados neste trabalho.

Notação	Breve Descrição do Método
NeuralRinger	Método atualmente empregado no HLT do ATLAS para elétrons.
CutBased	Método baseado cortes empregado no HLT do ATLAS para elétrons até 2017.
Ring-GA	Método de treinamento sem pré-processamento R_p usando Algoritmo Genético.
Ring-PSO	Método de treinamento sem pré-processamento R_p usando PSO.
Ring-CMAES	Método de treinamento sem pré-processamento R_p usando CMAES.
RingRp	Método de treinamento com mapeamento R_p usando <i>Backpropagation</i> .
RingRp-GA	Método de treinamento com mapeamento R_p usando Algoritmo Genético.
RingRp-PSO	Método de treinamento com mapeamento R_p usando PSO.
RingRp-CMAES	Método de treinamento com mapeamento R_p usando CMA-ES.

4.4 Parâmetros de Avaliação de Desempenho

Neste trabalho, para avaliar o desempenho do sistemas de classificação são utilizados o índice SP (soma-produto) [56], a taxa de acerto de classificação de elétrons (Probabilidade de Detecção) e a taxa de erro de classificação de falsos elétrons (Probabilidade de Falso Alarme). O valor máximo do índice SP , é uma medida de avaliação de desempenho utilizada para escolher o patamar de decisão que equilibra (em média) o máximo acerto de classificação para ambas as classes do discriminador binário, definido pela Equação 4.6:

$$SP = \sqrt{\frac{P_D + (1 - P_F)}{2}} \times \sqrt{P_D \times (1 - P_F)} \quad (4.6)$$

sendo P_D e P_F as probabilidades de detecção e falso alarme, calculadas por meio das Equações 3.13 e 3.14, respectivamente, para um determinado patamar de decisão.

O valor máximo do índice SP para um determinado classificador é encontrado por meio da variação do patamar de decisão em toda faixa de excursão (que no caso deste trabalho varia entre -1 e 1). Para expressar o desempenho dos sistemas de classificação avaliados serão apresentados para cada parâmetro SP , P_D e P_F o valor médio e o desvio padrão associados, considerando as inicializações de treinamento, realizadas pelo procedimento de validação cruzada adotado¹³.

Conforme discutido no capítulo 2, o LHC tem em seu plano de atualizações aumentar constantemente a luminosidade e a energia de colisão do centro de massa para explorar potencialmente a ocorrência de partículas ainda não verificadas experimentalmente. Entretanto isso provoca o aumento da taxa de empilhamento de sinais no sistema de calorimetria, prejudicando a medição das assinaturas dos eventos e consequentemente degradando a eficiência dos algoritmos de seleção de eventos. A ilustração da distorção no sinal característico (linha de cor magenta) provocada pelo empilhamento de sinais pode ser observada na Figura 4.13, que mostra dois pulsos (linhas de cor preta e vermelha) gerados por partículas que interagiram com o TileCal, espaçadas em 50 ns [124].

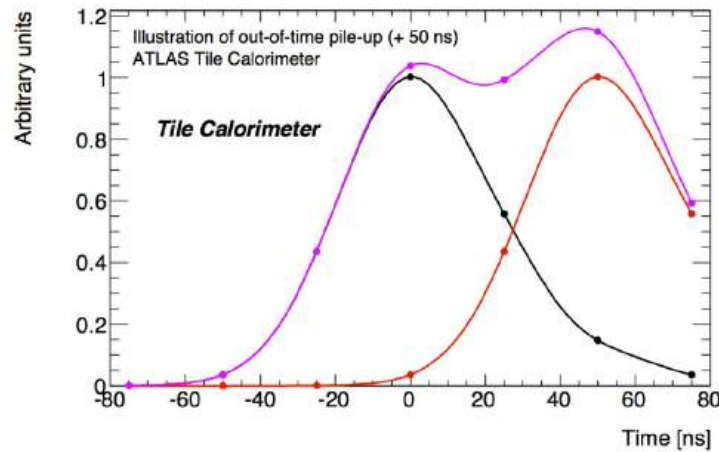


Figura 4.13: Ilustração do efeito do empilhamento (linha de cor magenta) provocado pela incidência de partículas que geraram os sinais em cor preta e vermelha, espaçadas em 50 ns. Extraído de [124].

No conjuntos de dados, o índice de empilhamento é variável, como por exemplo pode ser observado na Figura 4.14, onde são apresentadas as distribuições do índice de empilhamento para elétrons e jatos, do conjunto de dados da subdivisão do espaço de fase $E_{T2} \times \eta_0$. Neste contexto, o limiar de decisão dos classificadores precisa ser ajustado conforme a variação da taxa de empilhamento, de modo a manter sua eficiência de detecção de elétrons constante. Assim, após o treinamento

¹³Sendo a média e o desvio padrão contabilizados apenas pelos classificadores que obtiveram melhor desempenho a cada 10 inicializações no procedimento de validação cruzada.

os classificadores com maior índice SP são selecionados para realização do ajuste do limiar de decisão, em função do número médio de interações por cruzamento de feixes $\langle\mu\rangle$.

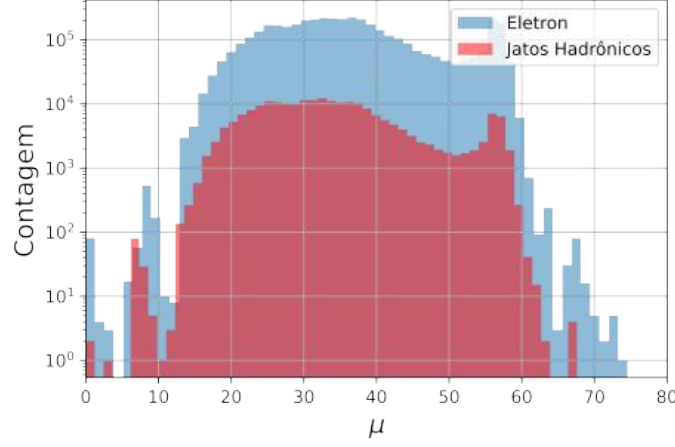


Figura 4.14: Distribuição do índice de empilhamento de sinais $\langle\mu\rangle$, para a subdivisão do espaço de fase $E_{T2} \times \eta_0$

Para realizar este ajuste, a função de ativação tangente hiperbólica do neurônio de saída é retirada, e substituída por uma função linear, em que os coeficientes a e b (Equação 4.7) são ajustados para manter a probabilidade de detecção semelhante a referência (taxa de P_D do *NeuralRinger*) em função de $\langle\mu\rangle$.

$$\delta(\langle\mu\rangle) = a \langle\mu\rangle + b, \quad (4.7)$$

em que a e b são os coeficientes da curva. Histogramas com os sinais de saída da rede (elétrons e ruído de fundo), com a curva linear ajustada são exibidos para mostrar o impacto do ajuste realizado em função da taxa de empilhamento.

Capítulo 5

Resultados

Neste capítulo são apresentados os resultados obtidos com a aplicação da metodologia proposta, visando melhorias na etapa rápida de seleção *online* de elétrons do ATLAS. Para o projeto dos classificadores, foram utilizados dados experimentais de colisões do LHC ocorridas no ano de 2017.

Uma etapa de pré-processamento especialista é proposta, com o intuito de destacar nos sinais de entrada os diferentes aspectos discriminantes da largura do chuveiro de partículas produzido por elétrons e por jatos hadrônicos. O método de classificação de elétrons de referência, empregado no HLT do ATLAS (NeuralRinger) é comparado com os métodos de treinamento de redes neurais propostos neste trabalho, que, de modo integrado, permitem realizar o treinamento dos pesos do classificador neural e simultaneamente otimizar os coeficientes da etapa de pré-processamento aos sinais em anéis.

Este capítulo está organizado do seguinte modo: inicialmente, foi avaliado o desempenho dos classificadores neurais, utilizando os sinais com intervalo de energia $30 \leq E_T[GeV] < 40$. Nesta faixa de energia, foram realizadas análises na região do barril central do ATLAS, região do *crack* e região mais externa da coordenada $|\eta|$. O objetivo desta análise foi investigar o impacto do pré-processamento proposto considerando a variação da granularidade das células em relação ao ponto de interação das partículas. Para cada método de treinamento, foram avaliados os desempenhos obtidos para os pontos de operação *Tight* e *VeryLoose*. Após esta análise, os métodos que apresentaram melhor desempenho foram selecionados para avaliação em todo o espaço de fase¹. É realizada uma discussão sobre os valores encontrados dos coeficientes que ponderam a energia e a distância dos anéis ao centro. Uma análise de eficiência geral é realizada, para verificar o impacto dos métodos de treinamento usando todos os sinais disponíveis nos conjuntos de dados utilizados. Os classi-

¹As partições do espaço de fase consistem nos pares de intervalos de energia e pontos de interação da partícula na coordenada η do ATLAS ($E_T \times \eta$), nos quais o *ensemble* de redes neurais atua na etapa rápida do HLT.

ficadores que obtiveram melhor desempenho em cada ponto de operação avaliado foram selecionados para o ajuste da função de ativação do neurônio de saída em função do empilhamento de sinais. Este procedimento foi realizado para encontrar um mapeamento linear, que mantenha a mesma taxa de probabilidade de detecção do classificador neural, considerando as variações do índice de empilhamento. Os impactos do ajuste pelo índice de empilhamento, também são discutidos para os pontos de operação *Tight* e *VeryLoose*.

5.1 Comparação de Desempenho entre os Métodos de Treinamento

Para realizar a comparação do desempenho encontrado pelos métodos de treinamento avaliados neste trabalho, foi definida uma série de etapas. Para verificar a melhor quantidade de neurônios a serem utilizados na camada oculta, variou-se o número de neurônios ocultos entre 2 e 10. Os métodos de treinamento que utilizam algoritmos bio-inspirados, foram executados de dois modos distintos. No primeiro modo, o treinamento foi realizado sem a etapa de pré-processamento R_p (chamados de Ring-GA, Ring-PSO, Ring-CMAES). No segundo modo, o pré-processamento R_p foi incluído, onde os coeficientes α e β , que ponderam a energia e a distância do anel ao centro (ver Equação 4.1), foram ajustados juntamente com os pesos do classificador neural (métodos aqui chamados de RingRp-GA, RingRp-PSO, RingRp-CMAES). Os treinamentos foram realizados para os pontos de operação *Tight* e *VeryLoose*, utilizando a taxa de falso alarme como critério de parada.

No caso do método de treinamento proposto baseado em retropropagação (RingRp), também foi realizado o procedimento de análise do número de neurônios na camada oculta. Conforme descrito na subseção 4.1.2, neste método o índice SP foi adotado como critério de parada e após o treinamento, foram selecionados os classificadores com melhor desempenho obtido para os pontos de operação *Tight* e *VeryLoose*.

A Figura 5.1 apresenta para cada método de treinamento bio-inspirado, os *boxplots* com a taxa de falso alarme obtida no ponto de operação *Tight* ($P_D[\%] = 99,83$). Cada *boxplot* apresenta o melhor desempenho encontrado entre as inicializações de treinamento, nos 10 sorteios do procedimento de validação cruzada, usando o conjunto de dados de teste da subdivisão do espaço de fase $E_{T2} \times \eta_0$. Na Figura 5.1 (a), são apresentados os resultados da taxa de falso alarme em função no número de neurônios, com os métodos Ring-GA e RingRp-GA. Uma redução do falso alarme é notada com a adoção da etapa de pré-processamento, com o método RingRp-GA encontrando a menor média com 9 neurônios na camada oculta ($P_F[\%] = 0,98 \pm 0,07$),

enquanto o Ring-GA apresenta menor média de falso alarme, utilizando 8 neurônios ocultos ($P_F[\%] = 1,09 \pm 0,05$).

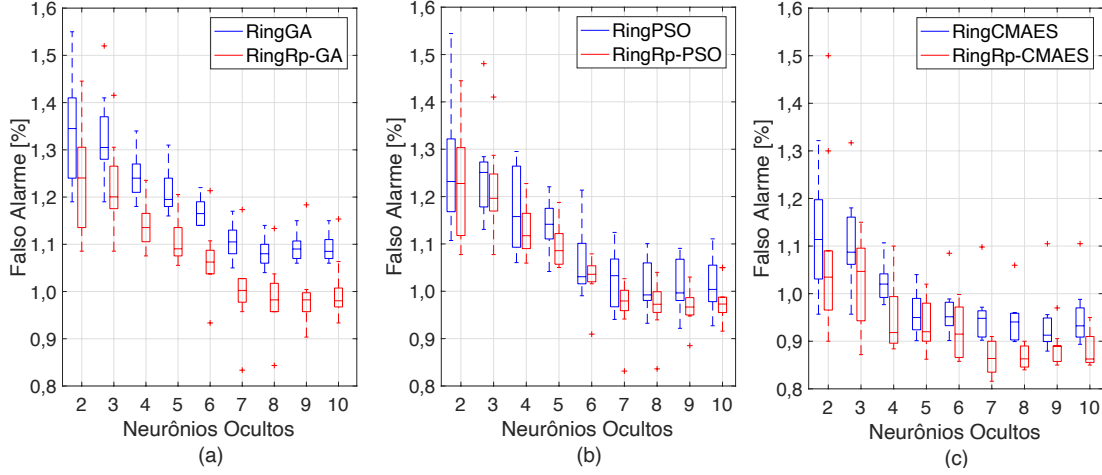


Figura 5.1: *Boxplot* da taxa de falso alarme para o ponto de operação *Tight* ($P_D[\%] = 99,83$), encontrada para cada método de treinamento bio-inspirado avaliado, em função do número de neurônios na camada oculta. Resultados obtidos usando o conjunto de dados da subdivisão do espaço de fase $E_{T2} \times \eta_0$.

Na Figura 5.1 (b), são apresentados os resultados da taxa de falso alarme com os métodos Ring-PSO e RingRp-PSO, em função no número de neurônios ocultos. Semelhante ao observado com os métodos baseado em GA, não há redução significativa da taxa de falso alarme a partir de 7 neurônios na camada oculta. A taxa de falso alarme apresenta redução com método de treinamento com a etapa de pré-processamento, onde o método RingRp-PSO encontra a menor média com 9 neurônios na camada oculta ($P_F[\%] = 0,97 \pm 0,09$), enquanto o Ring-PSO apresenta menor média de falso alarme, utilizando 8 neurônios ocultos ($P_F[\%] = 1,01 \pm 0,09$).

Dentre os algoritmos bio-inspirados avaliados, os métodos Ring-CMAES e RingRp-CMAES, foram os que encontraram melhor desempenho. Conforme indicado na Figura 5.1 (c), observa-se que, ao adotar a etapa pré-processamento R_p , a taxa de falso alarme apresenta redução, onde o método RingRp-CMAES obteve em média $P_F[\%] = 0,87 \pm 0,09$, usando 8 neurônios na camada oculta. O método Ring-CMAES alcançou $P_F[\%] = 0,92 \pm 0,03$, com a adoção de 9 neurônios ocultos.

A Figura 5.2 apresenta os *boxplots* da taxa de falso alarme, em função do número de neurônios na camada oculta, obtidos pelo método de treinamento RingRp no ponto de operação *Tight*. Nota-se que com 7 neurônios na camada oculta, obtém-se a menor média da taxa de falso alarme ($P_F[\%] = 0,82 \pm 0,04$), se comparado às demais quantidades de neurônios avaliadas. Percebe-se um ligeiro aumento na dispersão do falso alarme, para uma quantidade de neurônios ocultos acima de 7.

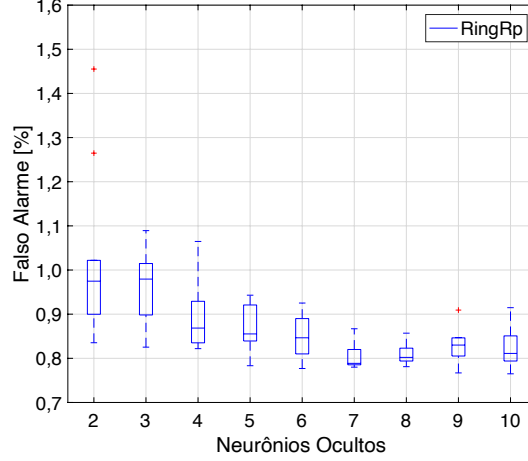


Figura 5.2: *Boxplot* da taxa de falso alarme para o ponto de operação *Tight* ($P_D[\%] = 99,83$), encontrada para o método RingRp, em função do número de neurônios na camada oculta. Resultados obtidos usando o conjunto de dados da subdivisão do espaço de fase $E_{T2} \times \eta_0$.

A Tabela 5.1 resume os resultados encontrados (média e desvio padrão) pelos métodos de treinamento avaliados. São apresentados os resultados obtidos para os índices $SP[\%]$ e $P_F[\%]$, computados para a mesma taxa de $P_D[\%]$ na qual o NeuralRinger opera, no ponto de operação *Tight* na etapa rápida do HLT.

Tabela 5.1: Comparação de desempenho entre os classificadores avaliados para o ponto de operação *Tight*. Os classificadores foram treinados para a subdivisão do espaço de fase $E_{T2} \times \eta_0$.

Discriminador	SP [%]	P _D [%]	P _F [%]	Neurônios Ocultos
NeuralRinger (Ref.)	99,11 ± 0,06	99,18 ± 0,05	0,96 ± 0,04	5
CutBased	95,94	99,18	7,24	-
Ring-GA	99,05 ± 0,06	99,18 ± 0,06	1,09 ± 0,05	8
Ring-PSO	99,08 ± 0,08	99,18 ± 0,09	1,01 ± 0,09	7
Ring-CMAES	99,11 ± 0,03	99,18 ± 0,03	0,92 ± 0,03	9
RingRp	99,16 ± 0,04	99,18 ± 0,04	0,82 ± 0,04	7
RingRp-GA	99,09 ± 0,06	99,18 ± 0,06	0,99 ± 0,06	9
RingRp-PSO	99,10 ± 0,09	99,18 ± 0,09	0,97 ± 0,09	7
RingRp-CMAES	99,14 ± 0,05	99,18 ± 0,05	0,87 ± 0,05	9

Os classificadores treinados com o algoritmo Ring-CMAES apresentaram uma redução de 0,04 ponto percentual na média da taxa de falso alarme, quando comparado com o NeuralRinger. Quando o pré-processamento R_p foi adotado, o método RingRp-CMAES apresentou melhor desempenho entre os algoritmos bio-inspirados projetados, alcançando uma redução de 0,09 ponto percentual na taxa de falso

alarme se comparado ao resultado obtido pelo NeuralRinger.

Dentre os métodos propostos, o método de treinamento RingRp apresentou desempenho com maior rejeição ao ruído de fundo, alcançando a taxa de $P_F = 0,82\% \pm 0,04$, que equivale a redução de 0,14 ponto percentual em relação ao NeuralRinger. A redução da taxa de falso alarme no ponto de operação *Tight* trata-se de um resultado importante para o experimento, visto que este ponto de operação é usado principalmente para selecionar eventos com maior probabilidade de serem realmente elétrons, ou seja, produz maior rejeição à falsos elétrons.

Na Tabela 5.2 são apresentados os resultados dos classificadores avaliados na subdivisão do espaço de fase $E_{T2} \times \eta_0$, para o ponto de operação *VeryLoose*. Este ponto de operação permite uma maior eficiência na coleta de elétrons, entretanto permite também uma maior aprovação de falsos elétrons, quando comparado ao ponto de operação *Tight*. Neste contexto, a redução do falso alarme é importante para permitir que a maior parte dos eventos coletados sejam compostos por elétrons, com reduzida presença de falsos elétrons. Os resultados indicam que, dentre os métodos de classificação propostos, o RingRp encontrou melhor desempenho, apresentando uma redução de 0,13 ponto percentual na taxa de falso alarme quando comparado ao falso alarme obtido pelo NeuralRinger. O método RingRp-CMAES também apresentou melhoria na redução do falso alarme, alcançando uma redução de 0,08 ponto percentual em relação a taxa de falso alarme apresentada pelo NeuralRinger.

Tabela 5.2: Comparação de desempenho entre os classificadores avaliados para o ponto de operação *VeryLoose*. Os classificadores foram treinados na faixa $E_{T2} \times \eta_0$.

Discriminador	SP [%]	P _D [%]	P _F [%]	Neurônios Ocultos
NeuralRinger (Ref.)	98,99 ± 0,10	99,83 ± 0,1	1,85 ± 0,10	5
CutBased	81,87	99,83	34,30	-
Ring-GA	98,91 ± 0,11	99,83 ± 0,11	1,99 ± 0,11	8
Ring-PSO	98,95 ± 0,08	99,83 ± 0,09	1,91 ± 0,09	7
Ring-CMAES	98,99 ± 0,05	99,83 ± 0,05	1,84 ± 0,04	5
RingRp	99,04 ± 0,08	99,83 ± 0,08	1,72 ± 0,08	7
RingRp-GA	98,96 ± 0,09	99,83 ± 0,07	1,91 ± 0,07	6
RingRp-PSO	98,99 ± 0,09	99,83 ± 0,07	1,84 ± 0,07	7
RingRp-CMAES	99,02 ± 0,04	99,83 ± 0,04	1,77 ± 0,04	7

Pode-se observar que os métodos de treinamento RingRp e RingRp-CMAES, ambos com a etapa de pré-processamento, superaram o desempenho do NeuralRinger, nos pontos de operação avaliados para a subdivisão do espaço de fase $E_{T2} \times \eta_0$. Os métodos RingRp-GA e RingRp-PSO não apresentaram melhorias significativas (resultados com média e desvio padrão próximos ao NeuralRinger). Os métodos propostos realizam a otimização simultânea dos coeficientes α e β , da etapa de

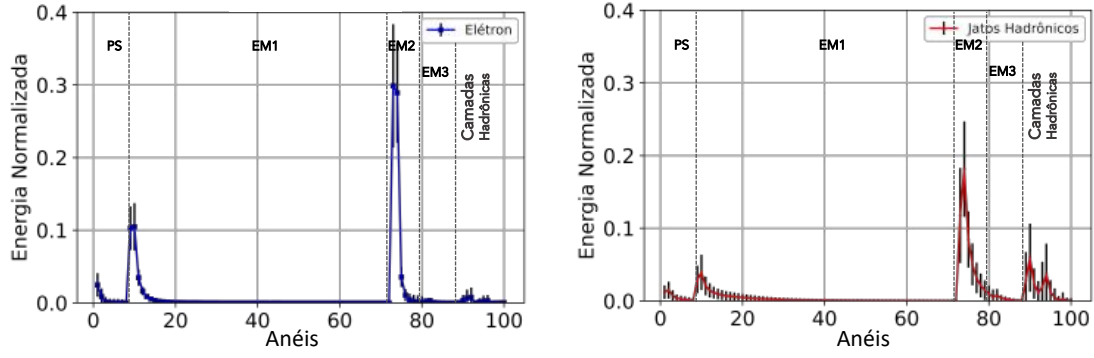
pré-processamento, que ponderam a energia do anel e a sua distância ao centro, respectivamente. Na Tabela 5.3 são apresentados os valores (média e desvio padrão) encontrados para os coeficientes, pelo distintos métodos de treinamento avaliados. Observa-se que os valores encontrados por cada método são próximos, indicando concordância na convergência destes parâmetros. Conforme foi observado nos pontos de operação *Tight* e *VeryLoose*, o método RpRing obteve melhor desempenho quando comparado ao NeuralRinger. Neste contexto, foi realizado uma análise do perfil dos sinais em anéis com o pré-processamento R_p .

Tabela 5.3: Valores (média e desvio padrão) dos coeficientes α e β encontrados pelos métodos de treinamento propostos. Os classificadores foram treinados na faixa $E_{T2} \times \eta_0$.

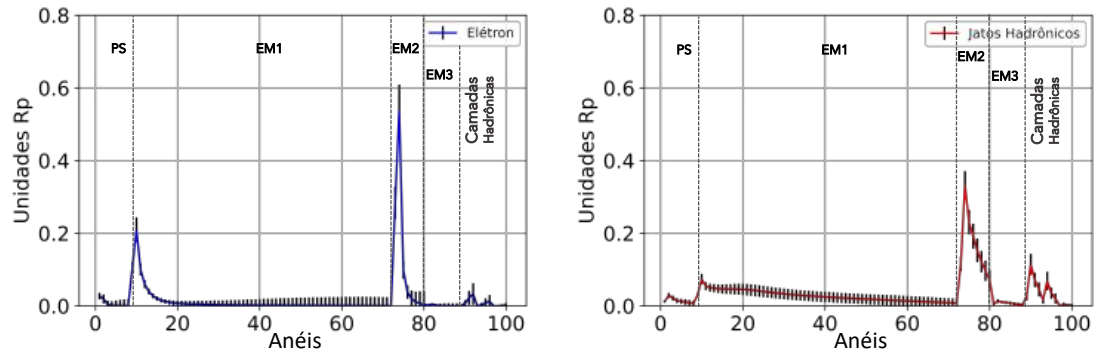
	α	β
RingRp-GA	$1,13 \pm 0,11$	$0,99 \pm 0,08$
RingRp-PSO	$1,09 \pm 0,03$	$0,98 \pm 0,10$
RingRp-CMAES	$1,07 \pm 0,04$	$0,95 \pm 0,07$
RingRp	$1,05 \pm 0,02$	$0,93 \pm 0,03$

Considerando que o método RingRp, dentre os avaliados, foi o que encontrou o melhor desempenho, uma análise do perfil médio dos sinais em anéis pré-processados foi realizada, utilizando os coeficientes α e β encontrados. O perfil médio dos sinais em anéis originais, do espaço de fase $E_{T2} \times \eta_0$, e os sinais em anéis com pré-processamento, podem ser observados nas Figuras 5.3 (a) e 5.3 (b) respectivamente. Nos sinais em anéis originais, é observado para o caso de elétrons (sinais em azul) deposição de energia nos primeiros anéis da camada EM1 (anéis 8 a 16), e uma maior de posição de energia nos primeiros anéis da camada EM2 (72 a 76). Para o caso dos jatos, há deposição de energia com menor amplitude nos primeiros anéis da camada EM1, e essa baixa amplitude é visualizada até aproximadamente o vigésimo anel, indicando maior espalhamento de energia, mesmo com baixa intensidade. No perfil de jatos também é observado uma acentuada deposição de energia na camada EM2 e uma maior intensidade de energia presente nas camadas hadrônicas (se comparado ao perfil de elétrons).

No perfil de elétrons com pré-processamento R_p (linha azul na Figura 5.3 (b)), não são visualizadas diferenças significativas em seu perfil médio, quando comparado ao perfil original dos anéis (ver linhas azuis na Figura 5.3 (a)). Para o caso do perfil médio de jatos (linha vermelha na Figura 5.3 (b)), é observado que o mapeamento R_p oferece um leve destaque para anéis com baixa energia presentes na camada EM1 (destaque observado principalmente do anel 8 ao anel 40, se comparado ao perfil de anéis originais para jatos). Este fato indica que o pré-processamento ajuda



(a) Perfil médio dos sinais em anéis



(b) Perfil médio dos sinais dos anéis com o mapeamento R_p

Figura 5.3: Perfil (média e desvio padrão) dos sinais em anéis do espaço de fase $E_{T2} \times \eta_0$ (a) e o perfil dos sinais em anéis processados pelo mapeamento R_p (b). A esquerda é apresentado o perfil de anéis para assinaturas de elétrons e a direita o perfil de anéis para assinaturas de jatos.

a destacar as energias medidas nos anéis mais distantes ao centro, fato que ocorre mais frequentemente nas assinaturas de jatos.

Os coeficientes que ponderam a energia e a distância entre os anéis foram $\alpha = 1,03$ e $\beta = 0,92$, encontrados pelo procedimento de otimização do método RingRp (*backpropagation* - ADAM), cuja variação ao longo do treinamento do classificador pode ser observada na Figura 5.4, sendo o losango vermelho o par de valores finais encontrados ao final do procedimento de otimização. Esta curva foi obtida pelo classificador que obteve maior índice SP entre os dez classificadores selecionados no procedimento de validação cruzada.

Os mesmos procedimentos que envolveram a definição do número de neurônios na camada oculta, e as comparações entre os métodos realizados anteriormente foram replicados para as demais subdivisões do espaço de fase. Neste contexto, a Tabela 5.4 apresenta os resultados dos classificadores avaliados para o ponto de operação

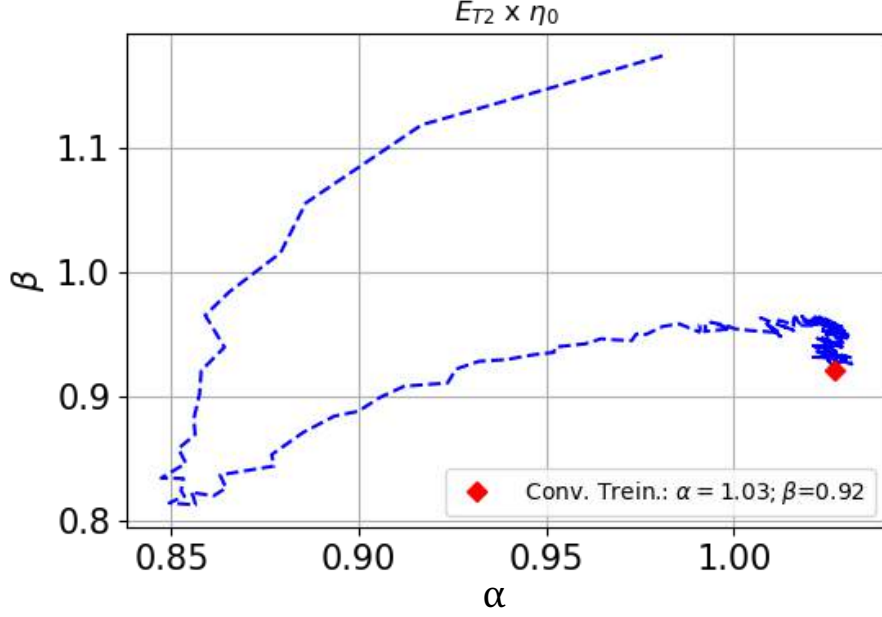


Figura 5.4: Variação dos valores de α e β ocorrida durante o procedimento de otimização do classificador neural treinado pelo método RingRp, que obteve melhor desempenho do índice SP. Espaço de fase $E_{T2} \times \eta_0$.

Tight, no espaço de fase $E_{T2} \times \eta_2$ (*crack*)². Esta região da coordenada pseudorapidez apresenta uma redução na quantidade de sensores, quando comparada região central do barril, pois trata-se de um espaço reservado para passagem de cabos, fibras ópticas e conexões utilizadas na transmissão de dados entre os dispositivos do detector. Logo, a eficiência de detecção para essa faixa em η é diminuída em relação ao barril central e as taxas de falso alarme são em geral mais altas.

Dentre os classificadores treinados neste espaço de fase, o método RingRp encontrou melhor desempenho, apresentando uma redução na taxa de falso alarme de 0,59 ponto percentual ($P_F[\%] = 1,81 \pm 0,14$), quando comparado ao NeuralRinger ($P_F[\%] = 2,40 \pm 0,22$). O método RingRp-CMAES destacou-se dentre os métodos que utilizam algoritmos bio-inspirados, obtendo desempenho superior ao NeuralRinger, onde alcançou uma redução de 0,51 ponto percentual na taxa de falso alarme ($P_F[\%] = 1,89 \pm 0,22$). Quando comparado ao RingRp, o RingRp-CMAES apresentou taxa de falso alarme superior em 0,08 ponto percentual. Os métodos RingRp-GA e RingRp-PSO apresentaram também resultados interessantes, reduzindo a taxa de falso alarme em 0,18 e 0,39 ponto percentual respectivamente em relação ao NeuralRinger.

Na Tabela 5.5 são exibidos os resultados dos classificadores avaliados para o ponto de operação *VeryLoose*, no espaço de fase $E_{T2} \times \eta_2$. Os resultados do método RingRp

²Intervalo que compreende eventos com energia transversa em $30 \leq E_T[GeV] < 40$ e posição de interação no calorímetro na coordenada pseudorapidez $1,37 \leq |\eta| < 1,54$.

Tabela 5.4: Comparação de desempenho entre os classificadores avaliados para o ponto de operação *Tight*. Os classificadores foram treinados na faixa $E_{T2} \times \eta_2$.

Discriminador	SP [%]	P _D [%]	P _F [%]	Neurônios Ocultos
NeuralRinger (Ref.)	97, 22 ± 0, 06	96, 84 ± 0, 05	2, 40 ± 0, 22	5
CutBased	88, 53	96, 81	19, 37	-
Ring-GA	97, 03 ± 0, 08	99, 84 ± 0, 11	2, 77 ± 0, 08	8
Ring-PSO	97, 17 ± 0, 07	99, 84 ± 0, 07	2, 49 ± 0, 07	8
Ring-CMAES	97, 25 ± 0, 03	99, 84 ± 0, 03	2, 30 ± 0, 03	8
RingRp	97, 50 ± 0, 08	96, 81 ± 0, 08	1, 81 ± 0, 14	6
RingRp-GA	97, 06 ± 0, 16	99, 84 ± 0, 16	2, 22 ± 0, 16	7
RingRp-PSO	97, 23 ± 0, 11	99, 84 ± 0, 11	2, 01 ± 0, 11	8
RingRp-CMAES	99, 14 ± 0, 08	99, 84 ± 0, 08	1, 89 ± 0, 08	10

(que obteve melhor desempenho dentre os avaliados) indicam uma redução do falso alarme de 0,85 ponto percentual em relação ao NeuralRinger. O RingRp-CMAES foi o método que, dentre os algoritmos bio-inspirados, obteve desempenho levemente inferior ao RingRp, com aumento do falso alarme de 0,04 ponto percentual comparado ao RingRp, e redução na taxa de falso alarme de 0,81 ponto percentual em relação ao NeuralRinger. Nota-se também, que os métodos Ring-PSO e Ring-CMAES (sem pré-processamento R_p) encontram resultados com taxa de falso alarme inferior ao NeuralRinger, reduzindo em 0,13 e 0,18 ponto percentual respectivamente.

Tabela 5.5: Comparação de desempenho entre os classificadores avaliados para o ponto de operação *VeryLoose*. Os classificadores foram treinados na faixa $E_{T2} \times \eta_2$.

Discriminador	SP [%]	P _D [%]	P _F [%]	Neurônios Ocultos
NeuralRinger (Ref.)	97, 40 ± 0, 08	97, 97 ± 0, 08	3, 16 ± 0, 08	5
CutBased	84, 41	97, 96	28, 13	-
Ring-GA	98, 91 ± 0, 18	97, 97 ± 0, 18	3, 24 ± 0, 18	8
Ring-PSO	98, 95 ± 0, 12	97, 97 ± 0, 12	3, 03 ± 0, 12	7
Ring-CMAES	98, 99 ± 0, 11	97, 97 ± 0, 08	2, 98 ± 0, 08	5
RingRp	97, 83 ± 0, 12	97, 97 ± 0, 12	2, 31 ± 0, 12	5
RingRp-GA	98, 96 ± 0, 09	97, 97 ± 0, 17	2, 94 ± 0, 17	6
RingRp-PSO	98, 99 ± 0, 09	97, 97 ± 0, 11	2, 47 ± 0, 11	10
RingRp-CMAES	99, 02 ± 0, 10	97, 97 ± 0, 10	2, 35 ± 0, 10	10

Na Figura [5.5](#) é apresentado o perfil médio dos sinais em anéis originais (a) e pré-processados pelo mapeamento R_p (b), para o espaço de fase $E_{T2} \times \eta_2$. Neste espaço de fase, o perfil médio dos sinais originais dos anéis apresentam deposição de energia nas camadas hadrônicas (anéis 92 a 10), o que pode implicar no aumento da taxa de falsos elétrons (perfis de deposição de energia parecidos). O perfil médio dos

anéis originais de elétrons apresenta deposição principalmente nos primeiros anéis da camada PS (0 ao 4) e nos primeiros anéis das camadas EM1 (8 a 16) e EM2 (72 a 76). O perfil médio de jatos nos anéis originais indica deposição de energia com baixa amplitude do anéis 1 (camada PS) até o anel 30 (camada EM1), indicando um maior espalhamento lateral (com menor energia) em torno do anel central de sua respectiva camada.

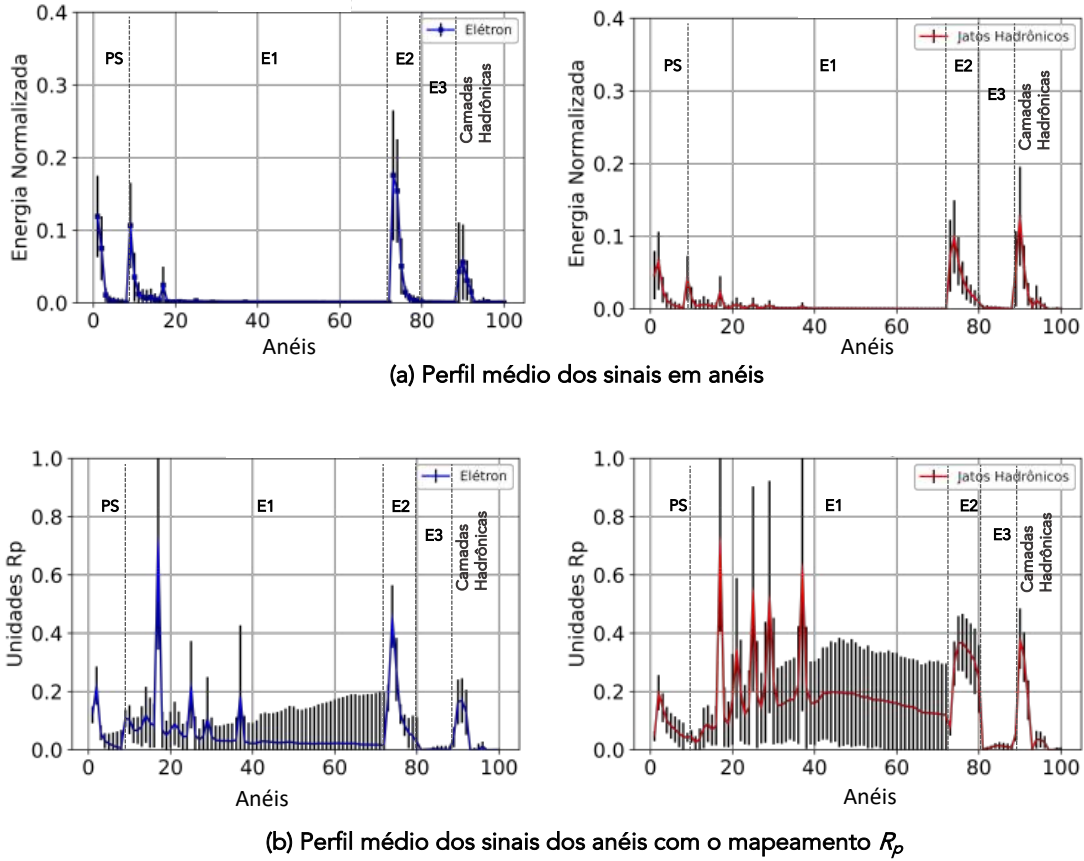


Figura 5.5: Perfil (média e desvio padrão) dos sinais em anéis do espaço de fase $E_{T2} \times \eta_2$ (a) e o perfil dos sinais em anéis processados pelo mapeamento R_p (b). A esquerda é apresentado o perfil de anéis para assinaturas de elétrons e a direita o perfil de anéis para assinaturas de jatos.

No caso do perfil de elétrons processados pelo mapeamento R_p (Figura 5.5 (b)), houve elevado destaque nos anéis próximos ao centro cuja deposição de energia foi acentuada. Anéis mais externos (20 ao 40) foram também destacados. As assinaturas de jatos pré-processadas, tiveram forte destaque nos anéis mais externos, revelando com maior amplitude a energia contida nos anéis que se afastam do anel central de sua respectiva camada. Este comportamento é esperado, baseado nos aspectos teóricos que indicam que jatos apresentam perfil lateral mais largo em relação a elétrons [21], e o mapeamento R_p ajuda a destacar esta característica discri-

minante. Os coeficientes utilizados na etapa de pré-processamento foram $\alpha = 1,02$ e $\beta = 1,56$, cuja as variações ao longo do procedimento de otimização podem ser observadas na Figura 5.6.

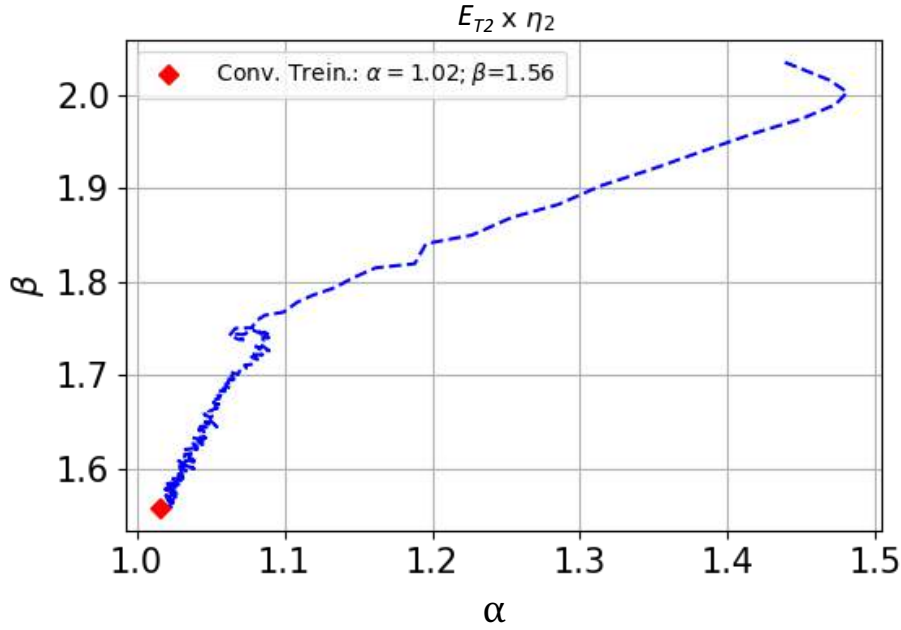


Figura 5.6: Variação dos valores de α e β ocorrida durante o procedimento de otimização do classificador neural, que obteve melhor desempenho do índice SP. Espaço de fase $E_{T2} \times \eta_2$.

Os resultados encontrados dos classificadores projetados para o espaço de fase $E_{T2} \times \eta_4$ podem ser observados na Tabela 5.6, para o ponto de operação *Tight*. Este espaço de fase é principalmente caracterizado pela mudança de granularidade das células na coordenada $|\eta|$, quando comparado ao barril central [7]. As células da primeira camada do calorímetro eletromagnético no intervalo $|\eta_4|$, apresentam dimensões maiores ($\Delta\eta \times \Delta\phi = 0,1 \times 0,1$), em relação ao barril central ($\Delta\eta \times \Delta\phi = 0,003 \times 0,1$), conforme indicado na Tabela 2.2, o que influencia no perfil médio dos sinais em anéis gerados nessa região. Outro aspecto desta região é a ausência da camada PS, que é limitada em $|\eta| < 1,8$, assim o perfil de deposição de energia não conta com os primeiros 8 anéis. Tais limitações, em geral, reduzem a eficiência de detecção de elétrons nessa região, quando comparado ao desempenho obtido na região do barril central. Os resultados encontrados neste espaço de fase indicam que o método RingRp encontrou uma redução de 0,11 ponto percentual na taxa de falso alarme quando comparado com o resultado obtido pelo NeuralRinger. O método RingRp-CMAES foi o segundo método com melhor desempenho, no qual obteve redução de 0,06 ponto percentual da taxa de falso alarme em relação ao desempenho do NeuralRinger.

Na Tabela 5.7 são apresentados resultados dos classificadores avaliados no espaço

Tabela 5.6: Comparação de desempenho entre os classificadores avaliados para o ponto de operação *Tight*. Os classificadores foram treinados na faixa $E_{T2} \times \eta_4$.

Discriminador	SP [%]	P _D [%]	P _F [%]	Neurônios Ocultos
NeuralRinger (Ref.)	96,01 ± 0,06	95,70 ± 0,06	3,67 ± 0,25	5
CutBased	74,45	95,68	44,10	-
Ring-GA	95,90 ± 0,08	95,70 ± 0,09	3,89 ± 0,17	6
Ring-PSO	96,00 ± 0,07	95,70 ± 0,07	3,69 ± 0,18	6
Ring-CMAES	96,03 ± 0,03	95,70 ± 0,03	3,63 ± 0,12	7
RingRp	96,07 ± 0,12	95,71 ± 0,12	3,56 ± 0,12	10
RingRp-GA	96,01 ± 0,18	95,70 ± 0,18	3,66 ± 0,18	7
RingRp-PSO	96,02 ± 0,17	95,70 ± 0,17	3,65 ± 0,18	10
RingRp-CMAES	96,04 ± 0,13	95,70 ± 0,13	3,61 ± 0,13	10

de fase $E_{T2} \times \eta_4$, para o critério de operação *VeryLoose*. Neste ponto de operação o método RingRp encontrou melhor desempenho, com redução da taxa de falso alarme em comparação com a referência. O RingRp apresentou redução de 0,1 ponto percentual na taxa de falso alarme ($P_F[\%] = 2,25 \pm 0,07$), se comparado aos resultados obtidos pelo NeuralRinger ($P_F[\%] = 2,35 \pm 0,18$). Os resultados indicam que o método RingRp-CMAES também supera o método de referência, onde obteve redução na taxa de falso alarme de 0,04 ponto percentual ($P_F[\%] = 2,31 \pm 0,07$), comparado ao NeuralRinger. Observa-se que os métodos RingRp e RingRp-CMAES encontram menor dispersão (desvio padrão) em seus resultados, e relação aos demais métodos avaliados, indicando uma melhor convergência aos valores encontrados.

Tabela 5.7: Comparação de desempenho entre os classificadores avaliados para o ponto de operação *VeryLoose*. Os classificadores foram treinados na faixa $E_{T2} \times \eta_4$.

Discriminador	SP [%]	P _D [%]	P _F [%]	Neurônios Ocultos
NeuralRinger (Ref.)	97,17 ± 0,14	96,70 ± 0,16	2,35 ± 0,18	5
CutBased	73,78	96,71	46,03	-
Ring-GA	97,09 ± 0,18	96,71 ± 0,18	2,52 ± 0,11	4
Ring-PSO	97,14 ± 0,15	96,70 ± 0,15	2,41 ± 0,13	7
Ring-CMAES	97,19 ± 0,11	96,71 ± 0,11	2,35 ± 0,07	5
RingRp	97,22 ± 0,07	96,70 ± 0,07	2,25 ± 0,07	9
RingRp-GA	97,11 ± 0,09	96,71 ± 0,17	2,48 ± 0,17	7
RingRp-PSO	97,16 ± 0,09	96,70 ± 0,11	2,37 ± 0,11	8
RingRp-CMAES	97,20 ± 0,07	96,70 ± 0,07	2,31 ± 0,07	8

Pode-se observar na Figura 5.7 o perfil médio dos sinais em anéis originais (a) e o perfil médio dos sinais em anéis com pré-processamento R_p (b). O mapeamento R_p foi realizado com valores os $\alpha = 1,06$ e $\beta = 1,68$ (ver Figura 5.8). É observado

que o mapeamento R_p oferece destaque a energia dos anéis mais externos medidos na camada EM1. Para o perfil médio de elétrons pré-processados esse destaque é visualizado entre os anéis 12 até o anel 40, enquanto no perfil médio de jatos, todos os anéis da camada EM1 são destacados, quando comparados aos anéis originais.

O perfil encontrado com os sinais pré-processados indica o efeito do mapeamento R_p na base de informação anelada, permitindo destacar o espalhamento lateral da energia, sobretudo para as assinaturas de jatos, nos quais tal característica ocorre com mais frequência.

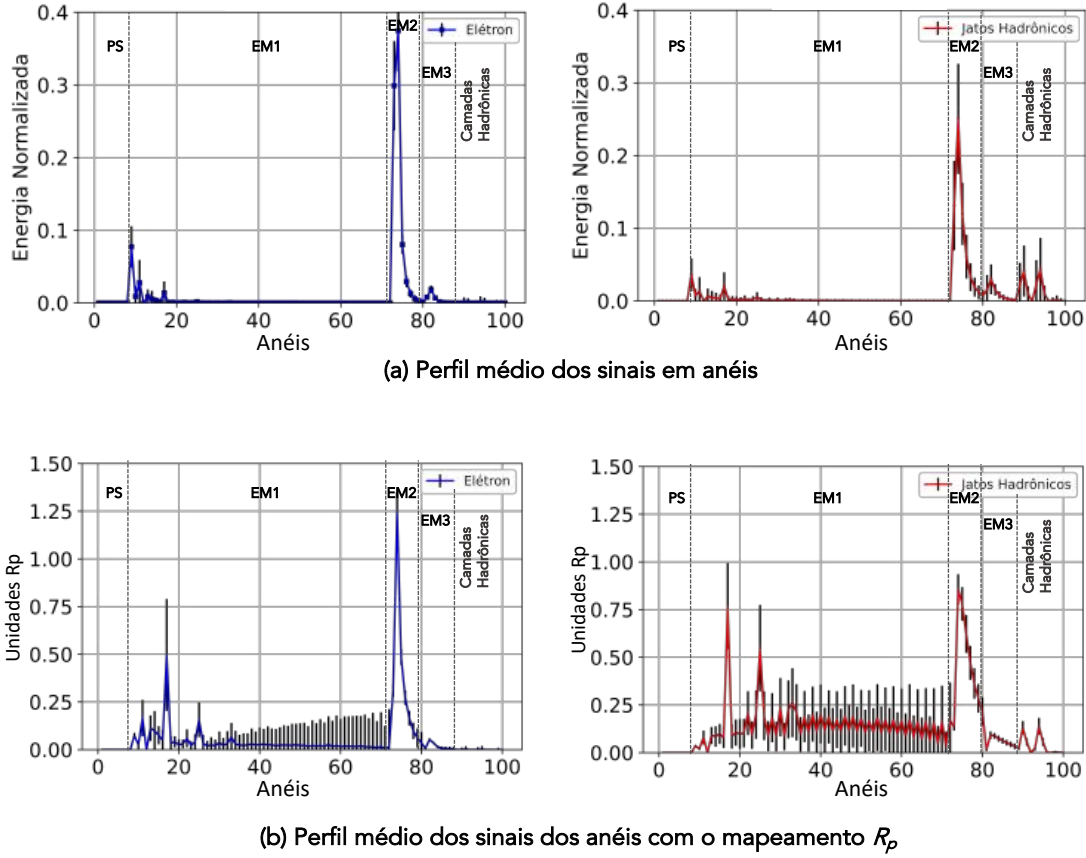


Figura 5.7: Perfil (média e desvio padrão) dos sinais em anéis do espaço de fase $E_{T2} \times \eta_4$ (a) e o perfil dos sinais em anéis processados pelo mapeamento R_p (b). A esquerda é apresentado o perfil de anéis para assinaturas de elétrons e a direita o perfil de anéis para assinaturas de jatos.

Com base nos resultados obtidos pelos classificadores no intervalos de $E_T \times \eta$ avaliados, foram selecionados os métodos que obtiveram melhor desempenho, para avaliar sua atuação em todo o espaço de fase disponível. Na Tabela 5.8 são apresentados os resultados obtidos em todo o espaço de fase para os métodos de treinamento RingRp, RingRp-CMAES e o NeuralRinger, com operação do critério *Tight*.

Observa-se que nos resultados obtidos no espaço de fase $E_{T0} \times \eta_0$, ambos os

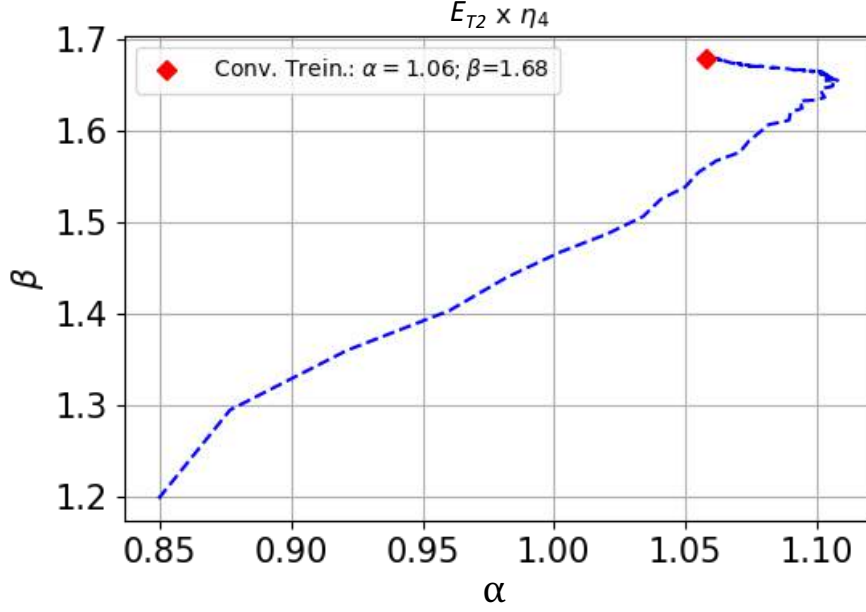


Figura 5.8: Variação dos valores de α e β ocorrida durante o procedimento de otimização do classificador neural, que obteve melhor desempenho do índice SP. Espaço de fase $E_{T2} \times \eta_4$.

métodos RingRp e RingRp-CMAES superam os resultados encontrados pelo NeuralRinger ($P_F[\%] = 2,59\% \pm 0,08$), com uma redução na taxa de falso alarme de 0,32 ($P_F[\%] = 2,27 \pm 0,09$) e 0,08 ($P_F[\%] = 2,51 \pm 0,07$) respectivamente. Para os demais intervalos de $|\eta|$ na faixa de energia $15 \leq E_T[GeV] < 20$, o método RingRp supera os em termos de redução na taxa de falso alarme quando comparado com os demais métodos no mesmo ponto de operação de P_D . Pode-se destacar os resultados encontrados pelo RingRp nos espaços de fase $E_{T0} \times \eta_2$ (região do *crack*) e $E_{T0} \times \eta_3$ com uma redução na taxa de falso alarme de 1,64 ($P_F[\%] = 7,93 \pm 0,52$) e 1,06 ($P_F[\%] = 3,54 \pm 0,52$) respectivamente quando comparado ao NeuralRinger.

Na faixa de energia $20 \leq E_T[GeV] < 30$, são também observadas melhorias em relação a taxa de falso alarme, pelo método RingRp, se comparado aos demais métodos avaliados. Comparando com os resultados obtidos pelo NeuralRinger no intervalo $0,00 \leq |n| < 0,08$, o método RingRp encontra uma taxa de $P_F = 1,45 \pm 0,09$, que representa a diminuição de 0,2 ponto percentual em relação a taxa de falso alarme do NeuralRinger ($P_F = 1,65 \pm 0,09$). Neste espaço de fase, o método RingRp-CMAES obtém desempenho semelhante ao NeuralRinger, e não apresentou melhoria significativa. Destacam-se os resultados obtidos pelo RingRp na região do *crack*, com redução na taxa de falso alarme de 0,91 ponto percentual ($P_F = 2,63 \pm 0,30$), quando comparado a taxa obtida pelo NeuralRinger ($P_F = 3,54 \pm 0,30$).

Considerando os classificadores treinados para a faixa de energia $40 \leq E_T[GeV] < 50$, observam-se melhorias de desempenho obtidas pelo método RingRp, em todos

os intervalos em η avaliados. Destaca-se a redução do falso alarme na região do *crack* em 0,54 ponto percentual ($P_F = 1,48 \pm 0,29$), se comparado a taxa de falso alarme encontrada pelo NeuralRinger ($P_F = 2,02 \pm 0,25$).

Para a faixa de energia $E_T > 50$ GeV, verifica-se também contribuições do método RingRp na rejeição de falsos elétrons. Os resultados mais destacados ocorreram nos intervalos de η_2 e η_3 onde foram obtidas taxas de $P_F = 0,59 \pm 0,03$ e $P_F = 0,42 \pm 0,04$ respectivamente, enquanto o NeuralRinger obteve nessas faixas resultados $P_F = 1,12 \pm 0,12$ e $P_F = 1,22 \pm 0,21$. Embora o método RingRp-CMAES não tenha superado o RingRp, seu desempenho melhora a rejeição de falsos alarme em relação ao Neural Ringer, com uma redução de também superou o desempenho do NeuralRinger, obtendo uma redução de 0,3 ponto percentual ($P_F = 0,79 \pm 0,04$) na faixa η_2 e uma redução de 0,15 ponto percentual ($P_F = 1,07 \pm 0,04$) na faixa η_3 comparado ao NeuralRinger nessas faixas.

		E_{T0}			E_{T1}			E_{T2}			E_{T3}			E_{T4}		
	Método	P_D [%]	SP [%]	P_F [%]	P_D [%]	SP [%]	P_F [%]	P_D [%]	SP [%]	P_F [%]	P_D [%]	SP [%]	P_F [%]	P_D [%]	SP [%]	P_F [%]
η_0	NeuralRinger	97,78±0,07	97,63±0,07	2,59±0,08	98,68±0,10	98,52±0,10	1,65±0,11	99,18±0,05	99,11±0,06	0,96±0,04	99,57±0,08	99,40±0,1	0,77±0,03	99,63±0,02	99,67±0,02	0,29±0,03
	RingRp-CMAES	97,78±0,07	97,63±0,07	2,51±0,07	98,68±0,10	98,53±0,10	1,63±0,10	99,18±0,06	99,12±0,06	0,93±0,06	99,57±0,09	99,39±0,09	0,80±0,09	99,63±0,02	99,64±0,02	0,35±0,02
	RingRp	97,77±0,09	97,75±0,09	2,27±0,09	98,68±0,09	98,62±0,09	1,45±0,09	99,18±0,04	99,15±0,04	0,87±0,04	99,56±0,08	99,41±0,08	0,74±0,08	99,63±0,01	99,69±0,01	0,26±0,01
η_1	NeuralRinger	97,78±0,01	96,70±0,10	4,37±0,20	98,88±0,00	97,86±0,03	3,16±0,07	94,43±0,00	98,67±0,04	2,09±0,07	99,68±0,00	99,02±0,07	1,63±0,15	99,67±0,00	99,27±0,06	1,12±0,12
	RingRp-CMAES	97,78±0,16	96,83±0,16	4,11±0,16	98,88±0,08	97,87±0,08	3,14±0,08	99,43±0,11	98,69±0,11	2,04±0,11	99,67±0,14	99,03±0,14	1,62±0,14	99,67±0,04	99,44±0,04	0,79±0,04
	RingRp	97,78±0,22	96,97±0,22	3,85±0,22	98,89±0,09	98,01±0,09	2,86±0,09	99,43±0,08	98,76±0,08	1,91±0,08	99,67±0,12	99,07±0,12	1,53±0,12	99,67±0,03	99,54±0,03	0,59±0,03
η_2	NeuralRinger	95,78±0,02	93,09±0,29	9,57±0,58	94,51±0,04	95,49±0,15	3,54±0,30	96,84±0,05	97,22±0,12	2,40±0,22	97,27±0,07	97,62±0,14	2,02±0,25	98,28±0,00	98,53±0,11	1,22±0,21
	RingRp-CMAES	95,81±0,93	93,58±0,93	8,62±0,93	94,52±0,30	95,69±0,30	3,14±0,30	96,84±0,23	97,33±0,23	2,15±0,23	97,26±0,23	97,68±0,23	1,90±0,23	98,28±0,04	98,85±0,04	1,07±0,04
	RingRp	95,80±0,52	93,93±0,52	7,93±0,52	94,53±0,30	95,95±0,30	2,63±0,30	96,84±0,24	97,50±0,24	1,81±0,24	97,27±0,29	97,89±0,29	1,48±0,29	98,28±0,04	98,93±0,04	0,42±0,04
η_3	NeuralRinger	97,40±0,01	96,40±0,09	4,60±0,18	98,23±0,00	97,54±0,14	3,15±0,29	98,98±0,00	98,29±0,12	2,40±0,24	99,22±0,00	98,66±0,10	1,89±0,21	99,32±0,00	98,75±0,13	1,80±0,25
	RingRp-CMAES	97,40±0,15	96,65±0,15	4,09±0,15	98,23±0,11	97,58±0,11	3,06±0,11	98,97±0,10	98,39±0,10	2,20±0,10	99,22±0,12	98,76±0,12	1,69±0,12	99,32±0,04	98,87±0,04	1,57±0,04
	RingRp	97,40±0,11	96,93±0,11	3,54±0,11	98,23±0,10	97,78±0,10	2,67±0,10	98,97±0,13	98,46±0,13	2,04±0,13	99,22±0,09	98,82±0,09	1,58±0,09	99,32±0,02	98,95±0,02	1,38±0,02
η_4	NeuralRinger	92,98±0,13	94,06±0,20	4,86±0,45	95,70±0,06	96,01±0,13	3,67±0,25	96,69±0,14	97,17±0,16	2,35±0,01	97,12±0,29	97,77±0,24	1,55±0,29	96,00±0,06	97,60±0,06	0,79±0,10
	RingRp-CMAES	93,05±0,67	94,25±0,67	4,75±0,67	95,67±0,31	96,04±0,31	3,59±0,31	96,71±0,25	97,18±0,25	2,34±0,25	97,15±0,30	97,76±0,30	1,55±0,30	96,00±0,04	97,70±0,04	0,57±0,04
	RingRp	93,03±0,81	94,22±0,81	4,58±0,81	95,54±0,44	96,04±0,44	3,47±0,44	96,69±0,22	97,19±0,22	2,31±0,22	97,09±0,27	97,80±0,27	1,48±0,27	96,00±0,03	97,78±0,03	0,41±0,03

Tabela 5.8: Resultados obtidos para o treinamento dos classificadores neurais em todo o espaço de fase. São indicados os valores para o ponto de operação *Tight* encontrados para as taxas de P_D , P_F e SP (média e desvio padrão), dos classificadores neurais com melhor desempenho em cada sorteio da validação cruzada.

A análise em todos os espaços de fase também foi realizada para o ponto de operação *VeryLoose*. A Tabela 5.9 apresenta os resultados do treinamento dos classificadores, com os diferentes métodos avaliados, com limiar de decisão ajustados para este ponto de operação. Em geral, os métodos propostos RingRp e RingRp-CMAES indicam melhoria de desempenho na rejeição de falsos elétrons. No ponto de operação *VeryLoose*, os classificadores operam com uma maior eficiência de detecção de elétrons, entretanto em contrapartida, as taxas de falso alarme são geralmente maiores. Neste contexto, os métodos propostos ganham destaque por oferecer menores taxas de falso alarme, característica que é importante para o sistema de filtragem *online*, pois permite a redução da ocupação das fazendas computacionais do experimento, que estariam processando eventos não relevantes.

Observa-se que, nos resultados obtidos no espaço de fase $E_{T0} \times \eta_0$, ambos os métodos RingRp e RingRp-CMAES encontram resultados que reduzem a taxa de falso alarme se comparados ao NeuralRinger. O método RingRp-CMAES obteve uma taxa de $P_F[\%] = 3,64\% \pm 0,20$, enquanto o método RingRp obteve o valor de $P_F[\%] = 3,32\% \pm 0,15$, que configura uma redução de 0,09 e 0,41 ponto percentual respectivamente quando comparados ao NeuralRinger ($P_F[\%] = 3,73\% \pm 0,11$).

Destacam-se os resultados obtidos pelo método RingRp nos intervalos η_2 e η_3 , que indicam uma redução de 0,48 ($P_F[\%] = 5,38\% \pm 0,23$) e 2,83 ($P_F[\%] = 15,00\% \pm 1,20$) ponto percentual respectivamente, quando comparados aos resultados encontrados pelo NeuralRinger com $P_F[\%] = 5,84\% \pm 0,37$ para a faixa η_2 e $P_F[\%] = 17,83\% \pm 2,52$ na faixa η_3 (região do *crack*).

Para as faixas de energia mais altas, os classificadores treinados pelo método RingRp encontram uma menor taxa de falso alarme em relação ao NeuralRinger, com destaque no espaço de fase $E_{T3} \times \eta_0$ (região do barril central do ATLAS), onde obteve uma redução de 0,32 ponto percentual ($P_F[\%] = 1,66\% \pm 0,11$) comparado ao NeuralRinger ($P_F[\%] = 1,98\% \pm 0,13$).

		E_{T0}			E_{T1}			E_{T2}			E_{T3}			E_{T4}		
	Método	P_D [%]	SP [%]	P_F [%]	P_D [%]	SP [%]	P_F [%]	P_D [%]	SP [%]	P_F [%]	P_D [%]	SP [%]	P_F [%]	P_D [%]	SP [%]	P_F [%]
η_0	NeuralRinger	98,86±0,00	97,55±0,06	3,73±0,11	99,50±0,00	98,41±0,08	2,67±0,16	99,83±0,00	98,94±0,07	1,95±0,13	99,94±0,00	98,98±0,06	1,98±0,13	99,98±0,00	99,22±0,20	1,54±0,39
	RingRp-CMAES	98,86±0,20	97,61±0,20	3,64±0,20	99,50±0,13	98,46±0,13	2,58±0,13	99,83±0,10	98,99±0,10	1,85±0,10	99,94±0,11	99,11±0,11	1,72±0,11	99,98±2,30	98,75±2,30	1,47±2,30
	RingRp	98,86±0,15	97,77±0,15	3,32±0,15	99,50±0,12	98,58±0,12	2,33±0,12	99,83±0,08	99,04±0,08	1,74±0,08	99,94±0,11	99,14±0,11	1,66±0,11	99,98±0,07	99,52±0,07	0,93±0,07
η_1	NeuralRinger	98,78±0,00	96,46±0,19	5,84±0,37	99,62±0,00	97,34±0,07	4,92±0,14	99,91±0,00	97,90±0,09	4,08±0,19	99,97±0,00	97,89±0,16	4,17±0,32	99,93±0,00	99,00±0,10	1,92±0,19
	RingRp-CMAES	98,79±0,17	96,54±0,17	5,68±0,17	99,63±0,15	97,37±0,15	4,85±0,15	99,91±0,22	97,98±0,22	3,92±0,22	99,97±0,45	98,05±0,45	3,85±0,45	99,93±0,11	99,09±0,11	1,74±0,11
	RingRp	98,79±0,23	96,69±0,23	5,38±0,23	99,62±0,17	97,57±0,17	4,45±0,17	99,91±0,18	98,09±0,18	3,70±0,18	99,97±0,30	98,17±0,30	3,62±0,30	99,93±0,08	99,40±0,08	1,13±0,08
η_2	NeuralRinger	97,60±0,00	89,71±1,32	17,83±2,52	96,71±0,03	95,97±0,18	4,77±0,34	97,97±0,02	97,40±0,14	3,16±0,28	97,87±0,04	97,80±0,11	2,28±0,21	98,60±0,00	98,59±0,12	1,42±0,23
	RingRp-CMAES	97,62±1,60	90,57±1,60	16,20±1,60	96,70±0,45	96,29±0,45	4,11±0,45	97,96±0,30	97,63±0,30	2,71±0,30	97,89±0,23	97,89±0,23	2,12±0,23	98,60±0,04	98,96±0,04	0,68±0,04
	RingRp	97,61±1,20	91,19±1,20	15,00±1,20	96,70±0,34	96,57±0,34	3,55±0,34	97,99±0,21	97,82±0,21	2,35±0,21	97,88±0,34	98,08±0,34	1,73±0,34	98,60±0,04	99,07±0,04	0,47±0,04
η_3	NeuralRinger	98,25±0,00	96,41±0,12	5,42±0,24	99,57±0,00	97,24±0,19	5,06±0,38	99,91±0,00	97,53±0,23	4,83±0,45	99,97±0,00	97,73±0,20	4,49±0,32	99,98±0,00	96,64±0,64	6,65±1,25
	RingRp-CMAES	98,25±0,17	96,74±0,17	4,75±0,17	99,57±0,19	97,29±0,19	4,97±0,19	99,92±0,16	97,72±0,16	4,46±0,16	99,97±0,20	98,01±0,20	3,94±0,20	99,98±0,04	96,97±0,65	5,99±0,65
	RingRp	98,25±0,14	97,05±0,14	4,14±0,14	99,57±0,20	97,58±0,20	4,40±0,20	99,92±0,12	97,87±0,12	4,16±0,12	99,97±0,21	98,08±0,21	3,79±0,21	99,98±0,05	99,46±0,36	5,01±0,34
η_4	NeuralRinger	93,57±0,10	94,26±0,28	5,05±0,52	95,73±0,04	96,04±0,14	3,69±0,26	96,69±0,14	97,17±0,16	2,35±0,31	97,12±0,29	97,78±0,24	1,55±0,29	96,01±0,04	97,60±0,06	0,79±0,10
	RingRp-CMAES	93,63±0,89	94,36±0,89	4,90±0,89	95,72±0,31	96,05±0,31	3,62±0,06	97,60±0,06	97,20±0,25	2,34±0,25	97,15±0,30	97,80±0,30	1,55±0,30	96,01±0,06	97,65±0,04	0,68±0,04
	RingRp	93,58±0,91	94,34±0,91	4,89±0,91	95,72±0,35	96,13±0,35	3,58±0,35	96,72±0,22	97,20±0,22	2,32±0,22	97,09±0,27	97,76±0,27	1,38±0,27	96,01±0,04	97,76±0,04	0,47±0,04

Tabela 5.9: Resultados obtidos para o treinamento dos classificadores neurais em todo o espaço de fase. São indicados os valores para o ponto de operação *VeryLoose* encontrados para as taxas de P_D , P_F e SP (média e desvio padrão), dos classificadores neurais com melhor desempenho em cada sorteio da validação cruzada.

Dentre os métodos de treinamento avaliados, o RingRp obteve, em geral, desempenho superior em todos os espaços de fase. Na Tabela 5.10 são apresentados os valores (média e desvio padrão) dos coeficientes α e β encontrados pelo método RingRp pelas inicializações de treinamento que obtiveram melhor índice SP em cada sorteio, e em cada intervalo $E_T \times \eta$. Os coeficientes extraídos foram obtidos ao fim do procedimento de otimização (valor dos coeficientes quando houve a parada do treinamento).

Tabela 5.10: Média e desvio padrão dos coeficientes α e β encontrados pelo método RingRp dos classificadores neurais com melhor desempenho de índice SP nas inicializações de treinamento.

	Coeficientes	E_{T0}	E_{T1}	E_{T2}	E_{T3}	E_{T4}
η_0	α	$1,23 \pm 0,04$	$1,17 \pm 0,01$	$1,05 \pm 0,02$	$0,96 \pm 0,02$	$0,92 \pm 0,01$
	β	$1,20 \pm 0,04$	$1,03 \pm 0,04$	$0,93 \pm 0,03$	$0,88 \pm 0,07$	$0,85 \pm 0,03$
η_1	α	$1,17 \pm 0,01$	$1,14 \pm 0,02$	$1,02 \pm 0,04$	$0,91 \pm 0,05$	$0,89 \pm 0,03$
	β	$1,24 \pm 0,03$	$1,16 \pm 0,04$	$1,08 \pm 0,08$	$0,98 \pm 0,07$	$0,95 \pm 0,04$
η_2	α	$1,09 \pm 0,08$	$1,05 \pm 0,06$	$1,04 \pm 0,06$	$0,96 \pm 0,03$	$0,95 \pm 0,04$
	β	$1,55 \pm 0,11$	$1,44 \pm 0,28$	$1,42 \pm 0,18$	$1,41 \pm 0,15$	$1,38 \pm 0,14$
η_3	α	$1,28 \pm 0,04$	$1,20 \pm 0,02$	$1,11 \pm 0,01$	$1,01 \pm 0,02$	$0,98 \pm 0,02$
	β	$1,41 \pm 0,05$	$1,25 \pm 0,03$	$1,20 \pm 0,03$	$1,13 \pm 0,04$	$1,08 \pm 0,05$
η_4	α	$1,18 \pm 0,05$	$1,16 \pm 0,04$	$1,14 \pm 0,04$	$1,05 \pm 0,04$	$1,03 \pm 0,05$
	β	$1,47 \pm 0,28$	$1,44 \pm 0,08$	$1,41 \pm 0,04$	$1,40 \pm 0,28$	$1,39 \pm 0,13$

Nota-se que a média dos valores de α e β encontrados reduzem, a medida que a energia dos anéis aumentam. Uma possível interpretação para este resultado deve-se ao fato que em faixas mais altas de energia, as assinaturas de elétrons e jatos apresentam perfis laterais e longitudinais mais discriminativos, assim pré-processamento especialista tende a diminuir a influência dos parâmetros no mapeamento.

Os resultados indicam que, os valores médios para β aumentam comparando-se os intervalos η_0 e η_4 . O aumento do coeficiente β observado, entre estes intervalos na coordenada η pode estar relacionado com a redução da granularidade das células em η_4 . Neste caso, o pré-processamento especialista tenta destacar nos sinais em anéis a informação da distância ao centro, dado que ela aumenta nas células em faixas mais altas de $|\eta|$. A influência da granularidade pode ser observada na variável de calorimetria W_{η_2} , empregada pelo ATLAS para calcular a largura do chuveiro,

definida pela Equação 5.1:

$$W_{\eta 2} = \sqrt{\frac{\sum E_i \eta_i}{\sum E_i} - \left(\frac{\sum E_i \eta_i}{\sum E_i} \right)^2}, \quad (5.1)$$

sendo E_i a energia da célula i e η_i o valor da pseudorapidez da célula i .

Na Figura 5.9 são apresentadas as distribuições da grandeza $W_{\eta 2}$ para elétrons e jatos hadrônicos, na subdivisão do espaço de fase $E_{T2} \times \eta_0$ (Figura 5.9 (a)), e $E_{T2} \times \eta_4$ (Figura 5.9 (b)). É possível observar, por exemplo, que a largura do Chuveiro para elétrons é menor em η_0 , quando comparada à distribuição de $W_{\eta 2}$ para elétrons em η_4 . O mesmo ocorre com a distribuição de jatos, que na região do intervalo η_4 apresenta limites da distribuição entre $\approx 0,010$ e $0,020$, enquanto em η_0 ocorrem menores valores para a largura do Chuveiro, iniciando a distribuição em $\approx 0,005$.

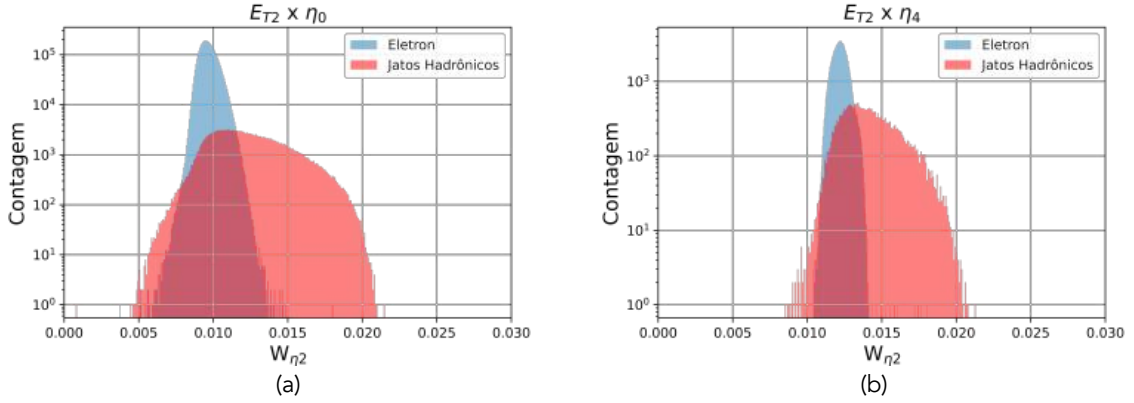


Figura 5.9: Distribuição da grandeza $W_{\eta 2}$ para elétrons e jatos hadrônicos, na subdivisão do espaço de fase $E_{T2} \times \eta_0$ (a) e $E_{T2} \times \eta_4$ (b).

Os valores médios de β encontrados para a região do *crack* (η_2) são, em geral, os mais elevados. Observa-se que a granularidade na primeira camada eletromagnética ($\Delta\eta \times \Delta\phi = 0,50 \times 0,1$), nesta região é a menor (células com maior área) se comparado as demais regiões do espaço de fase, assim, a hipótese de que o pré-processamento tenta aumentar a contribuição do coeficiente que pondera a distância ao centro, concorda com o aumento do valor de β ao longo da coordenada $|\eta|$, conforme discutido no parágrafo anterior. A largura do Chuveiro, descrita pela grandeza $W_{\eta 2}$ para a região do *crack* pode ser observada na Figura 5.10. Pode-se observar que na região do *crack*, a largura do Chuveiro para ambas as assinaturas de elétrons e jatos, apresentam perfil espalhado. A redução da granularidade e da quantidade de sensores nesta região ocasiona perfis com pouca distinção para esta grandeza de calorimetria.

Após o treinamento e seleção dos modelos com maior índice SP em cada sorteio, os conjuntos de sinais de todo o espaço de fase foram propagados pelos dez classificadores selecionados na etapa de validação cruzada. Esta estratégia foi realizada

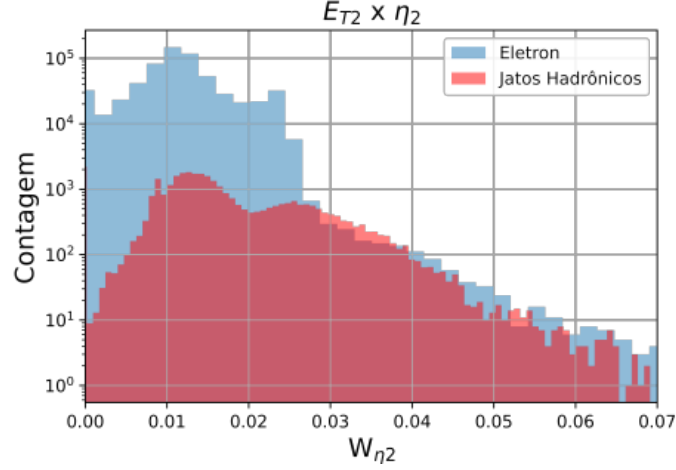


Figura 5.10: Distribuição da grandeza W_{η_2} para elétrons e jatos hadrônicos, na subdivisão do espaço de fase $E_{T2} \times \eta_2$ (*crack*).

para verificar a eficiência geral da etapa rápida de calorimetria com todo o conjunto de dados disponível.

Na Tabela 5.11 são apresentados os desempenhos obtidos pelos classificadores NeuralRinger, RingRp e RingRp-CMAES, considerando o ponto de operação *Tight* em cada espaço de fase. É observado que o método RingRp apresenta uma redução na taxa de falso alarme de 0,54 ponto percentual ($P_F[\%] = 0,79 \pm 0,01$) quando comparado ao resultado encontrado pelo NeuralRinger ($P_F[\%] = 1,33 \pm 0,05$).

Tabela 5.11: Valores dos indicadores de desempenho $SP[\%]$, $P_D[\%]$ e $P_F[\%]$, usando todo o conjunto de dados disponível como entrada para os classificadores do *ensemble*, com taxa de detecção ajustada para o ponto de operação *Tight*.

Método	$SP [\%]$	$P_D [\%]$	$P_F [\%]$
NeuralRinger	$98,86 \pm 0,00$	$99,05 \pm 0,00$	$1,33 \pm 0,05$
RingRp	$99,12 \pm 0,01$	$99,05 \pm 0,01$	$0,79 \pm 0,01$
RingRp-CMAES	$98,87 \pm 0,01$	$99,05 \pm 0,01$	$1,23 \pm 0,01$

Na Tabela 5.12 é apresentado a eficiência geral, considerando o ponto de operação *VeryLoose* usado em cada espaço de fase. É observado que o método RingRp apresenta uma redução na taxa de falso alarme de 1,37 ponto percentual ($P_F[\%] = 1,49 \pm 0,01$) quando comparado ao resultado encontrado pelo NeuralRinger ($P_F[\%] = 2,86 \pm 0,17$).

Tabela 5.12: Valores dos indicadores de desempenho $SP[\%]$, $P_D[\%]$ e $P_F[\%]$, usando todo o conjunto de dados disponível como entrada para os classificadores do *ensemble*, com taxa de detecção ajustada para o ponto de operação *VeryLoose*.

Método	$SP [\%]$	$P_D [\%]$	$P_F [\%]$
NeuralRinger	$98,40 \pm 0,00$	$99,67 \pm 0,00$	$2,86 \pm 0,17$
RingRp	$99,09 \pm 0,01$	$99,67 \pm 0,01$	$1,49 \pm 0,11$
RingRp-CMAES	$98,46 \pm 0,02$	$99,67 \pm 0,02$	$2,74 \pm 0,12$

5.2 Ajuste da Função de Saída do Neurônio pelo Nível de Empilhamento

A medida que o número médio de colisões por cruzamento de feixes aumenta, o empilhamento de sinais tende a ser crescente, podendo deteriorar o sinal medido pelos sensores no calorímetro [18]. O ajuste da saída do classificador neural em função do índice de empilhamento, permite manter a eficiência de detecção dos classificadores, mesmo com a variação da taxa de empilhamento³. Neste contexto, foram selecionados para cada espaço de fase o classificadores com melhor desempenho (índice SP) em cada ponto de operação.

Para realizar este procedimento, a função tangente hiperbólica do neurônio de saída é retirada e substituída por uma função linear, cuja inclinação em relação ao patamar de decisão é ajustada para manter a mesma eficiência de detecção em função das variações de empilhamento. A Figura 5.11 (a) e Figura 5.11 (b) apresentam os histogramas de saída do classificador neural, que opera no ponto de operação *Tight* ($P_D[\%] = 99,18$). O classificador foi treinado pelo método RingRp operando no espaço de fase $E_{T2} \times \eta_0$.

É possível observar que, o patamar de decisão do classificador é ajustado (pontos azuis) no intervalo entre ≈ 1 e $2,1$, para manter a mesma taxa de P_D , conforme a variação da taxa de empilhamento. A medida que o empilhamento de sinais aumenta, o classificador apresenta maior erro de confusão (elétrons sendo classificados como jatos) sendo necessário a diminuição do limiar de decisão para classificar corretamente os as assinaturas de elétrons.

Em alguns casos, o procedimento de ajuste da reta pelo índice de empilhamento pode não apresentar melhorias na P_D . Este fato ocorre pois a reta pode não passar exatamente pelos patamares de decisão, que garantem a mesma taxa de eficiência de detecção (pontos azuis). Em alguns casos, os pontos podem permanecer a es-

³Em alguns casos, a degradação dos sinais devido ao aumento do empilhamento pode confundir os classificadores, reduzindo sua eficiência.

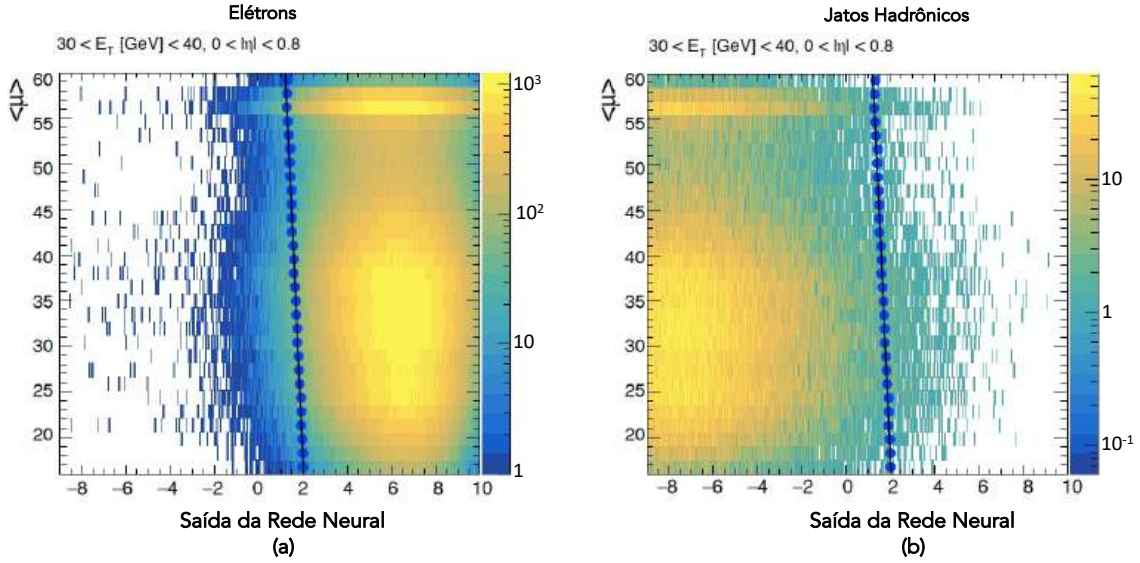


Figura 5.11: Histogramas bidimensionais para os sinais de saída do classificador neural para elétrons (a) e jatos hadrônicos (b). O classificador atua no espaço de fase $E_{T2} \times \eta_0$ no ponto de operação *Tight*. A curva em azul representa o ajuste do limiar de decisão da rede em função do nível de empilhamento, para manter a taxa de $P_D[\%] = 99,18$.

querda da reta (em cor preta), e ao utilizar o limiar proposto pela correção (pontos da abscissa da reta), a taxa de P_D pode cair ligeiramente. Nestes casos, onde não há melhorias na taxa de falso alarme para a mesma eficiência de detecção da referência (NeuralRinger), o procedimento de ajuste não é adotado e a função tangente hiperbólica é mantida na saída do classificador.

Os resultados do ajuste da função de saída em função do índice de empilhamento, para os classificadores no ponto de operação *Tight*, em todo o espaço de fase, podem ser observados na Tabela 5.13. As células com seta para baixo e para cima indicam, em pontos percentuais, a diminuição e o aumento respectivamente, dos índices P_D e P_F se comparado ao resultado obtido pelo classificador, com a função de ativação tangente hiperbólica. As células marcadas em cor vermelha, indicam as subdivisões do espaço de fase nas quais o procedimento de ajuste da função linear apresentou piora no desempenho do classificador.

No espaço de fase $E_{T2} \times \eta_0$ (barril central), o ajuste da função de saída pelo índice de empilhamento ocasionou uma redução na taxa de falso alarme, se comparada com o mesmo classificador sem o ajuste, resultando em uma nova taxa de $P_F[\%] = 0,78$ (0,09 ponto percentual menor que o classificador sem o ajuste). Em alguns casos, mesmo que o procedimento de ajuste tenha conseguido manter a mesma taxa de $P_D[\%]$ em relação a referência, eventualmente a taxa de falso alarme pode aumentar, como é observado no caso do espaço de fase $E_{T4} \times \eta_1$, no qual manteve-se a taxa de

		η_0	η_1	η_2	η_3	η_4
E_{T0}	P_D [%]	97,74 ($\downarrow$ -0,04)	97,73 ($\downarrow$ -0,05)	95,80	97,36 ($\downarrow$ -0,04)	92,93 ($\downarrow$ -0,11)
	P_F [%]	2,18	3,77	7,02 ($\downarrow$ -0,91)	3,53	4,36
E_{T1}	P_D [%]	98,66 ($\downarrow$ -0,02)	98,87 ($\downarrow$ -0,02)	94,44 ($\downarrow$ -0,08)	98,20 ($\downarrow$ -0,03)	95,65 ($\downarrow$ -0,03)
	P_F [%]	1,46	2,80	2,69	2,75	3,40
E_{T2}	P_D [%]	99,18	99,43	96,78 ($\downarrow$ -0,04)	98,96 ($\downarrow$ -0,02)	96,64 ($\downarrow$ -0,06)
	P_F [%]	0,78 ($\downarrow$ -0,09)	1,90 ($\downarrow$ -0,02)	1,95	2,04	2,33
E_{T3}	P_D [%]	99,57	99,68	97,23 ($\downarrow$ -0,04)	99,22	97,08 ($\downarrow$ -0,05)
	P_F [%]	0,73 ($\downarrow$ -0,01)	1,49 ($\downarrow$ -0,04)	1,62	1,56 ($\downarrow$ -0,02)	1,58
E_{T4}	P_D [%]	99,63	99,67	98,25 ($\downarrow$ -0,04)	99,32	95,95 ($\downarrow$ -0,07)
	P_F [%]	0,24 ($\downarrow$ -0,02)	0,65 ($\uparrow$ +0,08)	0,43	1,34 ($\downarrow$ -0,04)	0,72

Tabela 5.13: Taxas de P_D e P_F obtidas após o ajuste da função de saída dos classificadores neurais em função da taxa de empilhamento, para o ponto de operação *Tight*.

P_D [%] = 99,67, entretanto o falso alarme aumentou em 0,08 ponto percentual, se comparado ao mesmo classificador com a função tangente hiperbólica no neurônio de saída. Neste caso, a troca da função de saída não é realizada, para a operação do classificador no HLT do ATLAS.

A Figura 5.12 apresenta o histograma de saída do classificador neural para elétrons 5.12 (a) e jatos hadrônicos 5.12 (b). O classificador foi treinado para operação no espaço de fase $E_{T2} \times \eta_0$, atuando no ponto de operação *VeryLoose*. É possível observar que a maior parte dos pontos azuis (patamares de decisão que mantêm a taxa de P_D constante frente ao índice de empilhamento de sinais, permanece próxima da reta linear (em cor preta) ajustada. Neste caso, a taxa de falso alarme foi reduzida, em 0,04 ponto percentual quando comparada à taxa de falso alarme obtida pelo mesmo classificador, sem realizar o ajuste da função de saída do neurônio.

Na Tabela 5.14, são apresentados os resultados de P_D e P_F obtidos com o ajuste da função linear do neurônio de saída, em função do índice de empilhamento para o ponto de operação *VeryLoose*. Observa-se que o procedimento de ajuste da função linear de saída pelo índice de empilhamento, contribuiu para a redução do falso alarme na região central do barril ($0,00 \leq |\eta| < 0,80$), para os intervalos E_{T0} , E_{T1} , E_{T2} e E_{T3} respectivamente, comparados a taxa de falso alarme do mesmo classificador, com a função de ativação tangente hiperbólica no neurônio de saída. As reduções em pontos percentuais estão indicadas com as setas para baixo. Como pode ser observado na Tabela 5.14, as células em vermelho indicam os espaços de fase que não receberam contribuições para redução do falso alarme, mantendo a eficiência do ponto de operação do classificador.

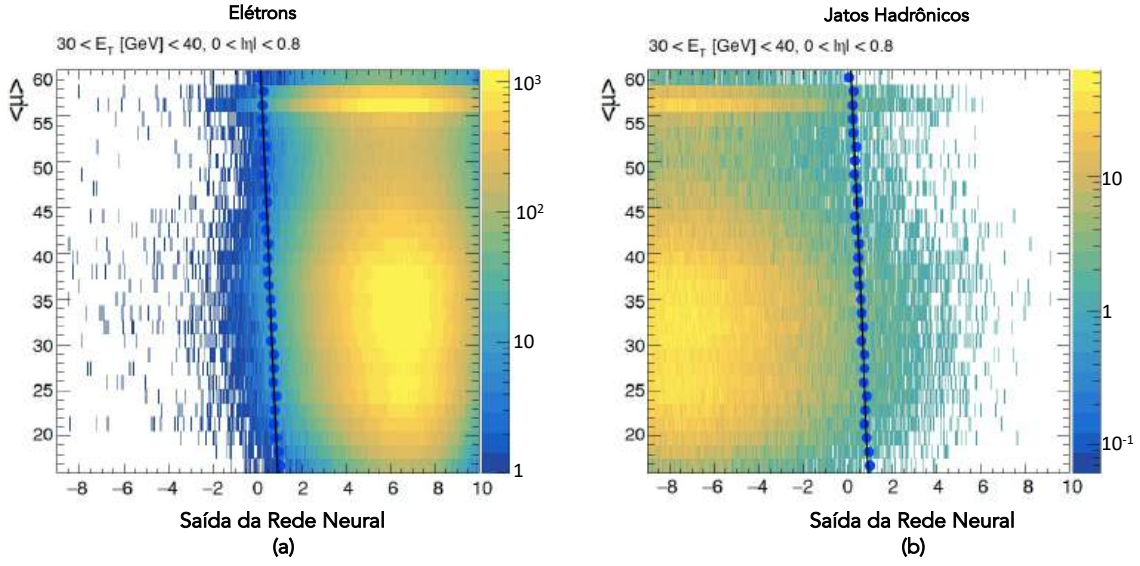


Figura 5.12: Histogramas bidimensionais para os sinais de saída do classificador neural para elétrons (a) e jatos hadrônicos (b). O classificador atua no espaço de fase $E_{T2} \times \eta_0$ no ponto de operação *VeryLoose*. A curva em azul representa o ajuste do limiar de decisão da rede em função do nível de empilhamento, para manter a taxa de $P_D[\%] = 99,18$.

		η_0	η_1	η_2	η_3	η_4
E_{T0}	$P_D[\%]$	98,86	98,78	97,62	98,22 ($\downarrow$ -0,03)	93,49 ($\downarrow$ -0,10)
	$P_F[\%]$	3,12 ($\downarrow$ -0,20)	5,17 ($\downarrow$ -0,21)	13,85 ($\downarrow$ -1,15)	4,18	4,59
E_{T1}	$P_D[\%]$	99,50	99,61 ($\downarrow$ -0,02)	96,66 ($\downarrow$ -0,05)	99,56 ($\downarrow$ -0,01)	95,68 ($\downarrow$ -0,05)
	$P_F[\%]$	2,25 ($\downarrow$ -0,08)	4,26	3,58	4,39	3,41
E_{T2}	$P_D[\%]$	99,83	99,91	97,94 ($\downarrow$ -0,03)	99,92	96,65 ($\downarrow$ -0,06)
	$P_F[\%]$	1,70 ($\downarrow$ -0,04)	3,57 ($\downarrow$ -0,13)	2,43	3,93 ($\downarrow$ -0,23)	2,34
E_{T3}	$P_D[\%]$	99,94	99,97	97,85 ($\downarrow$ -0,03)	99,97	97,08 ($\downarrow$ -0,05)
	$P_F[\%]$	1,60 ($\downarrow$ -0,04)	3,37 ($\downarrow$ -0,25)	1,78	3,46 ($\downarrow$ -0,33)	1,58
E_{T4}	$P_D[\%]$	99,98	99,93	98,58 ($\downarrow$ -0,02)	99,98	95,95 ($\downarrow$ -0,07)
	$P_F[\%]$	0,97 ($\uparrow$ +0,04)	1,20 ($\uparrow$ +0,07)	0,50	4,89 ($\downarrow$ -0,12)	0,72

Tabela 5.14: Taxas de P_D e P_F obtidas após o ajuste da função de saída dos classificadores neurais em função da taxa de empilhamento, para o ponto de operação *VeryLoose*.

5.3 Comentários e Discussão

Dentre os métodos avaliados, os algoritmos RingRp (método baseado em retropropagação) e o RingRp-CMAES (método que utiliza como otimizador o algoritmo CMA-ES) obtiveram desempenho superior ao NeuralRinger, com a redução na taxa de falsos elétrons nos pontos de operação *Tight* e *VeryLoose*. Com os métodos propostos foi possível obter maior rejeição ao falso alarme na região do barril central (região do detector com maior granularidade de células), como também em regiões

do calorímetro que apresentam número reduzido de sensores (*crack* e interconexões dos módulos). Comparando os métodos desenvolvidos, o método RingRp encontrou melhor desempenho, em geral, para todo o espaço de fase definido para operação no HLT. Melhorias de desempenho também foram observadas em faixas mais baixas de energia ($15 \leq E_T < 20$), na qual apresentam menor razão sinal/interferência.

Os resultados obtidos para os coeficientes α e β da etapa de pré-processamento, apresentaram distintos valores em função da energia dos anéis e da sua posição na coordenada η . Em média, os valores obtidos para o coeficiente α diminuem a medida que a faixa de energia aumenta. Em relação aos resultados encontrados para o coeficiente β foi possível observar maiores valores em média, para a região do *crack*, e para as regiões η_3 e η_4 . Esse resultado é um indicativo que o mapeamento R_p consegue capturar nos anéis a informação da largura no chuveiro, principalmente em regiões nas quais os anéis são gerados a partir de células com menor granularidade. A diferença da largura do chuveiro entre a área central do barril e área coberta em η_4 pôde ser observada pela distribuição da grandeza unidimensional W_{η_2} , atualmente empregada pelo ATLAS para descrever a largura do chuveiro. Observou-se que na região η_4 são encontrados maiores valores para W_{η_2} , indicando a influência da granularidade para a largura do chuveiro.

Para avaliar, de modo integrado, os métodos que obtiveram melhor desempenho, todo o conjunto de dados foi propagado pelos classificadores selecionados no procedimento de validação cruzada. Os resultados indicaram que, para os classificadores selecionados para operação no ponto de operação *Tight*, os métodos propostos RpRing e RpRing-CMAES encontraram uma redução de 0,54 e 0,1 ponto percentual respectivamente, se comparados ao método NeuralRinger. No ponto de operação *VeryLoose*, também notou-se uma redução do falso alarme com a utilização dos métodos propostos.

Um procedimento de ajuste da função de saída dos classificadores neurais treinados foi realizado para avaliar a influência do empilhamento de sinais no desempenho dos classificadores. A função de ativação tangente hiperbólica do neurônio de saída foi substituída por uma função linear, onde seus coeficientes foram ajustados para permitir que o classificador mantenha a mesma taxa de detecção (P_D), para as distintas taxas de empilhamento que ocorrem durante as colisões do LHC. Observou-se, em alguns casos, que o efeito do aumento de empilhamento pode confundir o classificador, requerendo modificações no limiar de decisão para não classificar incorretamente algumas assinaturas de elétrons. Algumas subdivisões do espaço de fase não apresentaram melhora com o procedimento de substituição da função de ativação do neurônio de saída, indicando diminuição na P_D ou aumento na P_F , quando comparado com o classificador com a tangente hiperbólica. Nestas regiões, como o procedimento provocou piora nos resultados, a função tangente hiperbólica

é utilizada no neurônio de saída para a implementação do classificador neural na infraestrutura de computação do HLT do ATLAS.

Capítulo 6

Conclusões e Trabalhos Futuros

Experimentos de física de altas energias (HEP - *High Energy Physics*) se propõem a investigar as características fundamentais da matéria, e ao longo dos anos equipamentos cada vez mais complexos vêm sendo construídos. Esse esforço multidisciplinar envolve profissionais de diversas áreas como física, matemática, química, engenharia e computação para o projeto, instalação, validação e operação. Muitos experimentos apresentam desafios crescentes em engenharia e computação, entre os quais pode-se destacar processamento de grandes volumes de dados, tratamento de sinais com alta dimensionalidade e classificação de eventos raros. Esse contexto é propício para o desenvolvimento e aplicação de técnicas de inteligência computacional, que nos últimos anos tem ganhado espaço para atuação principalmente na seleção *online* de eventos.

O experimento ATLAS (*A Toroidal LHC ApparatuS*), que atua no Grande Colisor de Hádrons (LHC - *Large Hadron Collider*), é o maior detector de propósito geral do LHC, sendo dotado de um sistema de seleção *online* de eventos que tem por objetivo selecionar partículas de interesse para o experimento, removendo a maior parte da informação já conhecida (ruído de fundo) que é produzida nas colisões do LHC. O programa de atualizações do LHC prevê para as próximas tomadas de dados, o crescimento da luminosidade (número de colisões por centímetro quadrado por segundo) e da energia das colisões, com o objetivo de aumentar a probabilidade de ocorrência de eventos físicos ainda não observados experimentalmente. Esse contexto exige que os sistemas de seleção sejam atualizados, para suportar a crescente taxa de produção da informação e o aumento do empilhamento de sinais nos sensores.

Dentre os diversos canais de eventos físicos investigados pelo ATLAS, elétrons são partículas que apresentam forte interesse para o experimento, pois são relacionados com o decaimento de raros fenômenos físicos. Para realizar a detecção de elétrons, o ATLAS possui um sistema de seleção *online* (*trigger*) dotado de dois

níveis sequenciais. No segundo nível de *trigger*, algoritmos de classificação e reconstrução realizam o processamento dos sinais medidos pelo sistema de calorimetria, que mede a energia das partículas. Em 2017, o ATLAS adotou o método NeuralRinger¹ como algoritmo principal para classificação de elétrons, na etapa rápida do segundo nível de *trigger*. O NeuralRinger realiza o processamento topológico do perfil de deposição de energia da partícula no sistema de calorimetria extraindo, em uma geometria aproximadamente cônica, informações do espalhamento lateral e longitudinal da energia da partícula que interage com os calorímetros. Na primeira etapa do método, os sinais medidos são organizados em anéis concêntricos em torno da célula com maior deposição de energia. Em seguida, os sinais servem de entrada para um conjunto de classificadores neurais treinados para atuação em intervalos específicos de energia e posição de interação da partícula no calorímetro (espaços de fase).

Elétrons e jatos hadrônicos, apresentam distintas características de espalhamento lateral de energia devido aos diferentes fenômenos envolvidos no processo de interação com a matéria. De modo a destacar nos sinais em anéis as diferenças do espalhamento lateral de energia dessas partículas, este trabalho de doutorado propôs a adoção de uma técnica de pré-processamento, que pondera a energia e a distância dos anéis com relação ao primeiro anel. Para otimizar os coeficientes que ponderam a energia e a distância do anel ao centro, foram propostos neste trabalho um conjunto de métodos de treinamento de redes neurais que otimizam simultaneamente os pesos da rede e os parâmetros livres da etapa de pré-processamento.

O treinamento das redes neurais foi realizado através de dois métodos. Um método baseado em algoritmos bio-inspirados, foi desenvolvido para otimizar o desempenho dos classificadores em pontos de operação específicos. Três algoritmos de otimização bio-inspirados foram avaliados (GA - Algoritmo Genético, PSO - Otimização por Enxame de Partículas e o CMAES - Estratégia Evolutiva por Adaptação da Matriz de Covariância). Uma função objetivo foi projetada para minimizar a taxa de falsos elétrons, dada a restrição da mínima probabilidade de detecção para atuação no sistema de *trigger*. Essa estratégia foi aplicada também com a otimização simultânea dos parâmetros da etapa de pré-processamento. Um método baseado em retropropagação (RingRp) também foi desenvolvido para ajuste simultâneo dos pesos da rede e dos coeficientes da etapa de pré-processamento. A ideia principal

¹O NeuralRinger evoluiu à partir de uma proposta de longa data do grupo brasileiro ATLAS-Brasil, cujo desenvolvimento vem sendo aprimorado nos últimos 30 anos. Os primeiros trabalhos visando a utilização de anéis para classificação de elétrons, foram apresentados ainda na fase de projeto do ATLAS [22]. Desde então, muitos pesquisadores tem contribuído para a melhoria de desempenho do método e extensão de sua aplicação conforme a atualização dos requisitos de operação do HLT.

destes desenvolvimentos foi permitir que o conhecimento especialista em calorimetria (largura do chuveiro) seja utilizado para guiar o algoritmo de aprendizagem, para uma melhor eficiência na classificação de elétrons.

Foi observado que os anéis da primeira camada eletromagnética foram destacados com a aplicação da etapa de pré-processamento, principalmente nas assinaturas de jatos hadrônicos. Este comportamento concorda com aspectos teóricos, que indicam um maior espalhamento lateral para as assinaturas de jatos. O pré-processamento foi capaz de destacar essa informação nos anéis, beneficiando o processo discriminativo entre elétrons e jatos. Os valores encontrados com o método RingRp para o coeficiente β (que pondera a distância do anel ao centro) apresentaram em média, um aumento à medida que o ponto de interação da partícula se afasta da região central do barril. Este resultado pode estar relacionado com a diminuição da granularidade das células nas regiões mais externas em η , onde aumenta-se a distância dos anéis ao centro devido à maior área das células do calorímetro nestas regiões. Os resultados indicam que o mapeamento R_p pôde capturar essa informação nos sinais em anéis, fazendo com que a distância ao centro tenha maior ponderação no processo discriminativo.

Adicionalmente às investigações realizadas para melhoria do sistema de *trigger*, o autor esteve envolvido no desenvolvimento da infraestrutura de computação do ATLAS, como parte do programa de atualização do experimento para a *Run3*. Uma demanda crescente do ambiente de operação do ATLAS é a certificação de que os sinais que são coletados durante as colisões estão livres de problemas de *hardware* ou *software*. Para isso, o HLT conta com um sistema de monitoração *online*, onde as principais grandezas que descrevem o perfil de deposição de energia são verificadas, e suas distribuições são exibidas para especialistas em qualidade de dados que acompanham a tomada de dados na sala de controle do ATLAS. Neste trabalho, foram desenvolvidos sistemas computacionais para monitoração *online* das *shower shapes* nos algoritmos de classificação e reconstrução do HLT. Os algoritmos desenvolvidos monitoram as saídas dos algoritmos de reconstrução, exportando para o sistema de avaliação e qualidade de dados as grandezas de interesse. Foram desenvolvidos algoritmos para cadeias de assinaturas de elétrons e fótons. Os algoritmos desenvolvidos foram validados e atualmente estão em operação no *software* principal de análise do experimento².

Considerando o bom desempenho encontrado pela aplicação da técnica de pré-

²O HLT é baseado no ambiente de *software offline* chamado Athena [125]. O Athena tem como principal infraestrutura de *software* o Gaudi [126], que consiste em um *framework* para processamento de dados para experimentos HEP. O Athena é baseado em componentes, onde cada um deles (por exemplo, algoritmo, serviço, ferramenta) é configurado por um conjunto de propriedades que podem ser definidas durante ao longo de uma aplicação.

processamento especialista, em conjunto com os métodos de treinamento propostos, novas abordagens podem ser exploradas em trabalhos futuros. Neste trabalho, o pré-processamento foi aplicado após a geração dos sinais em anéis. No entanto, o pré-processamento também pode ser realizado diretamente na informação das células, em uma etapa anterior a geração dos anéis. Essa estratégia pode oferecer vantagens para captura da informação lateral do chuveiro, uma vez que estaria realizando a leitura de mais sensores, entretanto pode demandar maior esforço computacional, que deve ser investigado. No trabalho [22], essa proposta ofereceu bons resultados para separação elétron/píon, quando aplicada em um calorímetro com uma única camada.

Nos trabalhos [46, 127] classificadores neurais treinados com a informação de camada do calorímetro foram investigados, e obtiveram bom desempenho se comparado aos classificadores neurais que recebem como entradas todos os anéis de modo concatenado. A saída dos 7 classificadores foi combinada por uma rede neural, que realizou a fusão da informação de cada classificador (abordagem segmentada). Uma estratégia que pode ser investigada, é o desenvolvimento de um classificador para cada camada, com o uso do pré-processamento especialista. Neste caso, os parâmetros α e β seriam otimizados para cada camada, explorando assim as diferenças de granularidade no espaço $\eta \times \phi$ de cada camada.

Neste trabalho, as distâncias entre os anéis foram consideradas unitárias, em cada camada onde considerou-se a posição da região de interesse como ponto de início da contagem. Esse procedimento indica uma mesma distância relativa ao centro em todas as camadas, não considerando os distintos tamanhos das células nas quais os anéis são formados. Uma abordagem que pode ser explorada é normalizar o vetor de distâncias pela células com maior dimensão. Isso permitiria capturar as diferenças entre as granularidades presentes em cada camada e mapear essa informação no vetor de distâncias dos anéis ao centro.

Em relação aos métodos de treinamento de redes neurais propostos, os resultados indicaram que o método de treinamento que otimiza os pesos da rede para um determinado ponto de operação obteve bom desempenho, quando comparado ao NeuralRinger. O método desenvolvido utilizou uma função objetivo que combina a minimização da taxa de falso alarme instantânea (na atual iteração), a divergência KL entre a saída da rede e a saída desejada, e um parâmetros de regularização em um único objetivo. Como os diferentes objetivos apresentam distintas escalas, o treinamento pode ser polarizado para atendimento do objetivo que apresente maior peso. Neste caso, para melhoria de desempenho, trabalhos futuros podem explorar uma proposta que utilize versões multi-objetivo dos algoritmos de otimização bio-inspirados utilizados.

O método de treinamento bio-inspirado proposto neste trabalho, encontra aplica-

ções também em outros contextos científicos, onde restrições de desempenho tenham impacto na operação de sistemas automáticos de classificação. No trabalho [128], uma extensão do método de treinamento com restrições foi apresentada, no contexto de suporte ao diagnóstico médico. Neste caso, o método proposto apresentou aumento de eficiência de classificação em problemas binários de detecção de doenças, considerando restrições para taxas de falsos positivos. Nos trabalhos [84, 129], o método de treinamento de redes neurais foi expandido para uma abordagem de classificação multiclasse, em problemas de identificação de defeitos em tubulações industriais. Foram codificadas restrições de desempenho para atendimento a diferentes taxas mínimas aceitáveis de falsos alarmes, em cada classe, guiando o treinamento do classificador para atendimento a estas restrições. O método proposto apresentou bom desempenho se comparado a diferentes algoritmos de otimização, baseados em retropropagação. Considerando a flexibilidade do método proposto, sua aplicação pode ser estendida a diferentes problemas científicos de classificação, nos quais restrições severas de operação precisam ser atendidas.

Referências Bibliográficas

- [1] EASSA, M., MOHAMED SELIM, I., DABOUR, W., et al., “Automated detection and classification of galaxies based on their brightness patterns”, *Alexandria Engineering Journal*, 2021.
- [2] CHARRIER, D., DALLIER, R., ESCUDIE, A., et al., “Radio detection of cosmic rays in [1.7–3.7] MHz: The EXTASIS experiment”, *Astroparticle Physics*, v. 113, pp. 6–21, 2019.
- [3] SITAREK, J., SOBCZYŃSKA, D., SZANECKI, M., et al., “Nature of the low-energy, γ -like background for the Cherenkov Telescope Array”, *Astroparticle Physics*, v. 97, pp. 1–9, 2018.
- [4] SHILON, I., KRAUS, M., BÜCHELE, M., et al., “Application of deep learning methods to analysis of imaging atmospheric Cherenkov telescopes data”, *Astroparticle Physics*, v. 105, pp. 44–53, 2019.
- [5] AARTSEN, M., ACKERMANN, M., ADAMS, J., et al., “Neutrinos below 100 TeV from the southern sky employing refined veto techniques to IceCube data”, *Astroparticle Physics*, v. 116, pp. 102392, 2020.
- [6] GARCIA, R., ANZORENA, M., VALDÉS-GALICIA, J., et al., “Particle identification and analysis in the SciCRT using machine learning tools”, *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, v. 1003, pp. 165326, 2021.
- [7] ATLAS-COLLABORATION, “The ATLAS Experiment at the CERN Large Hadron Collider”, *Journal of Instrumentation*, v. 3, n. 08, pp. S08003–S08003, aug 2008.
- [8] EVANS, L., BRYANT, P., “LHC Machine”, *Journal of Instrumentation*, v. 3, n. 08, pp. S08001–S08001, aug 2008.

- [9] ATLAS-COLLABORATION, “Evidence for Higgs boson decays to a low-mass dilepton system and a photon in pp collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector”, *Physics Letters B*, v. 819, pp. 136412, 2021.
- [10] ATLAS-COLLABORATION, “Electron efficiency measurements with the ATLAS detector using 2012 LHC proton–proton collision data”, *Eur. Phys. J. C*, v. 195, n. 77, pp. 1–45, 2017.
- [11] ATLAS-COLLABORATION, “Operation of the ATLAS trigger system in Run 2”, *Journal of Instrumentation*, v. 15, n. 10, pp. P10004–P10004, oct 2020.
- [12] ATLAS-COLLABORATION, “Projected ATLAS Electron and Photon Trigger Performance in Run 3”, *Proceedings of Science*, v. ICHEP2020, pp. 757, 2021.
- [13] ICECUBE-COLLABORATION, “Detection of a particle shower at the Glashow resonance with IceCube”, *Nature*, v. 591, n. 7849, pp. 220–224, 2021.
- [14] CMS-COLLABORATION, “The CMS experiment at the CERN LHC”, *Journal of Instrumentation*, v. 3, n. 08, pp. S08004–S08004, aug 2008.
- [15] ATLAS-COLLABORATION, “Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC”, *Physics Letters B*, v. 716, n. 1, pp. 1–29, 2012.
- [16] ATLAS-COLLABORATION, *Technical Design Report for the Phase-II Upgrade of the ATLAS Muon Spectrometer*, Tech. rep., CERN, Geneva, Sep 2017.
- [17] ATLAS-COLLABORATION, “Performance of the ATLAS trigger system in 2015”, *Eur. Phys. J. C*, v. 77, n. 317, pp. 1–53, 2017.
- [18] AAD, G., ABBOTT, OTHERS, “Performance of electron and photon triggers in ATLAS during LHC Run 2”, *Eur. Phys. J. C*, v. 80, pp. 47. 56 p, Sep 2019.
- [19] SPOLIDORO FREUND, W., COLLABORATION, A., *An Ensemble of Neural Networks for Online Electron Filtering at the ATLAS Experiment*, Tech. rep., CERN, Geneva, May 2019.
- [20] ACOSTA, D., BUONTEMPO, S., CALÔBA, L., et al., “Lateral shower profiles in a lead/scintillating fiber calorimeter”, *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, v. 316, n. 2, pp. 184–201, 1992.

- [21] WIGMANS, R., *Calorimetry: Energy Measurement in Particle Physics. International series of monographs on physics*, Clarendon Press, 2017.
- [22] SEIXAS, J., CALOBA, L., SOUZA, M., et al., “Neural second-level trigger system based on calorimetry”, *Computer Physics Communications*, v. 95, n. 2, pp. 143–157, 1996.
- [23] PAUL, A. L., BYRNE, P. C., “An efficient learning algorithm for the backpropagation artificial neural network”. In: *IEEE Proceedings on Southeastcon*, pp. 61–63, April 1990.
- [24] LIVAN, M., WIGMANS, R., *Calorimetry for Collider Physics, an Introduction. Uniftext for physics*, Springer, 2019.
- [25] SPINELLI, M., “A Instauração da Filosofia e da Ciência por Tales de Mileto”, *Revista Portuguesa de Filosofia*, v. 48, n. 1, pp. 99–124, 1992.
- [26] CARTLEDGE, P., *Demócrito: Demócrito e a Política atomista*. Editora Unesp: São Paulo, 2001.
- [27] DALTON, J., “III. On the absorption of gases by water and other liquids”, *The Philosophical Magazine*, v. 24, n. 93, pp. 15–24, 1806.
- [28] PHYSICS, N. L., “Nobel Prize in Physics 1903”, 2021.
- [29] OUTREACH, N. P., “Nobel Prize in Chemistry 1911”, 2021.
- [30] BOHR, N., “On the Constitution of Atoms and Molecules, Part I”, *Philosophical Magazine*, v. 26, pp. 1–25, 1913.
- [31] PACCHIONI, G., “CERN’s next big thing”, *Nature Reviews Physics*, v. 1, pp. 1–4, 2019.
- [32] SCHMITZ, W., *Particle, Field and Forces - A Conceptual Guide to Quantum Field Theory and the Standard Model. The Frontiers Collection*, Springer, 2019.
- [33] WIEDEMANN, H., *Particle Accelerator Physics II - NonLinear and Higher-Order Beam Dynamics*. Springer, 1999.
- [34] BRÜNING, O., BURKHARDT, H., MYERS, S., “The Large Hadron Collider”, *Progress in Particle and Nuclear Physics*, v. 67, n. 3, pp. 705–734, 2012.
- [35] WIEDEMANN, H., *Particle Accelerator Physics*. Springer, 2007.

- [36] ICECUBE-COLLABORATION, “The IceCube Neutrino Observatory: instrumentation and online systems”, *Journal of Instrumentation*, v. 12, n. 03, pp. P03012–P03012, mar 2017.
- [37] ACHARYYA, A., ADAM, R., ADAMS, C., et al., “Sensitivity of the Cherenkov Telescope Array to a dark matter signal from the Galactic centre”, *Journal of Cosmology and Astroparticle Physics*, v. 2021, n. 01, pp. 057–057, jan 2021.
- [38] ACHARYYA, A., AGUDO, I., ANGÜNER, E., et al., “Monte Carlo studies for the optimisation of the Cherenkov Telescope Array layout”, *Astroparticle Physics*, v. 111, pp. 35–53, 2019.
- [39] STEFANIK, S., NOSEK, D., DE LOS REYES, R., et al., “Atmospheric monitoring and inter-calibration of the telescope optical throughput efficiencies using the trigger rates of the Cherenkov Telescope Array”, *Astroparticle Physics*, v. 109, pp. 12–24, 2019.
- [40] WEINBERG, S., *The Quantum Theory of Fields. Vol1 Foundations*. Cambridge University Press, 2005.
- [41] WEINBERG, S., *The Quantum Theory of Fields. Vol2 Modern Applications*. Cambridge University Press, 2013.
- [42] ARBUZOV, A. B., “Quantum Field Theory and the Electroweak Standard Model”, *Proceedings, European School of High-Energy Physics*, 2017.
- [43] MOUCHE, P., “Overall view of the LHC.” *CERN Document Server*, Jun 2014, General Photo.
- [44] CASADO, M., “HL-LHC physics with the ATLAS detector”, *Nuclear and Particle Physics Proceedings*, v. 309-311, pp. 43–48, 2020, QCD 19 The 22nd International Conference on Quantum Chromodynamics.
- [45] MARCASTEL, F., “CERN’s Accelerator Complex. La chaîne des accélérateurs du CERN”, *CERN Document Server*, Oct 2013, General Photo.
- [46] DE SIMAS FILHO, E. F., *Análise Não-Linear de Componentes Independentes para uma Filtragem Online Baseada em Calorimetria de Alta Energia e com Fina Segmentação*, Ph.D. Thesis, Universidade Federal do Rio de Janeiro, Poli/PEE/COPPE, Rio de Janeiro/RJ, 2010.
- [47] ATLAS-COLLABORATION, “Performance of the ATLAS muon triggers in Run 2”, *Journal of Instrumentation*, v. 15, n. 09, pp. P09015–P09015, sep 2020.

- [48] PEQUENAO, J., SCHAFFNER, P., “How ATLAS detects particles: diagram of particle paths in the detector”, General Photo.
- [49] ATLAS-COLLABORATION, *ATLAS Liquid Argon Calorimeter Phase-II Upgrade: Technical Design Report*, Tech. rep., CERN, Geneva, Sep 2017.
- [50] ARTAMONOV, A., BAILEY, D., BELANGER, G., et al., “The ATLAS Forward Calorimeter”, *Journal of Instrumentation*, v. 3, n. 02, pp. P02010–P02010, feb 2008.
- [51] ATLAS-COLLABORATION, *Technical Design Report for the Phase-II Upgrade of the ATLAS Tile Calorimeter*, Tech. rep., CERN, Geneva, Sep 2017.
- [52] ATLAS-COLLABORATION, *ATLAS tile calorimeter: Technical Design Report. Technical design report. ATLAS*, CERN: Geneva, 1996.
- [53] ATLAS-COLLABORATION, “The ATLAS Data Acquisition and High Level Trigger system”, *Journal of Instrumentation*, v. 11, n. 06, jun 2016.
- [54] ATLAS-COLLABORATION, “Operation of the ATLAS trigger system in Run 2”, *Journal of Instrumentation*, v. 15, n. 10, pp. P10004–P10004, oct 2020.
- [55] ACHENBACH, R., ADRAGNA, P., ANDREI, V., et al., “The ATLAS Level-1 Calorimeter Trigger”, *Journal of Instrumentation*, v. 3, n. 03, pp. P03001–P03001, mar 2008.
- [56] DOS ANJOS, A. R., *Sistema Online de Filtragem em um Ambiente com Alta Taxa de Eventos*, Ph.D. Thesis, Universidade Federal do Rio de Janeiro, Poli/PEE/COPPE, Rio de Janeiro/RJ, 2006.
- [57] TORRES, R. C., *Sistema Online de Filtragem em um Ambiente com Alta Taxa de Eventos e Fina Granularidade*, Ph.D. Thesis, Universidade Federal do Rio de Janeiro, Poli/PEE/COPPE, Rio de Janeiro/RJ, 2010.
- [58] XAVIER, T. C., *Identificação Online de Sinais Baseada em Calorimetria de Altas Energias e com Fina Segmentação*, Ph.D. Thesis, Universidade Federal do Rio de Janeiro, Poli/PEE/COPPE, Rio de Janeiro/RJ, 2012.
- [59] FERRAZ, V. A., *Identificação de Elétrons Baseada em um Calorímetro de Altas Energias Finamente Segmentado*, Ph.D. Thesis, Universidade Federal do Rio de Janeiro, Poli/PEE/COPPE, Rio de Janeiro/RJ, 2018.

- [60] DA FONSECA PINTO, J. V., *Ring-shaped Calorimetry Information for a Neural EGamma Identification with ATLAS Detector*, Tech. rep., CERN, Geneva, Mar 2016.
- [61] ATLAS-COLLABORATION, *An Ensemble of Neural Networks for Online Filtering Implemented in the ATLAS Trigger System*, Tech. rep., CERN, Geneva, Jul 2018.
- [62] DE ARAÚJO, M. V., *Filtragem Online Baseada em Calorimetria de Altas Energias e com Alta Taxa de Eventos*, Master’s Thesis, Universidade Federal do Rio de Janeiro, Poli/PEE/COPPE, Rio de Janeiro/RJ, 2019.
- [63] DE ARAÚJO, M. V., COSTA, C. E. C., DA FONSECA PINTO, J. V., et al., “Avaliação da Contribuição Longitudinal de um Calorímetro Finamente Segmentado para a Filtragem Online Realizada por um Ensemble de Redes Neurais em Ambiente de Alta Taxa de Eventos”. In: *Anais do 14 Congresso Brasileiro de Inteligência Computacional*, pp. 1–6, Belém, PA, 2019.
- [64] ATLAS-COLLABORATION, T., “Monitoring and data quality assessment of the ATLAS liquid argon calorimeter”, *Journal of Instrumentation*, v. 9, n. 07, pp. P07024–P07024, 2014.
- [65] ATLAS-COLLABORATION, “ATLAS data quality operations and performance for 2015–2018 data-taking”, *Journal of Instrumentation*, v. 15, n. 04, pp. P04003–P04003, apr 2020.
- [66] ILCHENKO, Y., ALMENAR, C. C., CORSO-RADU, A., et al., “Data Quality Monitoring Display for ATLAS experiment at the LHC”, *Journal of Physics: Conference Series*, v. 219, n. 2, pp. 022035, apr 2010.
- [67] DOTTI, A., ADRAGNA, P., VITILLO, R. A., “The Online Histogram Presenter for the ATLAS experiment: A modular system for histogram visualization”, *Journal of Physics: Conference Series*, v. 219, n. 3, pp. 032037, apr 2010.
- [68] NUNES, A. S., MAMASHLI, F., KOZHEMIAKO, N., et al., “Classification of evoked responses to inverted faces reveals both spatial and temporal cortical response abnormalities in Autism spectrum disorder”, *NeuroImage: Clinical*, v. 29, pp. 102501, 2021.
- [69] DE OLIVEIRA E SOUZA FILHO, J. B., SANCHEZ, M., DE SEIXAS, J. M., et al., “Screening for active pulmonary tuberculosis: Development and

- applicability of artificial neural network models”, *Tuberculosis*, v. 111, pp. 94–101, 2018.
- [70] SILVA, L. C., SIMAS FILHO, E. F., ALBUQUERQUE, M. C., et al., “Embedded decision support system for ultrasound nondestructive evaluation based on extreme learning machines”, *Computers and Electrical Engineering*, v. 90, pp. 106891, 2021.
- [71] OLIVEIRA, M. A., SIMAS FILHO, E. F., ALBUQUERQUE, M. C., et al., “Ultrasound-based identification of damage in wind turbine blades using novelty detection”, *Ultrasonics*, v. 108, pp. 106166, 2020.
- [72] SHANMUGAN, K. S., BREIPOHL, A. M., *Random Signals: Detection, Estimation and Data Analysis*. v. 1. John Wiley and Sons, 1988.
- [73] MARZBAN, C., “The ROC Curve and The Area Under It As Performance Measures”, *Weather and Forecasting*, v. 24, pp. 1106–1114, 2004.
- [74] LAFRATE, F., *Artificial Intelligence and Big Data*. v. 8. Wiley: London, 2018.
- [75] HAYKIN, S., *Neural Networks: Principles and Practice*. v. 3. Prentice Hall, 2008.
- [76] KHAN, G. M., *Evolution of Artificial Neural Development. In Search of Learning Genes*. v. 725. Springer: Switzerland, 2018.
- [77] DU, K.-L., SWAMY, M. N. S., *Neural Networks and Statistical Learning*. v. 1. Springer: Montreal, 2014.
- [78] RUMELHART, E., HINTON, G. E., WILLIAMS, R. J., “Learning representations by back-propagating errors.” *Letters to Nature*, pp. 533–536, 1986.
- [79] MENG, X., JIANG, J., WANG, H., “AGWO: Advanced GWO in multi-layer perception optimization”, *Expert Systems with Applications*, v. 173, pp. 114676, 2021.
- [80] AWAD, M., KHANNA, R., *Efficient Learning Machines*. v. 1. Apress, 2015.
- [81] ZHOU, X.-H., OBUCHOWSKI, N. A., MCCLISH, D. K., *Statistical Methods in Diagnostic Medicine*. John Wiley and Sons: New York, 2002.
- [82] SIMAS FILHO, E. F., DE SEIXAS, J. M., CALÔBA, L. P., “Modified post-nonlinear ICA model for online neural discrimination”, *Neurocomputing*, v. 73, n. 16, pp. 2820–2828, 2010, 10th Brazilian Symposium on Neural Networks (SBRN2008).

- [83] ATLAS-COLLABORATION, “Electron reconstruction and identification in the ATLAS experiment using the 2015 and 2016 LHC proton–proton collision data at $\sqrt{s} = 13$ TeV”, *Eur. Phys. J. C*, v. 80, pp. 47. 56 p, Sep 2019.
- [84] SILVA, G. F., PURCINO DE SOUZA, E., SIMAS FILHO, E., et al., “Constrained Neural Classifier Training Method for Flaw Detection in Industrial Pipes Using Particle Swarm Optimization”, *International Journal of Innovative Computing and Applications*, v. 107, 2019.
- [85] CUNHA, A. G., TAKAHASHI, R., ANTUNES, C. H., *Manual de Computação Evolutiva e Meta-Heurísticas*. Editora UFMG: Belo Horizonte, 2013.
- [86] AL-BADARNEH, I., HABIB, M., ALJARAH, I., et al., “Neuro-evolutionary models for imbalanced classification problems”, *Journal of King Saud University - Computer and Information Sciences*, 2020.
- [87] CZAKO, Z., SEBESTYEN, G., HANGAN, A., “AutomaticAI – A hybrid approach for automatic artificial intelligence algorithm selection and hyperparameter tuning”, *Expert Systems with Applications*, v. 182, pp. 115225, 2021.
- [88] IDOWU, O. P., ADELOPO, O., ILESANMI, A. E., et al., “Neuro-evolutionary approach for optimal selection of EEG channels in motor imagery based BCI application”, *Biomedical Signal Processing and Control*, v. 68, pp. 102621, 2021.
- [89] ROSTAMI, M., BERAHMAND, K., NASIRI, E., et al., “Review of swarm intelligence-based feature selection methods”, *Engineering Applications of Artificial Intelligence*, v. 100, pp. 104210, 2021.
- [90] HOLLAND, J. H., *Adaptation in Natural and Artificial Systems*. v. 2. The MIT Press, 1992.
- [91] CHAKRABORTY, A., KAR, A. K., *Swarm Intelligence: A Review of Algorithms*. Springer International Publishing, 2017.
- [92] MIRJALILI, S., *Evolutionary Algorithms and Neural Networks - Theory and Applications*. v. 1. Springer: Australia, 2018.
- [93] MANDISCHER, M., “A comparison of evolution strategies and backpropagation for neural network training”, *Neurocomputing*, v. 42, n. 1, pp. 87–117, 2002, Evolutionary neural systems.

- [94] FERNANDES JR., F. E., YEN, G. G., “Pruning Deep Convolutional Neural Networks Architectures with Evolution Strategy”, *Information Sciences*, v. 552, pp. 29–47, 2021.
- [95] CHEN, Z., HUANG, A., QIANG, X., “Improved neural networks based on genetic algorithm for pulse recognition”, *Computational Biology and Chemistry*, v. 88, pp. 107315, 2020.
- [96] MUROVIĆ, T., TROST, A., “Genetically optimized massively parallel binary neural networks for intrusion detection systems”, *Computer Communications*, v. 179, pp. 1–10, 2021.
- [97] PANT, P., CHATTERJEE, D., “Prediction of clad characteristics using ANN and combined PSO-ANN algorithms in laser metal deposition process”, *Surfaces and Interfaces*, v. 21, pp. 100699, 2020.
- [98] EBRAHIMI, M., AZIMI, E., NASIRI SARVI, M., et al., “Hybrid PSO enhanced ANN model and central composite design for modelling and optimization of Low-Intensity magnetic separation of hematite”, *Minerals Engineering*, v. 170, pp. 106987, 2021.
- [99] DARWIN, C., *On the Origin of Species by Means of Natural Selection, or the Preservation of Favoured Races in the Struggle for Life*. London, 1859.
- [100] KENNEDY, J., EBERHART, R., “Particle Swarm Optimization.” *Proceedings of the IEEE International Conference on Neural Networks.*, pp. 1942–1945, 1995.
- [101] DONGSHU, W., DAPEI, T., LEI, L., “Particle Swarm Optimization algorithm: an overview”, *Soft Computing*, v. 22-2, pp. 387–408, 2018.
- [102] LIU, K., ZHANG, J., “Nonlinear process modelling using echo state networks optimised by covariance matrix adaption evolutionary strategy”, *Computers and Chemical Engineering*, v. 135, pp. 106730, 2020.
- [103] LIANG, Y., WANG, X., ZHAO, H., et al., “A covariance matrix adaptation evolution strategy variant and its engineering application”, *Applied Soft Computing*, v. 83, pp. 105680, 2019.
- [104] KHAN, N., “A parallel implementation of the covariance matrix adaptation evolution strategy”, *CoRR*, v. abs/1805.11201, 2018.
- [105] WHITESON, S., WHITESON, D., “Machine learning for event selection in high energy physics”, *Engineering Applications of Artificial Intelligence*, v. 22, n. 8, pp. 1203–1217, 2009.

- [106] VÁZQUEZ-ESCOBAR, J., HERNÁNDEZ, J., CÁRDENAS-MONTES, M., “Estimation of Machine Learning model uncertainty in particle physics event classifiers”, *Computer Physics Communications*, v. 268, pp. 108100, 2021.
- [107] BETHAPUDI, S., DESAI, S., “Separation of pulsar signals from noise using supervised machine learning algorithms”, *Astronomy and Computing*, v. 23, pp. 15–26, 2018.
- [108] FABIAN, X., BAULIEU, G., DUCROUX, L., et al., “Artificial neural networks for neutron/ γ discrimination in the neutron detectors of NEDA”, *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, v. 986, pp. 164750, 2021.
- [109] SPENCER, S., ARMSTRONG, T., WATSON, J., et al., “Deep learning with photosensor timing information as a background rejection method for the Cherenkov Telescope Array”, *Astroparticle Physics*, v. 129, pp. 102579, 2021.
- [110] GUILLÉN, A., BUENO, A., CARCELLER, J., et al., “Deep learning techniques applied to the physics of extensive air showers”, *Astroparticle Physics*, v. 111, pp. 12–22, 2019.
- [111] AKCHURIN, N., BEDESCHI, F., CARDINI, A., et al., “Particle identification in the longitudinally unsegmented RD52 calorimeter”, *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, v. 735, pp. 120–129, 2014.
- [112] AKCHURIN, N., CARRELL, K., HAUPTMAN, J., et al., “Hadron and jet detection with a dual-readout calorimeter”, *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, v. 537, n. 3, pp. 537–561, 2005.
- [113] NITTA, K., ALIU, E., ANDRINGA, S., et al., “The K2K SciBar detector”, *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, v. 535, n. 1, pp. 147–151, 2004, Proceedings of the 10th International Vienna Conference on Instrumentation.
- [114] NAGAI, Y., MATSUBARA, Y., ITOW, Y., et al., “First cosmic-ray measurements by the SciCRT solar neutron experiment in Mexico”, *Astroparticle Physics*, v. 59, pp. 39–46, 2014.

- [115] LAURITS, T., VEELKEN, C., KADASTIK, M., “Evolutionary algorithms for hyperparameter optimization in machine learning for application in high energy physics”, *The European Physical Journal C*, v. 81, pp. 1–9, 2019.
- [116] HMIDA, H., HAMIDA, S. B., BORGI, A., et al., “Scale Genetic Programming for large Data Sets: Case of Higgs Bosons Classification”, *Procedia Computer Science*, v. 126, pp. 302–311, 2018, Knowledge-Based and Intelligent Information and Engineering Systems: Proceedings of the 22nd International Conference, KES-2018, Belgrade, Serbia.
- [117] ADAM-BOURDARIOS, C., COWAN, G., GERMAIN, C., et al., “The Higgs boson machine learning challenge”. In: *Proceedings of the NIPS 2014 Workshop on High-energy Physics and Machine Learning*, v. 42, *Proceedings of Machine Learning Research*, pp. 19–55, Montreal, Canada, 13 Dec 2015.
- [118] PIERRE AUGER COLLABORATION, T., “The Pierre Auger Cosmic Ray Observatory”, *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, v. 798, pp. 172–213, 2015.
- [119] CAMERON, A. D., CHAMPION, D. J., BAILES, M., et al., “The High Time Resolution Universe Pulsar Survey – XVI. Discovery and timing of 40 pulsars from the southern Galactic plane”, *Monthly Notices of the Royal Astronomical Society*, v. 493, n. 1, pp. 1063–1087, 01 2020.
- [120] KUNCHEVA, L. I., *Combining Pattern Classifiers: Methods and Algorithms*. v. 2. John Wiley and Sons, 2014.
- [121] NGUYEN, D., WIDROW, B., “Improving the learning speed of 2-layer neural networks by choosing initial values of the adaptive weights”, *1990 IJCNN International Joint Conference on Neural Networks*, v. 24, pp. 21–26 vol.3, 1990.
- [122] KINGMA, D. P., BA, J., “Adam: A Method for Stochastic Optimization”, *CoRR*, v. abs/1412.6980, 2015.
- [123] SPECKMAYER, P., HÖCKER, A., STELZER, J., et al., “The toolkit for multivariate data analysis, TMVA 4”, *Journal of Physics: Conference Series*, v. 219, n. 3, pp. 032057, apr 2010.
- [124] KLIMEK, P., *Signal reconstruction performance with the ATLAS Hadronic Tile Calorimeter*, Tech. Rep. ATL-TILECAL-PROC-2012-007, CERN, Geneva, Aug 2012.

- [125] ATLAS-COLLABORATION, “Athena Framework”, *Zenodo*, April 2019.
- [126] BARRAND, G., BELYAEV, I., BINKO, P., et al., “GAUDI — A software architecture and framework for building HEP data processing applications”, *Computer Physics Communications*, v. 140, n. 1, pp. 45–55, 2001, CHEP2000.
- [127] DE SOUZA, E. E. P., *Pré-processamento Eficiente para Um Sistema Online de Classificação Baseado em Redes Neurais Artificiais*, Master’s Thesis, Universidade Federal da Bahia/PPGEE/UFBA, Salvador/BA, 2015.
- [128] DE SOUZA, E. E. P., DE SIMAS FILHO, E. F., FARIAS, P. C. M. A., “Método De Treinamento De Redes Neurais Através De Otimização Evolutiva Com Restrições Para Auxílio Ao Diagnóstico Médico”. In: *XXXVI Simpósio Brasileiro de Telecomunicações e Processamento de Sinais (SBrT2018)*, pp. 1–5, Anais do SBrT: Campina Grande/PB, 2018.
- [129] SILVA, G., SOUZA, E., FARIAS, P., et al., “A Constrained Neural Classifier for Pulsed Eddy Current based Flaw Detection in Industrial Pipes”. In: *Anais do 14 Congresso Brasileiro de Inteligência Computacional*, pp. 1–6, ABRICOM: Belém/PA, 2019.
- [130] ATLAS-COLLABORATION, “Electron and photon performance measurements with the ATLAS detector using the 2015–2017 LHC proton-proton collision data”, *Journal of Instrumentation*, v. 14, n. 12, pp. P12006–P12006, dec 2019.

Apêndice A

Produção Científica

Neste apêndice são apresentadas as publicações realizadas durante o desenvolvimento desta tese.

A.1 Publicações em Periódicos

1. ATLAS COLLABORATION. **Evidence for Higgs boson decays to a low-mass dilepton system and a photon in pp collisions at TeV with the ATLAS detector.** *Physics Letters B*, vol. 819, 2021.

Abstract: *A search for the Higgs boson decaying into a photon and a pair of electrons or muons with an invariant mass $m_{ll} < 30$ GeV is presented. The analysis is performed using 139 fb^{-1} of proton-proton collision data, produced by the LHC at a centre-of-mass energy of 13 TeV and collected by the ATLAS experiment. Evidence for the $H \rightarrow ll\gamma$ process is found with a significance of 3.2 over the background-only hypothesis, compared to an expected significance of 2.1 for the Standard Model prediction. The best-fit value of the signal-strength parameter, defined as the ratio of the observed signal yield to the one expected in the Standard Model, is $\mu = 1.5 \pm 0.5$. The Higgs boson production cross-section times the $H \rightarrow ll\gamma$ branching ratio for $m_{ll} < 30$ GeV is determined to be $8.7 - 2.7 + 2.8 \text{ fb}$.*

2. ATLAS COLLABORATION. **Measurements of $W^+W^- + \geq 1$ jet production cross-sections in pp collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector.** *Journal of High Energy Physics*, 2021, 3 (2021).

Abstract: *Fiducial and differential cross-section measurements of W^+W^- production in association with at least one hadronic jet are presented. These measurements are sensitive to the properties of electroweak boson self interactions and provide a test of perturbative quantum chromodynamics and the electroweak theory. The analysis is performed using proton-proton collision data collected at $\sqrt{s} = 13$ TeV with the ATLAS experiment, corresponding to an integrated luminosity of 139 fb^{-1} . Events are selected with exactly one oppositely charged electron-muon pair and at least one hadronic jet with a transverse momentum of $p_T > 30$ GeV and a pseudorapidity of $|\eta| < 4.5$. After subtracting the background contributions and correcting for detector effects, the jet-inclusive $W^+W^- + \geq 1$ jet fiducial cross-section and $W^+W^- + \text{jets}$ differential cross-sections with respect to several kinematic variables are measured. These measurements include leptonic quantities, such as the lepton*

transverse momenta and the transverse mass of the W^+W^- system, as well as jet-related observables such as the leading jet transverse momentum and the jet multiplicity. Limits on anomalous triple gauge-boson couplings are obtained in a phase space where interference between the Standard Model amplitude and the anomalous amplitude is enhanced.

3. ATLAS COLLABORATION. **A search for the decays of stopped long-lived particles at $\sqrt{s} = 13$ TeV with the ATLAS detector.** *Journal of High Energy Physics*, 2021, 173 (2021).

Abstract: A search for long-lived particles, which have come to rest within the ATLAS detector, is presented. The subsequent decays of these long-lived particles can produce high-momentum jets, resulting in large out-of-time energy deposits in the ATLAS calorimeters. These decays are detected using data collected during periods in the LHC bunch structure when collisions are absent. The analysed dataset is composed of events from proton-proton collisions produced by the Large Hadron Collider at a centre-of-mass energy of $\sqrt{s} = 13$ TeV and recorded by the ATLAS experiment during 2017 and 2018. The dataset used for this search corresponds to a total live time of 579 hours. The results of this search are used to derive lower limits on the mass of gluino R -hadrons, assuming a branching fraction $B(\tilde{g} \rightarrow q\bar{q}\tilde{\chi}_1^0) = 100\%$, with masses of up to 1.4 TeV excluded for gluino lifetimes of 10^{-5} to 10^3 s.

4. ATLAS COLLABORATION. **Measurements of sensor radiation damage in the ATLAS inner detector using leakage currents.** *Journal of Instrumentation*, vol. 6, 2021.

Abstract: Non-ionizing energy loss causes bulk damage to the silicon sensors of the ATLAS pixel and strip detectors. This damage has important implications for data-taking operations, charged-particle track reconstruction, detector simulations, and physics analysis. This paper presents simulations and measurements of the leakage current in the ATLAS pixel detector and semiconductor tracker as a function of location in the detector and time, using data collected in Run 1 (2010–2012) and Run 2 (2015–2018) of the Large Hadron Collider. The extracted fluence shows a much stronger $|z|$ -dependence in the innermost layers than is seen in simulation. Furthermore, the overall fluence on the second innermost layer is significantly higher than in simulation, with better agreement in layers at higher radii. These measurements are important for validating the simulation models and can be used in part to justify safety factors for future detector designs and interventions.

5. ATLAS COLLABORATION. **Two-particle azimuthal correlations in photonuclear ultraperipheral collisions at 5.02 TeV with ATLAS.** *Physical Review C*, vol. 104, 2021.

Abstract: Two-particle long-range azimuthal correlations are measured in photonuclear collisions using 1.7nb^{-1} of 5.02 TeV Pb+Pb collision data collected by the ATLAS experiment at the CERN Large Hadron Collider. Candidate events are selected using a dedicated high-multiplicity photonuclear event trigger, a combination of information from the zero-degree calorimeters and forward calorimeters, and from pseudorapidity gaps constructed using calorimeter energy clusters and charged-particle tracks. Distributions of event properties are compared between data and Monte Carlo simulations of photonuclear processes. Two-particle correlation functions are formed using charged-particle tracks in the selected events, and a template-fitting method is employed to subtract the nonflow contribution to the correlation.

Significant nonzero values of the second- and third-order flow coefficients are observed and presented as a function of charged-particle multiplicity and transverse momentum. The results are compared with flow coefficients obtained in proton-proton and proton-lead collisions in similar multiplicity ranges, and with theoretical expectations. The unique initial conditions present in this measurement provide a new way to probe the origin of the collective signatures previously observed only in hadronic collisions.

6. ATLAS COLLABORATION. **Search for new phenomena with top quark pairs in final states with one lepton, jets, and missing transverse momentum in pp collisions at $\sqrt{s}=13$ TeV with the ATLAS detector.** *Journal of High Energy Physics*, vol. 174, 2021.

Abstract: A search for new phenomena with top quark pairs in final states with one isolated electron or muon, multiple jets, and large missing transverse momentum is performed. Signal regions are designed to search for two, three, and four-body decays of the directly pair-produced supersymmetric partner of the top quark (*stop*). Additional signal regions are designed specifically to search for spin-0 mediators that are produced in association with a pair of top quarks and decay into a pair of dark-matter particles. The search is performed using the Large Hadron Collider proton-proton collision dataset at a centre-of-mass energy of $\sqrt{s} = 13$ TeV recorded by the ATLAS detector from 2015 to 2018, corresponding to an integrated luminosity of 139 fb^{-1} . No significant excess above the Standard Model background is observed, and limits at 95% confidence level are set in the *stop*-neutralino mass plane and as a function of the mediator mass or the dark-matter particle mass. Stops are excluded up to 1200 GeV (710 GeV) in the two-body (three-body) decay scenario. In the four-body scenario stops up to 640 GeV are excluded for a *stop*-neutralino mass difference of 60 GeV. Scalar and pseudoscalar dark-matter mediators are excluded up to 200 GeV when the coupling strengths of the mediator to Standard Model and dark-matter particles are both equal to one and when the mass of the dark-matter particle is 1 GeV.

A.2 Artigos Aceitos em Periódicos

1. SILVA, G., SOUZA, E., SIMAS, E., FARIAS P., ALBUQUERQUE M., SILVA I., FARIAS C. **Constrained Neural Classifier Training Method for Flaw Detection in Industrial Pipes Using Particle Swarm Optimization.** Submetido para: *International Journal of Innovative Computing and Applications*, 2020.

Abstract: A novel method for constrained training of multi-class artificial neural network classifiers is proposed in this work. The traditional training procedure is usually based on mean square error minimization and thus, all classes of interest are considered as having the same relevance for system performance. This is not always the case for real-world applications in which the class relevance may be unbalanced. In this paper, cost functions designed to introduce classification performance constraints for specific classes are presented and particle swarm optimization is used as global optimization method. The proposed method is applied to a non-destructive evaluation decision support problem using pulsed eddy currents signals. Experimental results obtained from thermally insulated industrial pipes indicate the efficiency of the proposed method in comparison to neural networks trained from the traditional back-propagation algorithm.

2. ATLAS COLLABORATION. **Search for new phenomena in pp collisions in final states with tau leptons, b-jets, and missing transverse momentum with the ATLAS detector.** Submetido para: Physics Review D, 2021.

Abstract: *A search for new phenomena in final states with hadronically decaying tau leptons, b-jets, and missing transverse momentum is presented. The analyzed dataset comprises pp collision data at a center-of-mass energy of $\sqrt{s} = 13$ TeV with an integrated luminosity of 139/fb, delivered by the Large Hadron Collider and recorded with the ATLAS detector from 2015 to 2018. The observed data are compatible with the expected Standard Model background. The results are interpreted in simplified models for two different scenarios. The first model is based on supersymmetry and considers pair production of top squarks, each of which decays into a b-quark, a neutrino and a tau slepton. Each tau slepton in turn decays into a tau lepton and a nearly massless gravitino. Within this model, top-squark masses up to 1.4 TeV can be excluded at the 95% confidence level over a wide range of tau-slepton masses. The second model considers pair production of leptoquarks with decays into third-generation leptons and quarks. Depending on the branching fraction into charged leptons, leptoquarks with masses up to around 1.25 TeV can be excluded at the 95% confidence level for the case of scalar leptoquarks and up to 1.8 TeV (1.5 TeV) for vector leptoquarks in a Yang–Mills (minimal-coupling) scenario. In addition, model-independent upper limits are set on the cross section of processes beyond the Standard Model.*

3. ATLAS COLLABORATION. **Search for charginos and neutralinos in final states with two boosted hadronically decaying bosons and missing transverse momentum in pp collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector.** Submetido para: Physics Review D, 2021.

Abstract: *A search for charginos and neutralinos at the Large Hadron Collider is reported using fully hadronic final states and missing transverse momentum. Pair-produced charginos or neutralinos are explored, each decaying into a high-pT Standard Model weak boson. Fully-hadronic final states are studied to exploit the advantage of the large branching ratio, and the efficient background rejection by identifying the high-pT bosons using large-radius jets and jet substructure information. An integrated luminosity of 139 fb⁻¹ of proton-proton collision data collected by the ATLAS detector at a center-of-mass energy of 13 TeV is used. No significant excess is found beyond the Standard Model expectation. The 95% confidence level exclusion limits are set on wino or higgsino production with varying assumptions in the decay branching ratios and the type of the lightest supersymmetric particle. A wino (higgsino) mass up to 1060 (900) GeV is excluded when the lightest SUSY particle mass is below 400 (240) GeV and the mass splitting is larger than 400 (450) GeV. The sensitivity to high-mass wino and higgsino is significantly extended compared with the previous LHC searches using the other final states.*

4. ATLAS COLLABORATION. **Search for new phenomena in final states with two leptons and one or no b-tagged jets at $\sqrt{s} = 13$ TeV using the ATLAS detector.** Submetido para: Physics Review Letters, 2021.

Abstract: *A search for new phenomena is presented in final states with two leptons and one or no b-tagged jets. The event selection requires the two leptons to have opposite charge, the same flavor (electrons or muons), and a large invariant mass. The analysis is based on the full Run-2 proton-proton collision dataset recorded at a center-of-mass energy of $\sqrt{s} = 13$*

TeV by the ATLAS experiment at the LHC, corresponding to an integrated luminosity of 139 fb^{-1} . No significant deviation from the expected background is observed in the data. A four-fermion contact interaction between two quarks (b,s) and two leptons (ee or $\mu\mu$), inspired by the B-meson decay anomalies, is used as a benchmark signal model. This model is characterized by the energy scale and coupling, Λ and g_* respectively. Contact interactions with Λ/g_* lower than 2.0 (2.4) TeV are excluded for electrons (muons) at the 95% confidence level, still far below the value which is favored by the B-meson decay anomalies. In addition, model-independent limits are set as a function of the selection on the dilepton invariant mass, which allows the results to be reinterpreted in other signal scenarios.

A.3 Artigos em Conferências

1. SOUZA, E., SIMAS E., FARIAS P., MARIN J., PINTO J., ARAÚJO M., SEIXAS J. **Informação Especialista para um Sistema de Filtragem *Online* com *Ensemble* de Redes Neurais em um Ambiente com Alta Taxa de Eventos**. Congresso Brasileiro de Inteligência Computacional, 2021, Joinville/SC.

Resumo: O experimento ATLAS é projetado para observar os constituintes fundamentais em colisões de partículas no Grande Colisor de Hádrons (LHC). Na tomada de dados a partir de 2021, espera-se aumentar a energia de colisão e a taxa de produção de informações, o que requer melhorias no sistema de filtragem *online*. O NeuralRinger é o algoritmo padrão de identificação *online* de elétrons no ATLAS (etapa rápida) e usa informações do calorímetro (medição de energia) organizadas em anéis concêntricos para alimentar um conjunto de redes neurais. Uma das principais características utilizadas para a discriminação de elétrons e jatos hadrônicos é o perfil lateral do chuveiro de partículas de energia decrescente medido no calorímetro, pois o chuveiro desenvolvido por partículas eletromagnéticas tende a ser menos largo lateralmente se comparado a chuveiros hadrônicos. Neste trabalho, é investigada a utilização do perfil lateral do chuveiro combinado com os anéis. O método proposto pondera os anéis gerados pelo NeuralRinger pela energia e distância até a célula mais energética. Os resultados indicaram uma redução do falso alarme de 1,37 pontos percentuais quando comparado ao desempenho do NeuralRinger em dados experimentais o que é relevante, haja vista a alta taxa de eventos gerados no ATLAS.

2. MARIN J., SIMAS E., FARIAS P., SOUZA, E., PINTO J., ARAÚJO M., SEIXAS J. **Conjunto de Redes Neurais Artificiais para Detecção *Online* de Fótons de Altas Energias**. Congresso Brasileiro de Inteligência Computacional, 2021, Joinville/SC.

Resumo: A correta identificação de partículas é um dos principais objetivos de experimentos de física de altas energias. Devido a alta taxa de eventos no Grande Colisor de Hádrons (LHC), o experimento ATLAS tem empregado técnicas baseadas em aprendizado de máquina a fim de encontrar eventos raros em grandes massas de dados. Entre eles está o NeuralRinger, um conjunto classificador de rede neural projetado para detecção rápida de elétrons com base em anéis concêntricos de calorímetro (sistema de medição de energia). Nesse contexto, o presente trabalho propõe a adaptação desta técnica, que opera identificando elétrons, para a detecção rápida de fótons. Os resultados com dados simulados mostram a eficiência do método proposto no ambiente do experimento ATLAS.

3. SIMAS E., FARIAS P., SOUZA, E., MARIN J., PINTO J., FREUND W., SEIXAS J.,

LAFORGE B. **Uma Proposta para Calibração do Sistema *Online* de Seleção de Eventos no Detector ATLAS utilizando Gradient Boosted Decision Trees**. Congresso Brasileiro de Inteligência Computacional, 2021, Joinville/SC.

Resumo: Experimentos de física de partículas lidam com um enorme volume de informação e uma complexa cadeia sequencial de processamento para seleção *online* (*trigger*) de eventos. No detector ATLAS do LHC (*Large Hadron Collider*), o sistema de *trigger* é responsável por escolher os eventos que serão gravados em mídia permanente para futura análise e opera de modo sequencial em dois níveis de seleção. O valor estimado da energia depositada pelas partículas no detector é um importante parâmetro para o processo de seleção *online*. Neste trabalho é proposto um método de calibração baseado em *gradient boosted decision trees ensemble* (GBDTE) para melhorar a qualidade da energia estimada na segunda etapa de *trigger* do detector ATLAS. Com o método proposto é possível, ao mesmo tempo, reduzir requisitos computacionais e aumentar a eficiência de seleção de partículas eletromagnéticas (elétrons e fótons).

4. PEDREIRA L., SOUZA, E., SIMAS E., FARIAS P. **Método de Treinamento de Classificadores Neurais com Restrições para Redução do Falso Alarme em Experimentos com Alta Taxa de Eventos**. X Conferência Nacional em Comunicações, Redes e Segurança da Informação, 2020, Natal/RN.

Resumo: Experimentos de Física de Altas Energias são caracterizados pela elevada produção de informação durante a operação e pela rara ocorrência de eventos de interesse, o que requer o uso de sistemas de classificação *online*. Neste contexto, os pontos de operação dos sistemas de classificação de eventos são ajustados para elevada taxa de detecção de sinal de modo a preservar a informação de interesse, entretanto, isso pode causar aumento na taxa de falso alarme. Este trabalho apresenta um método de treinamento de classificadores neurais, que permite a inclusão de restrições associadas diretamente a taxas de probabilidade de detecção, de modo a minimizar a taxa de falso alarme para o ponto de operação requerido. Os resultados preliminares com conjunto de dados do telescópio de Cherenkov MAGIC indicam redução da taxa de falso alarme de ≈ 3 pontos percentuais, considerando 99% de probabilidade de detecção, se comparado a distintas versões do algoritmo *backpropagation*.

5. SILVA, G., SOUZA, E., SIMAS, E., FARIAS P., ALBUQUERQUE M., SILVA I., FARIAS C. **A Constrained Neural Classifier for Pulsed Eddy Current based Flaw Detection in Industrial Pipes**. Congresso Brasileiro de Inteligência Computacional, 2019, Belém/PA.

Abstract: Decision support systems are important to improve the efficiency of nondestructive evaluation, specially for industrial equipment. Pulsed eddy-current is a magnetic method used for evaluation of metallic equipment. In this paper, is proposed the combination of pulsed eddy current evaluation, digital signal processing, and neural networks to detect flaws in industrial pipes. A novel method using particle swarm optimization is proposed for imposing performance constraints during neural classifier training process. Results obtained for experimental signals acquired from composite-insulated metallic industrial pipes presenting internal and external corrosion areas are used to validate the proposed method. A comparison to neural networks trained from the traditional back-propagation algorithm was presented.

6. CARVALHO J., SOUZA, E., SIMAS, E., FARIAS P. **Particle Swarm Localization for Mobile Robots Using a 2D Laser Sensor**. 8th Brazilian Conference on Intelligent Systems (BRACIS), 2019, Salvador/BA.

Abstract: Mobile robot localization is a complex task, specially in unstructured indoor environments, due to noise and wrong data association from sensors. The localization procedure is even harder when the vehicle has low confidence about its last pose estimate, situation that requires a Global Localization procedure. In this work, a Global Localization algorithm based on Particle Swarm Optimization (PSO) is integrated with a Pose Tracking algorithm, the Perfect Match, to obtain a robust localization technique. Results show that this approach can solve the problem of Global Localization with good performance.

7. SOUZA, E., SIMAS, E., FARIAS P. **Método de Treinamento de Redes Neurais através de Otimização Evolutiva com Restrições para Auxílio ao Diagnóstico Médico**. Simpósio Brasileiro de Telecomunicações e Processamento de Sinais (SBrT), 2018, Campina Grande/PB.

Resumo: Neste trabalho é proposto um método de treinamento de Redes Neurais Artificiais para problemas de classificação binária. A metodologia apresentada utiliza algoritmos de otimização evolutiva que ajustam iterativamente os pesos da rede, permitindo a imposição de restrições que guiam o treinamento para satisfação de critérios em função de medidas de desempenho de classificação. O método proposto apresentou aumento de eficiência de classificação em problemas binários de detecção de doenças, considerando restrições para taxas de falsos positivos. Os resultados encontrados foram baseados em conjuntos de dados experimentais que representam problemas de diagnóstico médico.

8. FONSECA M., SOUZA, E., SIMAS, E., SCOLARI A. **Classificador de Gestos por Redes Neurais Artificiais utilizando Sinais Mioelétricos**. Simpósio Brasileiro de Telecomunicações e Processamento de Sinais (SBrT), 2018, Campina Grande/PB.

Resumo: Neste trabalho é apresentado um sistema de classificação, baseado em Redes Neurais Artificiais, para discriminação de gestos de mão utilizando sinais eletromiográficos (EMG), que são obtidos através de uma braçadeira *MyoTM* composta por oito eletrodos que realizam a aquisição da informação da contração muscular durante a execução de movimentos padrões. É realizado um pré-processamento nos sinais adquiridos, extraindo sete características para cada eletrodo. Visando analisar o melhor desempenho do classificador, dois algoritmos de treinamento foram utilizados, o Resilient Backpropagation (Rprop) e Levenberg Maguard (LM), onde o algoritmo Rprop obteve cerca de 91,3% de acerto de classificação entre as quatro classes de movimentos estudados, em relação ao algoritmo LM. Os resultados experimentais são apresentados para validar o método proposto.

9. SOUZA, E., SIMAS, E., FARIAS P. **Método de Treinamento com Restrições para Classificadores Neurais Utilizando Otimização Evolutiva**. VIII Conferência Nacional em Comunicações, Redes e Segurança da Informação, 2018, Salvador/BA.

Resumo: Neste trabalho é proposto um método de treinamento de Redes Neurais Artificiais (RNA) para problemas de classificação binária. A metodologia apresentada utiliza algoritmos de otimização evolutiva que ajustam iterativamente os pesos da RNA, permitindo a imposição de restrições que guiam o treinamento para satisfação de critérios em função de medidas de desempenho de classificação. O método proposto foi aplicado a problemas de

detecção de eventos em experimentos de física de partículas, onde a imposição de restrições é muito importante para o desempenho do sistema de classificação.

10. SOUZA, E., SIMAS, E., FARIAS P., FREUND W., SEIXAS J. **Máquinas de Aprendizado Extremo para Classificação *Online* de Eventos no detector ATLAS**. Simpósio Brasileiro de Telecomunicações e Processamento de Sinais (SBrT), 2017, São Pedro/SP.

Resumo: O ATLAS é um dos detectores do LHC (*Large Hadron Collider*), e está localizado no CERN (Organização Europeia para a pesquisa Nuclear). Para adequada caracterização das partículas é preciso realizar uma precisa medição do perfil de deposição de energia à medida que ocorrem interações com o detector. No ATLAS os calorímetros são responsáveis por realizar a estimativa da energia das partículas e, neste sentido, utilizam mais de 100.000 sensores. Um dos discriminadores para a detecção *online* de elétrons utilizados no ATLAS é o Neural Ringer, no qual o perfil de deposição de energia é utilizado como entrada para um classificador neural tipo *perceptron* de múltiplas camadas. Este trabalho propõe o uso de Máquinas de Aprendizado Extremo (ELM) em substituição às redes do tipo *perceptron multilayer* no Neural Ringer. Os resultados obtidos de uma base de dados simulados apontam para uma significativa redução do tempo de treinamento, com desempenho de classificação semelhante.

A.4 Resumos em Conferências

1. SIMAS E., FARIAS P., SOUZA, E., MARIN J., PINTO J., FREUND W., SEIXAS J., LA-FORGE B. **Calibração de Energia para o *Trigger* de Elétrons no ATLAS usando Anéis Concêntricos de Informação de Calorimetria**. Workshop da RENAFAE: Projetos para o Futuro da Física de Altas Energias no Brasil, 2021, São Paulo/SP.
2. MARIN J., SIMAS E., FARIAS P., SOUZA, E., PINTO J., ARAÚJO M., FREUND W., SEIXAS J. **Framework Integrado para Análise de Dados e Treinamento de Modelos de Aprendizagem de Máquina que operam no *Trigger Online* do Experimento ATLAS**. Workshop da RENAFAE: Projetos para o Futuro da Física de Altas Energias no Brasil, 2021, São Paulo/SP.
3. FILHO L., ARAÚJO M., CERQUEIRA A., FERRAZ V., FERREIRA I., GONÇALVES D., MARIN J., PERALVA B., PINTO J., SEIXAS J., SILVA A., SIMAS E., SOUZA E., VIANNA R. **Contribuições para os Sistemas de *Trigger* do ATLAS em Operação no *Run3***. Workshop da RENAFAE: Projetos para o Futuro da Física de Altas Energias no Brasil, 2021, São Paulo/SP.
4. SOUZA, E., SIMAS E., FARIAS P., MARIN J., PINTO J., ARAÚJO M., SEIXAS J. **Sistemas de Monitoramento para o *Trigger* de Elétrons e Fótons no Experimento ATLAS/LHC**. Workshop da RENAFAE: Projetos para o Futuro da Física de Altas Energias no Brasil, 2021, São Paulo/SP.
5. SOUZA, E., SIMAS E., FARIAS P., MARIN J., PINTO J., ARAÚJO M., SEIXAS J. **Detecção *Online* de Elétrons em um Calorímetro Finamente Segmentado Usando**

Informação Especialista de Calorimetria e um Ensemble de Redes Neurais Artificiais. XLI Encontro Nacional de Física de Partículas e Campos (ENFPC), 2021, Campos do Jordão/SP.

6. SOUZA, E., SIMAS E., FARIAS P., MARIN J., PINTO J., ARAÚJO M., SEIXAS J. **Monitoração do Sistema de *Trigger* para Elétrons e Fótons no Experimento ATLAS.** XLI Encontro Nacional de Física de Partículas e Campos (ENFPC), 2021, Campos do Jordão/SP.
7. SIMAS E., FARIAS P., SOUZA, E., MARIN J., PINTO J., FREUND W., SEIXAS J., LAFORGE B. **Estudo do uso de medidas de assimetria de deposição de energia no processo de calibração no *Fast Step* do *Trigger* do experimento ATLAS.** XLI Encontro Nacional de Física de Partículas e Campos (ENFPC), 2021, Campos do Jordão/SP.
8. SIMAS E., PERALVA B., SOUZA, E., MARIN J., PINTO J., FREUND W., SEIXAS J., LAFORGE B. **Conjunto de Redes Neurais para Identificação de Fótons no Experimento ATLAS.** XLI Encontro Nacional de Física de Partículas e Campos (ENFPC), 2021, Campos do Jordão/SP.
9. SIMAS E., FARIAS P., SOUZA, E., MARIN J., PINTO J., FREUND W., SEIXAS J., LAFORGE B. **Calibração para energia de elétrons na etapa rápida do *trigger* do ATLAS usando anéis concêntricos de informação de calorimetria.** XLI Encontro Nacional de Física de Partículas e Campos (ENFPC), 2021, Campos do Jordão/SP.
10. SOUZA, E., SIMAS E., FARIAS P., PEDREIRA L., FREUND W., SEIXAS J. **Usando Informação Especialista para uma Filtragem *Online* com um Ensemble de Redes Neurais em um Ambiente Finamente Segmentado e com Alta Taxa de Eventos.** Congresso Brasileiro de Inteligência Computacional, 2019, Belém/PA.
11. SANTOS M., SOUZA E., SIMAS E., FARIAS P., FILHO L., SEIXAS J. **Uso de Algoritmos de Treinamento Rápido para o Discriminador Neural Ringer no Detector ATLAS.** XXXIX Encontro Nacional de Física de Partículas e Campos, 2018, Campos do Jordão/SP.

Apêndice B

Qualificação para Afiliação no Experimento ATLAS

B.1 Descrição da Tarefa de Qualificação

Para se tornar membro da colaboração ATLAS e obter o direito de ser autor de suas publicações, os candidatos precisam realizar uma tarefa com importante necessidade para o experimento. A tarefa de qualificação realizada pelo autor desta tese, foi orientada por dois supervisores, sendo o supervisor no grupo brasileiro o Prof. Dr. José Manoel de Seixas (Universidade Federal do Rio de Janeiro/COPPE, RJ/Brasil), e o supervisor indicado pelo ATLAS, o Prof. Dr. Fernando Gabriel Monticelli (Universidade Nacional de La Plata/CONICET, La Plata, Argentina).

O processo de qualificação foi realizado em um período de um ano, tendo início em 01/11/2019, com a seguinte descrição:

Project Description:

To port online monitoring into egammaBuilder and superClusterBuilder algorithms. Currently the online monitoring histograms are published by hypothesis algorithms. Monitoring by the reconstruction algorithms are also needed in order to validate and commission their performance running online. Re-optimisation of the Run 3 ringer tunes $>15\text{GeV}$ with new L1Calo/precision/offline selections, potentially adding new calorimeter shower shape variables to the ringer structure (for example the shower width). Use monitoring tools to validate Run 3 ringer optimisation. When the task is complete, the online monitoring of egammaBuilder and superClusterBuilder will be present in output histogram files. All the progress will be discussed in the Trigger Egamma meetings, and documented as slides in presentations.

O período de qualificação se encerrou em 06/11/2020, quando obteve-se aprovação, constando o seguinte relatório final do supervisor técnico:

Final Report Comments:

Edmar successfully implemented the online monitoring on offline algorithms instantiated at HLT. After trying modifying offline algorithms it was suggested (and decided) to write stand alone monitoring algorithms that pick up the output of the offline algorithms and publish the histograms. Edmar wrote, developed and implemented such algorithms and are currently running in master branch of athenaMT for electrons and photons at precision step (namely topoEgammabuilder). He is currently working on implementing such algorithms for intermediate steps in the e/y reconstruction chain. When we defined this task we supposed would have available Run-3 samples to perform ringer optimization. Those are not available, but still the critical part of the task is complete and Edmar agreed to finish polishing what is pending of the online monitoring task, like tuning ranges and bins of histograms.

Adicionalmente à tarefa de qualificação, o autor desta tese realizou atividades técnicas e operacionais para avaliação de qualidade de dados (DQ - *Data Quality*) dos reprocessamentos de dados experimentais. O autor foi responsável por acompanhar as discussões junto aos demais grupos de trabalho do ATLAS, relacionados com ferramentas de monitoração e DQ, sendo o *DQ Contact Person* (contato principal para monitoração e avaliação de qualidade de dados) do grupo de trabalho *TrigEgamma Signature*. Algumas das principais implementações realizadas durante o período de qualificação, são descritas brevemente a seguir.

B.2 Monitoração *Online* nos Algoritmos de Reconstrução

Conforme discutido no escopo desta tese, as características do formato do chuveiro de partículas (grandezas de calorimetria), e da deposição de energia no sistema de calorimetria, carregam informações que com base em modelos de física teórica, é possível utilizá-las para discriminação elétron/jatos. Tais variáveis, também são utilizadas em tarefas de calibração da instrumentação do ATLAS, bem como avaliação de qualidade dos dados obtidos durante as colisões. Neste contexto, algoritmos de monitoração são empregados para prover histogramas das grandezas de calorimetria ao *framework* de exibição gráfica, utilizado pela equipe de turno na sala de controle do ATLAS.

A monitoração *online* é realizada no HLT, em algoritmos de classificação de eventos e algoritmos de reconstrução. Nos algoritmos de classificação, além das grandezas de calorimetria, são monitoradas variáveis que revelam o tempo de execução dos algoritmos, e as variáveis de decisão dos testes de hipótese aplicados aos eventos. Para a monitoração dos algoritmos de classificação, a declaração das variáveis a serem monitoradas é realizada internamente no algoritmo, utilizando classes

e métodos que dão suporte para exportar as informações monitoradas.

Os algoritmos de reconstrução na etapa final do HLT do ATLAS (etapa precisa) são projetados para operar de modo muito próximo¹ ao algoritmos usados na reconstrução no ambiente *offline*. Para isso, os mesmos algoritmos de reconstrução do *offline* são instanciados, e configurados de tal modo que restrições de tempo e CPU do HLT sejam atendidas. Estes algoritmos tem sido extensivamente validados pela comunidade do ATLAS responsáveis pelas análises físicas, sendo utilizados para recalculas as grandezas de calorimetria (ver Tabela 2.5) que descrevem a deposição de energia no sistema de calorimetria.

Para realizar a monitoração das variáveis de calorimetria na etapa precisa, e ao mesmo tempo, preservar as propriedades fundamentais dos algoritmos de reconstrução do *offline*, algoritmos foram desenvolvidos para monitorar a saída dos algoritmos de reconstrução, e publicar as variáveis desejadas para monitoramento ao sistema de avaliação e qualidade de dados do ATLAS, conforme indicado na Figura B.1.

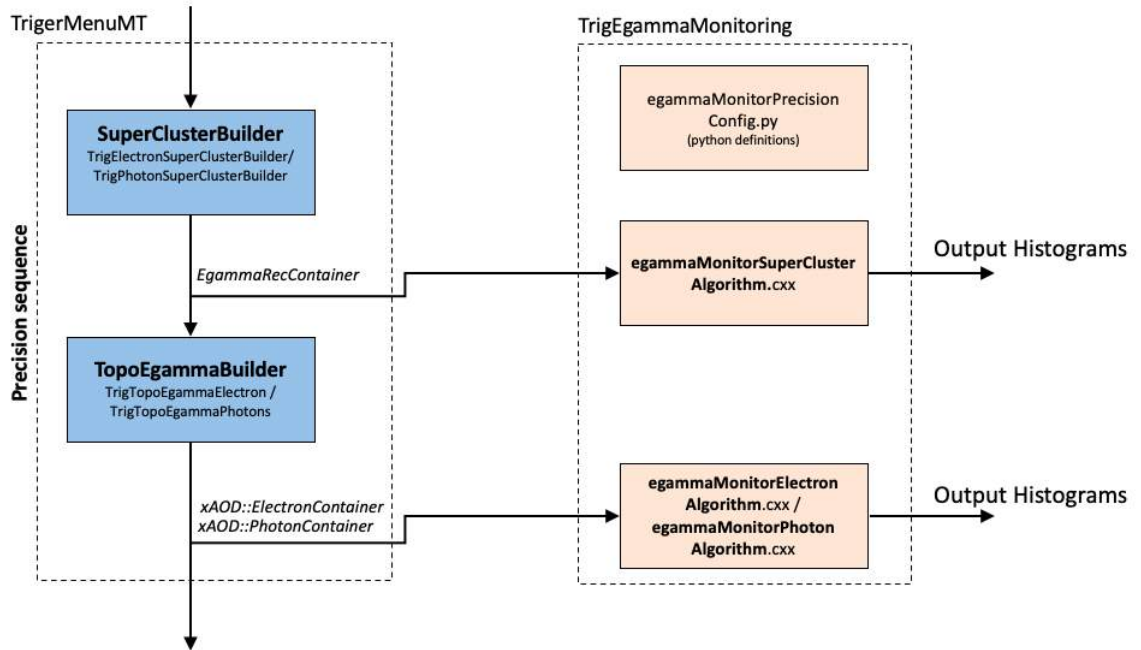


Figura B.1: Diagrama com o fluxo de operação dos algoritmos implementados para monitoração da etapa de reconstrução no HLT (etapa precisa).

Foram monitorados os algoritmos responsáveis pela geração dos *superclusters*².

¹Devido às restrições temporais do HLT nem todas as variáveis de calorimetria avaliadas no ambiente *offline* podem ser reconstruídas.

²Em substituição ao algoritmo de janela deslizante explorado no ATLAS para a reconstrução de aglomerados de tamanho fixo de células do calorímetro, a reconstrução *offline* de elétrons e fótons foi aprimorada para usar aglomerados dinâmicos de tamanho variável, chamados superaglomerados (*superclusters*). Os *clusters* dinâmicos mudam de tamanho conforme necessário para recuperar energia de fótons por *bremsstrahlung* [21] ou de elétrons de conversões de fótons, melhorando o

e responsáveis pela última etapa no HLT que realiza a reconstrução das grandezas *shower shapes*. A implementação contou com módulos responsáveis pela monitoração de elétrons e fótons, nos quais foram integrados à um *framework* dedicado para fins de monitoração *online* e *offline*. Na Figura B.2 observa-se o histograma de E_T produzido pelo algoritmo de monitoração desenvolvido, em um arquivo de teste obtido após a emulação do HLT.

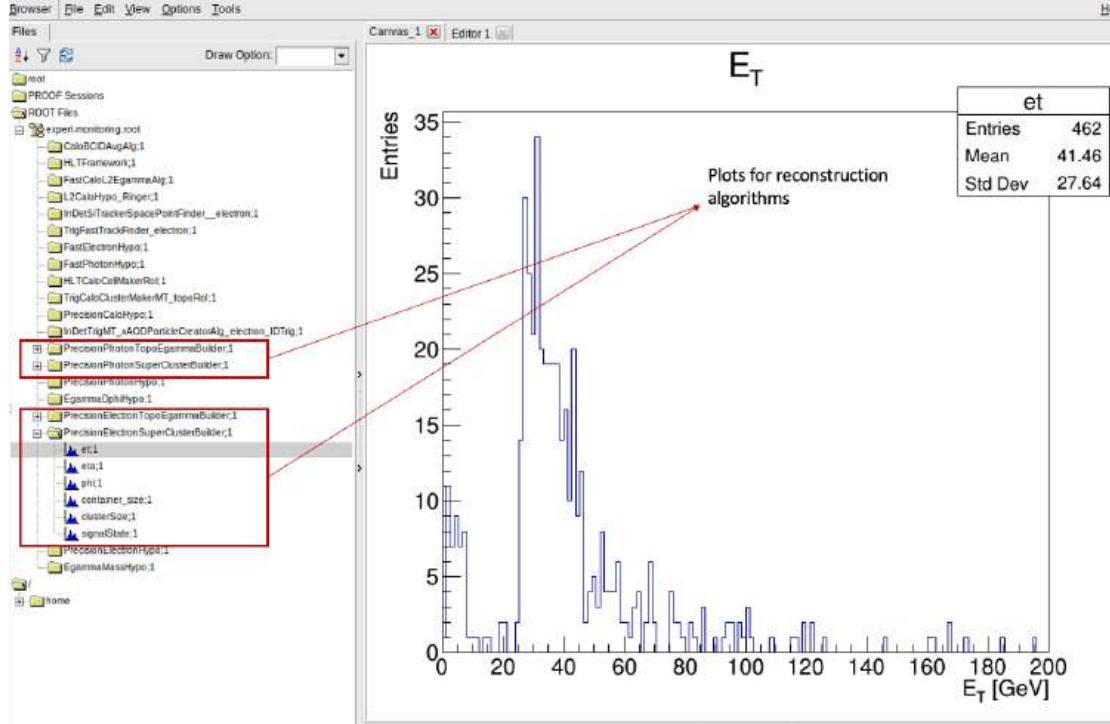


Figura B.2: Arquivo de teste para verificação dos histogramas de saída das variáveis selecionadas para monitoração.

Os algoritmos foram implementados na infraestrutura de computação do experimento ATLAS³, nos quais foram validados para monitoração da etapa precisa em cadeias de elétrons e fótons.

desempenho dos algoritmos de calibração [130].

³O HLT é baseado no ambiente de *software offline* chamado Athena [125]. O Athena tem como principal infraestrutura de *software* o Gaudi [126], que consiste em um *framework* para processamento de dados para experimentos HEP.