

IMPLEMENTAÇÃO E AVALIAÇÃO DE UM ALGORITMO DE ESTIMAÇÃO DE DENSIDADES BASEADO NO MKDE APLICADO À CLASSIFICAÇÃO DE ELÉTRONS NO DETECTOR ATLAS

IGOR ABRITTA COSTA*, DAVID DE MELO SOUZA*, RAFAEL ANTUNES NÓBREGA*, JOSÉ MANOEL DE SEIXAS†

**Universidade Federal de Juiz de Fora
Juiz de Fora-MG, Brasil*

†*Universidade Federal do Rio de Janeiro
Rio de Janeiro-RJ, Brasil*

Emails: igorabritta@gmail.com, david.melo@engenharia.ufjf.br, ranobr@gmail.com, seixas@lps.ufrj.br

Abstract— The electron identification is of fundamental importance for the physics demanded by ATLAS experiment due to the presence of these in the final process of interest particles decay. In particle physics environment, the occurrence probability of relevant electrons are very low compared to particles considered background, requiring algorithms with detection efficiency index and background rejection each time better. In this work, the methods applied in the electron identification in ATLAS experiment will be reviewed and possible optimizations will be evaluated from the data produced by the ATLAS experiment. Concentrated in the offline context, the work reproduces the method based on Likelihood and proposes an improvement with the use of multivariate technique MKDE (Multivariate Kernel Density Estimation), capable of mitigate the error inserted in consideration of independence between discriminating variables entered by the method Likelihood as is the case of ATLAS.

Keywords— Electron identification, Likelihood, Multivariate KDE

Resumo— A identificação de elétrons é de fundamental importância para a física que se deseja observar no experimento ATLAS, devido à presença destes no processo final de decaimento de partículas de interesse. Nesse ambiente de física de partículas, a probabilidade de ocorrência de elétrons, que são de interesse, são baixíssimas em relação à partículas consideradas ruído de fundo, exigindo algoritmos com índices de eficiência de detecção e rejeição de ruído de fundo cada vez melhores. Nesse trabalho, os métodos aplicados na identificação de elétrons no Experimento ATLAS serão revisados e possíveis otimizações serão avaliadas. Concentrado no contexto *offline*, o trabalho reproduz o método baseado em verossimilhança e propõe uma melhoria com o uso da técnica multivariada MKDE (do inglês, Multivariate Kernel Density Estimation), capaz de mitigar o erro inserido na consideração de independência entre as variáveis discriminantes, inserida em aplicações do método de verossimilhança, como é o caso no ATLAS.

Palavras-chave— Identificação de Elétrons, Verossimilhança, KDE Multivariado.

1 Introdução

Em ambientes industriais ou em laboratórios de pesquisa, a análise multivariada tem se mostrado uma ferramenta poderosa na resolução de vários problemas ligados a estimação e seleção de eventos. Problemas que requerem soluções deste tipo ocorrem em várias áreas do conhecimento.

Nas últimas décadas, vários experimentos geradores de enorme quantidade de dados foram iniciados, fazendo com que a importância da estimação de densidade, utilizando-se de métodos não paramétricos, crescesse consideravelmente. Os Experimentos ligados ao acelerador de partículas, conhecido como LHC (do inglês *Large Hadron Collider*), representam alguns deles (CERN, 2015). Em geral, a física experimental de altas energias é um dos ramos da ciência que mais exige dos sistemas de processamento, tendo em vista que os eventos de interesse são muito raros e contaminados com alto nível de ruído de fundo, devendo o sistema ser capaz de eliminar a maior quantidade possível dos eventos de ruído de fundo sem

descartar os raros eventos de interesse. Isso faz com que o desempenho dos algoritmos aplicados para este fim seja essencial e uma busca contínua pela otimização dos mesmos faz-se indispensável.

Nos anos de 2013 (Collaboration et al., 2013) e 2014 (Collaboration et al., 2014), a colaboração ATLAS publicou análises e resultados da técnica de verossimilhança aplicada na identificação de elétrons. Contudo, o grupo utiliza uma simplificação da formulação matemática desse método, assumindo independência entre as variáveis. Portanto, possibilita a investigação do impacto desta simplificação no desempenho do método.

Este trabalho se concentra na identificação de elétrons, no detector ATLAS, analisando os subprodutos das colisões geradas no LHC. Mais especificamente, na estimação de densidades conjuntas não-paramétricas, no intuito de mitigar o problema de dependência entre as variáveis utilizadas pelo algoritmo de identificação por verossimilhança. Foi desenvolvido um algoritmo para estimação de densidades uni e multidimensional, efetuando a comparação dos mesmos.

Este documento está organizado da seguinte maneira: a Seção 2 apresenta uma síntese da identificação de elétrons no detector ATLAS, na Seção 3 uma discussão sobre estimação de densidades de Núcleo (KDE) será apresentada, a Seção 4 apontará os resultados e análises; e finalmente na Seção 5 as conclusões serão pontuadas.

2 Identificação de elétrons no experimento ATLAS

Com 46 metros de comprimento, 25 metros de altura, 25 metros de largura e 7 mil toneladas, o detector ATLAS é o maior detector de partículas já construído (Figura 1). Situa-se em uma caverna a 100 metros, no CERN, como pode ser visto na Figura 2, próximo da cidade de Meyrin, na Suíça.

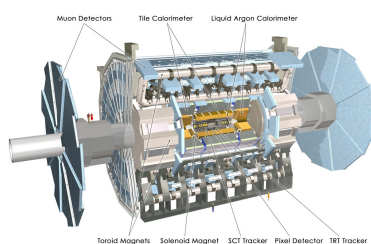


Figura 1: Detector ATLAS. Extraído de (CERN, 2016).

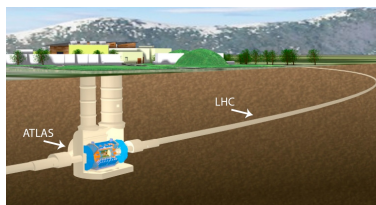


Figura 2: Vista subterrânea do detector ATLAS. Extraído de (CERN, 2016).

Feixes de prótons são acelerados em direções opostas pelo LHC e colidem no centro do Detector ATLAS, gerando subprodutos sob a forma de novas partículas, que se espalham em todas as direções. Por ser um detector de uso geral, o ATLAS deve ser capaz de identificar os mais diversos tipos de partículas assim como medir suas energias e trajetórias. Para isto, seis subsistemas de detecção diferentes foram dispostos em camadas ao redor do ponto de colisão, fazendo com que a reconstrução completa dos eventos fosse possível.

2.1 Perfil dos eventos do ATLAS

O perfil dos eventos de interesse do ATLAS e suas particularidades ditaram as características de construção e modulação do detector. Cada componente deste equipamento foi especificado para medir um conjunto de propriedades das partículas

geradas na colisão (Lippmann, 2012). Um exemplo do processo de interação das principais partículas com o detector pode ser visto na Figura 3. Os múons, prótons e elétrons deixam sinal no detector de traço; os múons, elétrons, fótons e prótons no calorímetro eletromagnético; os múons, prótons, nêutrons no calorímetro hadrônico; e a câmara de múons somente detecta os múons.

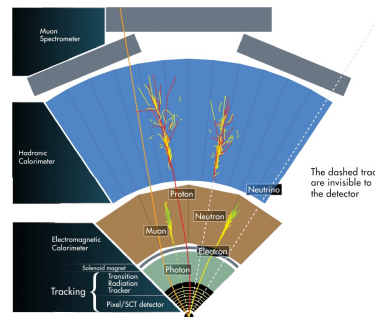


Figura 3: Assinatura das partículas no detector ATLAS. Extraído de (Landua, 2009).

2.2 Sistema de filtragem do ATLAS

No intuito de alcançar novas descobertas, utilizando-se de eventos raros, o experimento efetua colisões com uma taxa muito alta de operação, gerando um conjunto muito grande de informações, onde grande parte desses eventos podem ser descartados a fim de evitar o armazenamento de dados não relevantes (Watts, 2003).

Nesse contexto, faz-se necessário um Sistema de Filtragem *Online*, capaz de separar os eventos considerados importantes e armazená-los para uma análise *Offline*. Essas análises utilizam algoritmos mais complexos e criteriosos para selecionar as partículas de interesse da análise proposta.

2.3 Algoritmos de identificação de elétrons

Uma boa performance dos algoritmos na reconstrução e identificação de elétrons é um requisito fundamental para o sucesso do programa científico do experimento ATLAS, uma vez que as principais assinaturas dos processos eletrofracos são os léptons (Alison, 2014), sendo utilizados para inúmeras análises, como, por exemplo, as medidas de precisão do modelo padrão, a descoberta do bóson de *Higgs* e a busca por uma nova física além do Modelo Padrão.

Os elétrons isolados produzidos nos processos físicos de interesse estão sujeitos a uma grande quantidade de ruídos de fundo provenientes de hádrons identificados equivocadamente, fótons convertidos e elétrons não-isolados. Por essa razão, é de primordial importância alcançar uma eficiente identificação de elétrons, sobre todo o detector, e manter uma grande rejeição de ruído de fundo.

A colaboração ATLAS utiliza dois algoritmos para identificação de elétrons: O T2Calo, que atua na parte rápida da seleção, usando a informação de calorimetria, que é combinada com a de traço, e um algoritmo baseado em verossimilhança, que atua na parte de precisão.

O T2Calo efetua a identificação de elétrons utilizando cortes nas PDFs (do inglês *Probability Density Function*) de variáveis altamente discriminantes provenientes do detector. Já o algoritmo baseado em verossimilhança consiste na estimação das densidades dessas variáveis e cálculo do discriminante utilizando a PDF conjunta.

3 Estimação de densidade

O KDE (do inglês *Kernel Density Estimation*) é uma técnica não-paramétrica para estimação de densidades de distribuição onde cada observação é ponderada pela distância em relação a um valor central, o núcleo. Após a estimação da densidade de um conjunto de dados, tem-se, supostamente, conhecimento sobre a probabilidade de ocorrência de eventos futuros, podendo ser usado em problemas de classificação (Scott, 2015).

3.1 Estimação de densidade univariada

Uma das maneiras mais simples de abordar esse problema é a estimação por histograma, equacionado como segue:

$$\hat{f}_h(x) = \frac{1}{nh} \sum_{i=1}^n \sum_j I(x_i \in B_j) I(x \in B_j) \quad (1)$$

onde,

$$B_j = [x_o + (j-1)h, x_o + jh) \quad j \in \mathbb{Z} \quad (2)$$

Esse método lida com os parâmetros de origem (x_o) e tamanho do *bin* (h), sendo que ambos, se escolhidos de maneira errada, podem resultar em uma má estimação. No intuito de diminuir esse problema, é possível efetuar a estimação pela média de sub-histogramas deslocados, tornando o método independente da origem (Seather, 1992). Entretanto, o problema de otimização do parâmetro h continua e sua otimização foi considerada em vários trabalhos, como por exemplo: (Abramson, 1982), (Silverman, 1986) e (Comaniciu et al., 2001).

Outra forma de estimação de densidades muito utilizada na literatura é o uso de um estimador por núcleo, que de forma direta, aplica uma função *Kernel* em cada evento, considerando a contagem de eventos em sua vizinhança. O estimador é dado pela Equação 3 e foi Rosenblatt em 1956 o primeiro autor que considerou esse modelo.

$$\hat{f}(x_i) = \frac{1}{nh} \sum_{k=1}^n K\left(\frac{(x_i - X_k)}{h}\right) \quad (3)$$

onde $K(u)$ é a função *Kernel*, h é a largura de banda e n é o número de eventos em torno de x_i . Para garantir que esse modelo retorne uma função de densidade, o *Kernel* deve satisfazer $\int_{-\infty}^{+\infty} K(u) du = 1$ e $K(u) \geq 0 \quad \forall \quad u \in \mathbb{R}$ (Funções *Kernels* de ordem maiores que 2 não satisfazem a última propriedade).

Alguns exemplos de funções *kernel* de segunda ordem são (Scott, 2015): Triangular, Epanechnikov, Quartic, Triweight e Gaussiana, sendo que a última foi a escolhida para esse trabalho.

Em ambas as abordagens, histogramação e KDE, fica clara a dependência da estimação em relação ao parâmetro h de largura de banda. Na literatura existem algumas formas de otimização da estimação de densidades por meio da escolha ótima do parâmetro h . Nesse trabalho, a escolha do parâmetro h foi feita utilizando o método L_p , que é uma ideia bastante comum para medir a diferença entre duas funções, dado pela Equação 4:

$$L_p = \left(\int |f - g|^p \right)^{1/p} \quad (4)$$

onde p é o parâmetro a ser escolhido. Quando $p = 2$ na Equação 4, temos:

$$ISE(\hat{f}_h) = \int [\hat{f}_h(x) - f(x)]^2 dx \quad (5)$$

que é, então, chamada de Integral do Erro Quadrático (do inglês, *Integrated Squared Error* (ISE)), sendo assim, para o cálculo da Integral do Erro Quadrático Médio (do inglês, *Mean Integrated Squared Error* (MISE)), utiliza-se o valor esperado de $f(x)$, obtendo a Equação 6.

$$MISE(\hat{f}_h) = \int E\{\hat{f}_h(x) - f(x)\}^2 dx \quad (6)$$

Do mesmo modo que o Erro Quadrático Médio (do inglês, *Mean Squared Error* (MSE)), que é a medida do erro médio de um certo ponto x , a expressão do MISE pode ser decomposta em termos de *bias* e variância, como mostrada na Equação 7.

$$\begin{aligned} MSE(\hat{f}_h) &= E\{\hat{f}_h(x) - f(x)\}^2 \\ &= Bias(\hat{f}_h(x))^2 + var(\hat{f}_h(x)) \\ &\simeq \left(\frac{1}{2}f''(x)h^2\mu_2(K)\right)^2 + \\ &\quad + \frac{f(x)R(K)}{nh} + o\left(\frac{1}{nh} + h^4\right) \end{aligned} \quad (7)$$

Na Equação 7, o último termo é devido ao erro de truncamento da expansão de Taylor, como sugerido em (Hansen, 2004) faz-se uma aproximação assintótica no MSE, desconsiderando esse termo, a equação recebe o nome de AMSE. Quando se trata de distribuições e não de pontos, a medida de erro é chamada de MISE, e sua aproximação assintótica AMISE, dada pela Equação 8.

$$\begin{aligned} AMISE &= \int AMSE(\hat{f}(x)) dx \\ &= \frac{1}{4}h^4\mu_2(K)^2R(f'') + \frac{1}{nh}R(K) \end{aligned} \quad (8)$$

onde μ_2 representa a variância do *Kernel* de segunda ordem e $R(\cdot)$ a rugosidade.

Temos então que o primeiro termo (*bias*) aumenta proporcionalmente ao parâmetro h , já o segundo termo (variância) diminui de maneira inversamente proporcional ao aumento de h , com isso observa-se claramente um conflito entre reduzir a variância e o *bias* de forma simultânea, visto que a escolha de um h pequeno para garantir um menor *bias* ocasiona uma variância grande (Scott, 2015).

Na Equação 8, fazendo a primeira derivada igual a zero, pode-se calcular a largura de banda ótima, dada pela Equação 9.

$$h_{AMISE} = \left[\frac{R(K)}{\mu_2(K)^2 R(f'') n} \right]^{1/5} \quad (9)$$

Como a Equação 9 mostra, o problema de calcular a largura de banda ótima via AMISE, com o conhecimento somente da distribuição, é circular. Assumiu-se uma distribuição normal com variância σ^2 para a escolha automática desse parâmetro (Silverman, 1986). Com isso o parâmetro de largura de banda fixa h que minimiza o AMISE é:

$$h = 1,06\sigma n^{-1/5} \quad (10)$$

A largura de banda h , como foi abordado, é um parâmetro crucial, e pode ser utilizado de duas formas, fixa ou variável. A segunda abordagem possibilita a utilização de uma largura de banda $h(x_i)$ para cada ponto de x_i em que desejamos estimar a probabilidade $f_{h_i}(x_i)$. Esse estimador é conhecido como *balloon estimator* e tem a forma:

$$\hat{f}_{h_i}(x_i) = \frac{1}{nh(x_i)} \sum_{k=1}^n K\left(\frac{(x_i - X_k)}{h(x_i)}\right) \quad (11)$$

onde $h(x_i)$ é uma largura de banda que varia de acordo com o ponto x_i .

Para calcular o parâmetro de largura de banda variável $h(x_i)$ utiliza-se a Equação 12, como sugerido por (Abramson, 1982).

$$h(x_i) = \frac{h}{\sqrt{f_p(x_i)}} \quad (12)$$

onde, h é uma largura de banda fixa e $f_p(x_i)$ é a probabilidade de x_i na PDF. Para estimar esse último parâmetro fez-se um estudo sobre métodos de validação cruzada, que utilizam os critérios de ISE e MISE; e então utilizando um algoritmo proposto por (Shimazaki and Shinomoto, 2007) de estimação da binagem ótima para um histograma, que aborda o problema fazendo uso de uma função custo que minimiza o critério MISE. Entretanto, essa escolha não mostrou um bom resultado devido as variedades nas formas das distribuições analisadas nesse trabalho. Para contornar este problema foi necessário inserir outro parâmetro, o λ , chamado de constante de proporcionalidade proposto por (Comaniciu et al., 2001), como mostram as equações 13 e 14:

$$h(x_i) = h \left[\frac{\lambda}{f_p(x_i)} \right]^{\frac{1}{2}} \quad (13)$$

sendo λ dado por:

$$\lambda = e^{n^{-1} \sum_{i=1}^n \log(f_p(x_i))} \quad (14)$$

3.2 Estimação de densidade Multivariada

O conceito de KDE univariado precisa ser ampliado quando precisa-se investigar a dependência entre as variáveis do problema. Tendo como base a teoria univariada, uma generalização do KDE é mostrada na Equação 15:

$$\begin{aligned} f_{h_1, h_2, \dots, h_v}(x_1, x_2, \dots, x_v) &= \\ &= \frac{1}{n} \sum_{k=1}^n \frac{1}{h_1} \frac{1}{h_2} \dots \frac{1}{h_v} K\left(\frac{(x_1 - X_{k1})}{h_1}\right) \\ &K\left(\frac{(x_2 - X_{k2})}{h_2}\right) \dots K\left(\frac{(x_v - X_{kv})}{h_v}\right) \end{aligned} \quad (15)$$

O raciocínio é análogo e, em nosso algoritmo, foi utilizado a forma generalizada matricial. Primeiro, a largura de banda foi modelada como:

$$H = \text{diag}(h_1, h_2, \dots, h_v) \quad (16)$$

Foi escolhida a largura de banda fixa como proposto por (Wand and Jones, 1995), sendo uma generalização multidimensional de h .

$$h_j = \left(\frac{4}{d+2} \right)^{\frac{1}{(d+4)}} n^{\frac{-1}{(d+4)}} \sigma_j \quad (17)$$

onde, d representa o número de dimensões do problema, j é o subíndice da respectiva dimensão, n é o número de eventos e σ_j é o desvio padrão dos eventos da dimensão j . Note-se que, para $d=1$, a Equação 17 coincide com o método proposto por Silverman, dado pela Equação 10.

Sendo $x = x_1, x_2, \dots, x_v$, tem-se a fórmula para o KDE Multivariado, denominado de MKDE:

$$\begin{aligned} f_H(x) &= \frac{1}{n} \sum_{k=1}^n \frac{1}{\det(H)} K\{H^{-1}(x - X_k)\} \\ &= \frac{1}{n} \sum_{k=1}^n K_H(x - X_k) \end{aligned} \quad (18)$$

com,

$$K_H(a) = \frac{1}{\det(H)} K(H^{-1}a), a = (x - X_k) \quad (19)$$

4 Resultados

Uma boa estimação da densidade conjunta das variáveis discriminantes é central para o método de Verossimilhança e, devido a isto, este trabalho se concentrou principalmente neste tópico. Para tal, dados de Monte Carlo gerados pela Colaboração ATLAS foram utilizados.

O menu utilizado contempla 13 vari veis discriminantes e os dados foram divididos em 12 regi es, tr s em E_t e quatro em η , ponderando sobre a estat stica dispon vel e regi es de interesse, sendo que os resultados aqui mostrados s o referentes a regi o 1, que compreende os dados com $E_t < 20\text{GeV}$ e $|\eta| < 0.8$, interessantes por ser onde mais existe contamina  o de ru do de fundo.

O conjunto de dados foi dividido igualmente em treinamento e valida  o. O conjunto de treinamento   utilizado para estimar as PDFs e no c lculo de par metros (por exemplo, a matrix de AUC (do ingl s, *Area Under the Curve*) na Figura 6) utilizados na verossimilhan a, j  o conjunto de valida  o   utilizado para avaliar o desempenho dos algoritmos constru dos.

A implementa  o do algoritmo de estima  o de densidades foi efetuado em conformidade com as formula  es matem ticas descritas na Se  o 3.1 e seu resultado pode ser visto na Figura 4, onde s o mostradas 4 distribui  es, duas de sinal e duas de ru do de fundo, sendo que a escolha da melhor binagem para visualiza  o do histograma foi feita utilizando a medida do χ^2 .

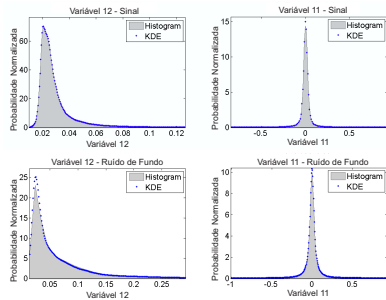


Figura 4: Histograma da vari vel e a sua PDF estimada pelo KDE. (Sup. Esq.) Vari vel 12 de sinal; (Sup. Dir.) Vari vel 11 de sinal; (Inf. Esq.) Vari vel 12 de ru do de fundo; (Inf. Dir.) Vari vel 11 de ru do de fundo.

De forma an loga, o resultado da implementa  o do algoritmo de estima  o de densidades multivariadas, para o caso bidimensional, pode ser observado na Figura 5, trazendo ao lado das mesmas, para fins de compara  o, as densidades estimadas a partir do m todo univariado, utilizando-se de suas marginais (do ingl s, *Marginal*) e do histograma normalizado da distribui  o.

Uma vez com as densidades estimadas, pode-se utiliz -las no algoritmo de identifica  o de el trons, baseado no m todo de verossimilhan a, e efetuar uma compara  o entre o uso das PDFs constru das de forma marginal e conjunta. Entretanto a escolha dos pares de vari veis discriminantes que s o utilizados para estima  o de densidade bivariada   de suma import ncia, sendo que o m todo de an lise aqui proposto   observar a diferen a entre AUC (do ingl s, *Area Under the Curve*) das an lises conjunta (multivariada) e

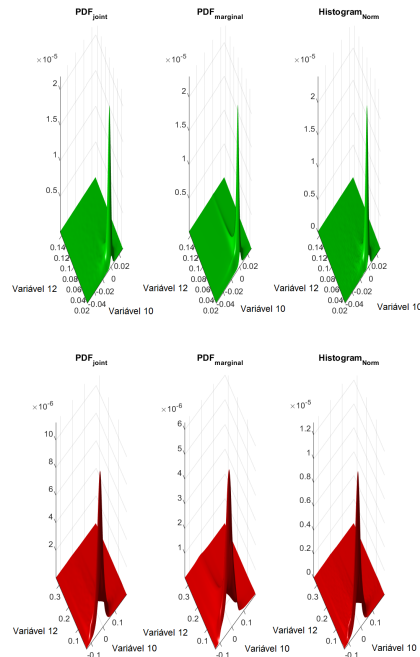


Figura 5: Gr fico que ilustra para um par de vari veis a diferen a entre a estima  o de densidade conjunta (*Joint*) e marginal (*Marginal*), com o histograma normalizado da distribui  o. (Superior) Sinal; (Inferior) Ru do de Fundo

marginal (univariada) para todos os pares de vari veis e ent o escolher as que apresentarem os maiores  ndices de melhoria, no intuito de otimizar a performance do algoritmo de classifica  o de eventos multivariado.

A Figura 6 apresenta a diferen a de AUC das ROCs obtidas a partir dos algoritmos de estima  o unidimensional e bidimensional aplicados a cada um dos pares poss veis entre as 13 vari veis discriminantes. Observa-se que os pares 4 e 7; 10 e 12; 5 e 6; 9 e 11 s o poss veis escolhas baseando-se nessa proposta.

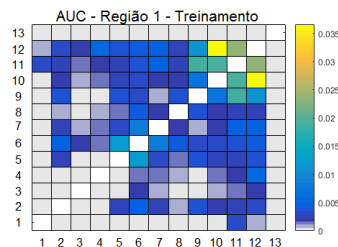


Figura 6: Gr fico de AUC com a combina  o de todos os pares de vari veis, para a Regi o 1.

A Figura 7 mostra a compara  o, via curva ROC, entre a utiliza  o da an lise Univariada e Bivariada para os dados de valida  o da regi o 1. Observando esses resultados, tem-se um indicativo que o m todo de an lise multivariada apresenta

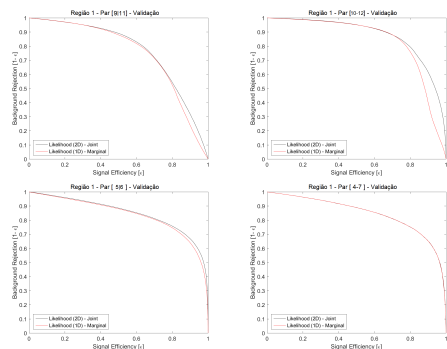


Figura 7: Curva ROC da Região 1 comparando a análise Univariada com a Bivariada para os eventos de CC. (Sup. Esq.) Par de variáveis - 9 e 11; (Sup. Dir.) Par de variáveis - 10 e 12; (Inf. Esq.) Par de variáveis - 5 e 6 e (Inf. Dir.) Par de variáveis - 4 e 7.

um pequeno ganho de performance, uma vez que este método diminui o impacto dos eventuais erros causados pela consideração de independência das variáveis discriminantes.

5 Conclusões

Este trabalho apresentou resumo dos estudos, implementação e análise relacionados a um algoritmo de estimação de densidade não-paramétrico baseado no MKDE aplicado à seleção de elétrons para o Experimento ATLAS usando a técnica de verossimilhança. A identificação de elétrons é de suma importância para o programa de física do ATLAS.

Os resultados apresentados indicam uma pequena melhora no uso de PDFs bidimensionais estimadas pelo método de MKDE na Verossimilhança. Atualmente há um novo conjunto de dados disponível para análise e será usado para confirmação desses resultados. Ressalta-se que os Experimentos do LHC estão em evolução e novas fases com energia e luminosidade maiores são previstas. Com isso, o aprofundamento deste trabalho pode levar a um aprimoramento do sistema de seleção de eventos com impacto direto nos resultados das pesquisas desenvolvidas nesse ambiente.

Agradecimentos

Ao Programa de Pós-Graduação em Engenharia Elétrica (PPEE) da Universidade Federal de Juiz de Fora (UFJF), a CAPES, ao CNPQ, a FAPERJ e a RENAFAP (MCTI) por darem suporte estrutural e financeiro para a realização deste artigo, e a Colaboração ATLAS pelas discussões frutíferas acerca deste trabalho.

Referências

Abramson, I. S. (1982). On bandwidth variation in kernel estimates-a square root law, *The an-*

nals of Statistics pp. 1217–1223.

Alison, J. (2014). *The road to discovery: Detector alignment, electron identification, particle misidentification, ww physics, and the discovery of the Higgs Boson*, Springer.

CERN (2015). Experiments, <http://home.cern/about/physics>. Accessed in 20-04-2016.

CERN (2016). Estrutura do detector atlas, <https://kjende.web.cern.ch/kjende/pt/atlas.htm>. Accessed in 23-04-16.

Collaboration, A. et al. (2013). Description and performance of the electron likelihood tool at atlas using 2012 lhc data, *Technical report*, ATLAS-COM-PHYS-2013-378. <http://cds.cern.ch/record/1537410>.

Collaboration, A. et al. (2014). Electron efficiency measurements with the atlas detector using the 2012 lhc proton-proton collision data, ATLAS-CONF-2014-032.

Comaniciu, D., Ramesh, V. and Meer, P. (2001). The variable bandwidth mean shift and data-driven scale selection, *Computer Vision, 2001. ICCV 2001. Proceedings. Eighth IEEE International Conference on*, Vol. 1, IEEE.

Hansen, B. E. (2004). Nonparametric estimation of smooth conditional distributions, *Dept. of Economics University of Wisconsin*.

Landua, R. (2009). O lhc, <http://www.scienceinschool.org/pt/2008/issue10/lhchow>. Accessed in 15-04-2016.

Lippmann, C. (2012). Particle identification, *Nucl. Instrum. Meth.* **A666**: 148–172.

Scott, D. W. (2015). *Multivariate density estimation: theory, practice, and visualization*, John Wiley & Sons.

Seather, S. (1992). The performance of six popular bandwidth selection methods on some real data sets, *COMPUTATIONAL STATISTICS QUARTERLY* **7**: 225–225.

Shimazaki, H. and Shinomoto, S. (2007). A method for selecting the bin size of a time histogram, *Neural computation* **19**(6).

Silverman, B. W. (1986). *Density estimation for statistics and data analysis*, CRC press.

Wand, M. and Jones, M. (1995). Kernel smoothing, *Chapman & Hall, London*.

Watts, G. (2003). Review of triggering, *Nuclear Science Symposium Conference Record, 2003 IEEE*, Vol. 1, pp. 282–287 Vol.1.