

Máquinas de Aprendizado Extremo para Classificação Online de Eventos no detector ATLAS

Marton S. dos Santos, Edmar E. P. de Souza, Eduardo F. Simas Filho, Paulo C. M. A. Farias, Werner S. Freund, e José M. de Seixas

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.

Palavras-Chave—ELM, Redes Neurais, Processamento de Sinais, Detector ATLAS.

Abstract—ATLAS is one of the LHC (*Large Hadron Collider*) detectors, and is located at the European Organization for Nuclear Research (CERN). For proper characterization of the particles an accurate measurement of the energy deposition profile is required as interactions with the detector occur. In ATLAS, the calorimeters are responsible for estimating the energy of the particles and, in this sense, they use more than 100,000 sensors. One of the discriminators for the online detection of electrons used in ATLAS is the *Neural Ringer*, in which the energy deposition profile is used as input to a multi-layer perceptron neural classifier. This work proposes the use of Extreme Learning Machines (ELM) in substitution of perceptron neural networks in *Neural Ringer*. The results obtained from a simulated database point to a significant reduction in training time, maintaining similar classification performance.

Keywords—ELM, Neural Network, Signal Processing, ATLAS detector.

I. INTRODUÇÃO

A compreensão a respeito da constituição fundamental da matéria obteve evolução significativa nos últimos anos devido a comprovações, resultantes de experimentos de física de altas energias. O Grande Colisor de Hadrons (*Large Hadron Collider* - LHC) [1] é o maior acelerador de partículas em operação atualmente situado no CERN (Organização Europeia para a pesquisa Nuclear)[2]. O objetivo do LHC é analisar a estrutura fundamental da matéria, investigar as propriedades das partículas fundamentais propostas no Modelo Padrão [3] e também buscar por fenômenos desconhecidos.

Marton S. dos Santos, Edmar E. P. de Souza Eduardo F. de Simas Filho, Paulo C. M. A. Farias, PPGEE-UFBA, Salvador-Ba, Brasil, Werner S. Freund e José M. de Seixas: Laboratório de Processamento de Sinais - COPPE/Poli, UFRJ, RJ Brasil, E-mail: mssandes@gmail.com,

Para tanto, o LHC, conta com alguns detectores de partículas como ATLAS (*A Toroidal LHC ApparatuS*), ALICE (*A Large Ion Collider Experiment*), CMS (*Compact Muon Solenoid*) e LHCb (*Large Hadron Collider beauty*). O túnel do acelerador tem cerca de 27 km de comprimento com colisões ocorrendo a taxa de até 40×10^6 vezes por segundo [1].

O ATLAS é um detector de propósito geral, ver Figura 1, ou seja, possui capacidade de detectar diversos tipos de partículas. Para isso, foi projetado no formato cilíndrico e conta com os seguintes sub-detectores: detector de traços (identifica a trajetória das partículas carregadas eletricamente); calorímetro (medidor de energia altamente segmentado); e a câmara de múons (projetado especificamente para a detecção de múons). Como resultado dessa estrutura altamente segmentada e da alta taxa de colisões, é produzida uma taxa de dados da ordem de 60 TB/s, o que requer a detecção (ou *trigger*) online das assinaturas de interesse.

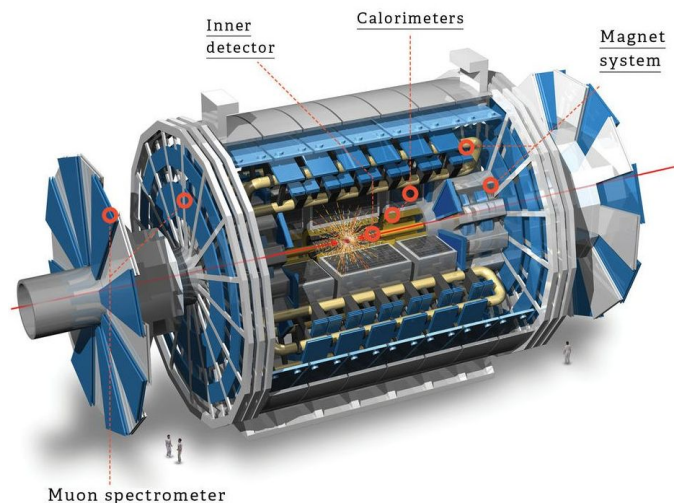


Fig. 1: Diagrama do Detector ATLAS, com destaque para seus sub-detectores. Fonte: [2].

O sistema de seleção ou filtragem *online* (*trigger*) do ATLAS [4] é responsável pela seleção dos eventos interessantes para o experimento e, também pela redução do ruído de fundo (assinaturas não relevantes) produzido nas colisões. Conforme mostrado na Figura 2, o sistema de *trigger* do ATLAS é composto por três estágios sequenciais de seleção.

O primeiro nível de *trigger* (*Level 1* ou L1), tem disponível uma janela de tempo de até $2,5 \mu s$ para tomada de decisão e utiliza informações dos calorímetros e das câmaras de múons para reduzir a taxa inicial de eventos para 75 kHz. O L1

é desenvolvido em *hardware* dedicado e tem a importante função de determinar as regiões onde mais provavelmente ocorreram eventos relevantes, ou como são denominadas: Regiões de Interesse (*Regions of Interest - RoIs*).

A partir dos sinais coletados nas RoIs, o nível 2 (L2) utiliza toda a granularidade do detector, com objetivo de reduzir ainda mais a taxa de eventos para 1 kHz. A partir do L2, inclusive, a implementação passa a ser por *software* e a tomada de decisão deve levar no máximo 40 ms.

O último estágio de filtragem *online*, o Filtro de Eventos (*Event Filter - EF*), tem como função refinar ainda mais a decisão realizada no L2. O EF possui um tempo de tomada de decisão de até 4 s, e reduzir a taxa de eventos para no máximo 300 Hz. As assinaturas aprovadas no EF são armazenados em mídia permanente para as posteriores análises *offline*.

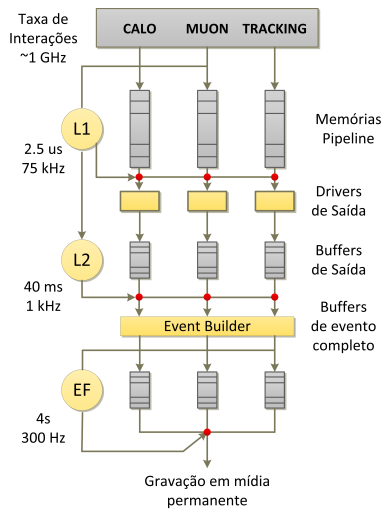


Fig. 2: Esquema do *trigger online* do ATLAS. Fonte: [5].

Os elétrons são partículas importantes para o ATLAS pois estão envolvidos em fenômenos de rara ocorrência (Ex. Bóson de Higgs) [6]. A identificação de elétrons, é realizada em meio a uma elevada taxa de ruído de fundo, basicamente composto de partículas hadrônicas, conhecidas como jatos.

Para aumentar a eficiência na identificação de elétrons no ATLAS foi proposto um sistema de classificação (*Neural Ringer*) [4], que organiza a região de interesse em anéis concêntricos de deposição de energia por camada do calorímetro, conforme ilustrado na Figura 3. A energia medida nos sensores de cada anel é somada, e essa informação é utilizada para alimentar um sistema de classificação baseado numa rede neural artificial tipo perceptron de múltiplas camadas (*multi-layer perceptron - MLP*) [7].

Um problema associado ao *Neural Ringer* é o elevado tempo de treinamento do sistema, que precisa ser repetido para as diferentes configurações de operação do detector. Este trabalho propõe a utilização de uma máquina de aprendizado extremo (*Extreme Learning Machine - ELM*) em substituição ao classificador MLP no discriminador *Neural Ringer*. As redes ELM apresentam estrutura semelhante à de uma rede MLP, porém seu tempo de treinamento é menor.

O trabalho está organizado conforme descrito a seguir. Na

Seção II são apresentados as técnicas de processamento de sinais utilizadas. As principais características do discriminador *Neural Ringer* são abordadas na Seção III. A Seção IV apresenta a metodologia utilizada. Os resultados obtidos com a base de dados são discutidos na Seção V. As conclusões são apresentadas na Seção VI.

II. TÉCNICAS UTILIZADAS

Nesta Seção são apresentados os principais aspectos das técnicas de processamento de sinais utilizadas neste trabalho.

A. Redes Neurais Artificiais

Redes Neurais Artificiais (RNA) são modelos matemáticos paralelos e distribuídos, compostos por unidades de processamento simples (chamadas de neurônios artificiais) que calculam determinadas funções matemáticas (normalmente não-lineares) [8]. Um neurônio é uma unidade de processamento de informação que é fundamental às operações de uma RNA [7].

A saída do neurônio j da camada l de uma RNA pode ser expressa conforme mostrado na Equação 1:

$$v_{lj} = \varphi \left(\sum_{i=1}^N \omega_{lij} x_i + b_{lj} \right) \quad (1)$$

onde $\varphi(\cdot)$ é a função de ativação, ω_{lij} são os pesos dos neurônios, b_{lj} o *bias* (ou tendência), e x_i as entradas.

Sistemas baseados em RNA vêm sendo utilizados com sucesso em diferentes aplicações como o suporte à decisão em inspeções não-destrutivas [9], diagnóstico biomédico [10] e seleção online de eventos em detector de partículas [5].

B. Máquinas de aprendizado extremo

As máquinas de aprendizado extremo (ELM) foram propostas inicialmente em [11]. Utilizando uma estrutura semelhante à de uma rede neural MLP com uma única camada oculta (*Single Layer Feedforward Network - SLFN*), o treinamento da ELM assume que é possível gerar aleatoriamente os pesos da camada de entrada e determinar analiticamente os melhores pesos para a camada oculta. Deste modo, o tempo de treinamento de uma ELM é consideravelmente reduzido, pois não existe um procedimento iterativo de retro-propagação de erro para o ajuste dos pesos do modelo.

Para um conjunto de M pares entrada-saída (x_i, y_i) com $x_i \in \mathbb{R}^{d_1}$ e $y_i \in \mathbb{R}^{d_2}$, a saída de uma SLFN com N neurônios na camada oculta é modelada pela Equação 2.

$$y_j = \sum_{i=1}^N \beta_i \varphi(w_i x_j + b_i), \quad j \in [1, M] \quad (2)$$

sendo φ a função de ativação, w_i e b_i os pesos e o *bias* da camada de entrada, respectivamente, e β_i os pesos da camada de saída. E reescrevendo a Equação 2, $\mathbf{H}\boldsymbol{\beta} = \mathbf{Y}$, sendo,

$$\mathbf{H} = \begin{pmatrix} \phi(w_1 x_1 + b_1) & \dots & \phi(w_N x_1 + b_N) \\ \vdots & \ddots & \vdots \\ \phi(w_1 x_M + b_1) & \dots & \phi(w_N x_M + b_N) \end{pmatrix}, \quad (3)$$

$$\text{e } \boldsymbol{\beta} = (\beta^T \dots \beta_N^T)^T \text{ e } \mathbf{Y} = (y_1^T \dots y_M^T)^T.$$

A solução baseia-se em determinar a matriz inversa generalizada de Moore-Penrose de \mathbf{H} , definida como $\mathbf{H}^\dagger = (\mathbf{H}^T \mathbf{H})^{-1} \mathbf{H}^T$, o que resulta em $\beta = \mathbf{H}^\dagger \mathbf{Y}$.

A ELM vem sendo utilizada em diferentes aplicações como, por exemplo, em [12], com o objetivo de desenvolver uma ferramenta que auxiliasse na extração de características de imagens de ressonância magnética no diagnóstico de mal de Alzheimer, em [13] e em [14] como aplicação robusta contra *outliers* classificadores, em [15] sua convergência foi (ELMRank) avaliada para generalização e montagem de rank em 2 duas bases de dados, uma para indentificação de filmes, piadas e livros não lidos, e na análise de QSAR (*Quantitative Structure-Activity Relationship*) e inibidores da enzima DHFR (*dihydrofolate reductase*) para detecção de doenças. Em [16] uma estrutura com duas camadas ocultas chamada TELM (*two-hidden-layer ELM*) foi avaliada e comparada em três problemas, regressão, classificação e numa base de dados MNIST (*Modified National Institute of Standards and Technology database*), e observaram que a TELM torna-se interessante para problemas complexos de regressão e classificação e na presença de recursos computacionais de armazenamento limitados. Foi demonstrado que uma ELM, assim como uma rede MLP é um aproximador universal [11]. Entretanto, em alguns casos a ELM requer um número maior de neurônios na camada oculta para resolver, com desempenho equivalente, o mesmo problema.

III. DETECÇÃO *online* DE ELÉTRONS NO ATLAS

No ATLAS, o *Trigger* de alto nível (HLT na sigla em inglês) é constituído pelos níveis L2 e EF. No nível 2 de *trigger* (L2), a informação chega das RoIs de forma a utilizar toda a granularidade do detector para que se reduza a taxa de eventos para 1 kHz. O tempo máximo de tomada de decisão não deve ultrapassar 40 ms.

Um sistema de detecção de elétrons alternativo que poderia substituir o seletor padrão operante no L2 do detector é o *Neural Ringer*, que realiza um pré-processamento topológico da informação de cada camada do calorímetro, localizando a posição com maior concentração energética na camada formando “anéis”, ver Figura 3, rodeando o ponto indicado. São gerados, então, 100 anéis para cada RoI previamente indicada pelo L1 nas diferentes camadas do calorímetro, a saber, PS = 08, E1 = 64, E2 = 08, E3 = 08, H0 = 04, H1 = 04, H2 = 04 [5].

Depois que os sinais são agrupados em anéis, esta resultante é normalizada em relação à energia total dos anéis. A característica resultante alimenta um classificador baseado em redes neurais artificiais (RNA) *feedforward* que apresenta duas camadas com um neurônio na saída. O *Neural Ringer* demonstrou índices melhores do que o sistema de detecção padrão do ATLAS pois reduz a quantidade de informação a ser analisada no sistema de calorimetria devido ao agrupamento de várias células para cada informação, o que reduz o equivalente de 1000 células de uma RoI, para 100 anéis. O anelador mantém também a interpretação física do processo; pois as relações de profundidade longitudinal e espalhamento lateral do chuveiro de partículas se mantém.

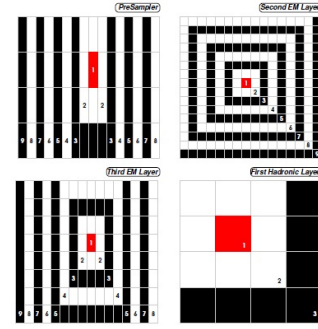


Fig. 3: Representação dos anéis do calorímetro nas RoIs. Fonte: [2].

O desenvolvimento de classificadores neurais utilizado pelo *Neural Ringer* demanda tempo de treinamento e de inicialização de um número significativo de redes neurais para cada colisão registrada. Pode-se citar o número de redes necessário à obtenção dos resultados apresentados neste trabalho, 80.000 (16 regiões, 50 sorteios e 100 inicializações). Outro fator a se considerar, são as atualizações periódicas que o detector passa, o que pode modificar as condições de operação e níveis de energia utilizados nas colisões. Observando esse contexto, a ELM pode ser utilizada como uma alternativa no processo de obtenção de redes com desempenho equivalente, porém com tempo de treinamento reduzido [11].

IV. METODOLOGIA

Dois técnicas foram utilizadas, redes neurais em avanço de camada única (*Feedforward*) utilizando como função de ativação, a função *tanh*. Os algoritmos foram comparados num PC com processador i7-4770 @ 3,40 GHz de 4 núcleos físicos e 2 lógicos por *core*, RAM com 11,6 GiB e HD com 1,8 TiB.

Nas redes MLP, foram utilizadas bibliotecas, já para avaliar a ELM, foi desenvolvido um algoritmo, em MATLAB, com base na teoria da técnica e nos trabalhos de [11], [15].

Nos treinos das redes com ELM e MLP foi utilizada a base de dados simulados por meio da técnica Monte Carlo. Nesta base, os dados estão organizados em 16 regiões, ver Tabela I, sendo 4 intervalos para o nível de energia transversa (E_T) e 4 intervalos para a posição no interior do detector (η). Esse fato, leva em consideração as diferenças ao longo da estrutura complexa do detector, como por exemplo as regiões por onde os cabos para fornecimento de energia para os mais de 100.000 sensores passam.

Os 04 valores para E_T e η , foram representados como, $E_T \in \{0, 1, 2, 3\}$ e $\eta \in \{0, 1, 2, 3\}$ respectivamente.

TABELA I: Segmentação base de dados simulados MC2014.

E_T [GeV]	[20, 30]	[30, 40]	[40, 50]	[50, 20.000]
η	[0,0, 0,8]	[0,8, 1,37]	[1,37, 1,54]	[1,54, 2,5]

O treinamento das redes foi realizado pelo método de validação cruzada [17]. Nesse método a base de dados foi subdividida em 10 subconjuntos, sendo 60% para treino e

40% para teste. Das possibilidades de subconjuntos disjuntos, sorteados 50 subconjuntos e para cada um 100 treinamentos realizados, totalizando 5000 redes para cada região.

Para avaliação de desempenho do discriminador foram utilizados a curva ROC (TreesS, 2001) e o maior índice SP obtido em cada treino, conforme definido na Equação 4. Sendo o maior índice SP o resultado da menor taxa de falhas para a maior taxa de acertos da rede.

$$SP = \sqrt{\frac{(Ef_e + Ef_j)}{2}} \times \sqrt{Ef_e \times Ef_j} \quad (4)$$

onde $Ef_e = PD$ e $Ef_j = 1 - FR$ são as eficiências obtidas, respectivamente, para elétrons e jatos (sendo PD a probabilidade de detecção de elétrons e FR probabilidade de classificar um jato hadrônico incorretamente).

V. RESULTADOS

As redes utilizando MLP com algoritmo de otimização RPROP (*Resilient back-Propagation*) foram treinadas variando o número de neurônios de 1 a 20. Para a ELM, o número de neurônios foi avaliado de 5 em 5 até 50. Nas 16 regiões da base de dados para as redes MLP foram treinadas 80 mil redes, utilizando o número de neurônios da camada oculta utilizado pela colaboração ATLAS, e 800 mil para ELM, visto que essa não é aplicada e o número que alcançasse desempenho similar ou superior era desconhecido. Na Tabela II são exibidos os números de neurônios utilizados em cada rede especialista que resultou em máximo SP.

TABELA II: Número de neurônios na camada oculta para melhores índices SP em cada região nas técnicas ELM e MLP.

(E_T, η)	(0,0)	(0,1)	(0,2)	(0,3)	(1,0)	(1,1)	(1,2)	(1,3)
nNeu _{MLP}	7	16	7	14	7	20	11	17
nNeu _{ELM}	40	45	35	40	35	50	50	30
(E_T, η)	(2,0)	(2,1)	(2,2)	(2,3)	(3,0)	(3,1)	(3,2)	(3,3)
nNeu _{MLP}	14	11	9	20	20	7	17	16
nNeu _{ELM}	40	45	30	45	45	40	40	50

Na Figura 4, são exibidos os melhores resultados obtidos de índice SP para cada região nas técnicas, MLP_{verm} e ELM_{preta}, nota-se, que a distribuição dos ensaios da ELM, em alguns casos resultaram em maior espalhamento entre o sorteio de melhor SP e o sorteio de menor SP, quando comparados com os resultados das redes MLP. Porém, apesar disso o desempenho obtido pela ELM em cada região foi similar ao obtido pelas redes MLP.

Nas Figura 5 e Figura 6, as curvas ROC para os melhores índices SP obtido em cada região. Nas 16 regiões (E_T, η) . Nota-se que o desempenho da ELM foi bastante similar ao do MLP, obtendo melhor resultado nas regiões (1,0), (1,1), (1,2), (2,1), (2,2), (2,3), (3,0), (3,1), (3,2) e (3,3).

A seguir, Tabela III são exibidos os tempos de treinamento utilizados nos melhores resultados exibidos da Figura 4 para cada região, nas técnicas ELM e MLP. Em todas as regiões é possível observar que a ELM é treinada em menor tempo, de 2 a 20 vezes mais rápido, quando comparada com o tempo de treinamento das redes MLP. Mesmo utilizando número de neurônios na camada oculta superior, em algumas regiões quase 6 vezes, o tempo utilizado para a ELM ainda foi inferior.

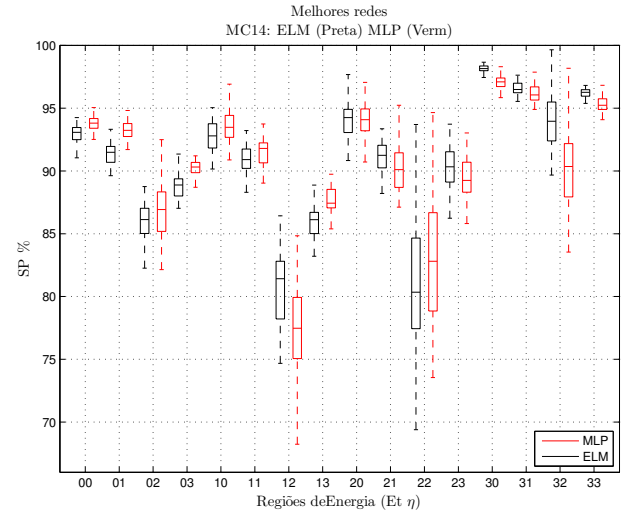


Fig. 4: Boxplot ELM \times MLP para cada região.

VI. CONCLUSÕES

A eficiência da detecção *online* de eventos de interesse é muito importante para a adequada operação de experimentos de física de altas energias. Para o ATLAS, um dos detectores do LHC, a identificação de elétrons é fundamental para possibilitar o estudo dos fenômenos físicos de interesse. Um dos discriminadores atualmente utilizados para a detecção de elétrons no ATLAS, o Neural Ringer, utiliza uma rede neural MLP alimentada por informações do sistema de calorimetria. Uma das limitações do Neural Ringer está relacionada com o longo tempo de treinamento decorrente da utilização de grandes bases de dados simulados e experimentais. Neste trabalho é proposta a utilização de máquinas de aprendizado extremo para substituir as redes MLP no Neural Ringer. Os resultados obtidos indicam que é possível reduzir o tempo de treinamento (pelo menos em 2 vezes) e manter desempenho de discriminação equivalente.

AGRADECIMENTOS

Os autores agradecem à FAPESB, FAPERJ e RENAFAP pelo apoio financeiro e aos colegas da Colaboração ATLAS pelas relevantes contribuições para esse trabalho.

REFERÊNCIAS

- [1] L. Evans and P. Bryant, "LHC machine," *Journal Of Instrumentation*, vol. 3, no. Ss08001, pp. 1–158, 2008.
- [2] CERN-Colaboration, "Cern homepage," Disponível Em :<<http://www.cern.ch/>>, 2016.
- [3] M. A. Moreira, "O modelo padrão da física de partículas," *Revista Brasileira De Ensino De Física*, vol. 31, no. 1, pp. 1306.01 – 1306.11, 2009.
- [4] A. dos Anjos, R. C. Torres, J. M. Seixas, B. Ferreira, and T. C. Xavier, "Neural triggering system operating on high resolution calorimetry information," *Nuclear Instruments and Methods In Physics Research A*, vol. 559, pp. 134–138, Apr 2006.
- [5] E. E. P. de Souza, E. F. Simas Filho, P. C. M. A. Farias, and J. M. Seixas, "Sistema neural para seleção online de eventos com pré-processamento através de dwt," in *Anais Do XX Congresso Brasileiro de Automática. CBA - Congresso Brasileiro De Automática*, 2014, pp. 1072 – 1079.
- [6] J. J. M. Pimenta, L. F. B. Belussi, E. R. T. Natti, and P. L. Natti, "O bóson de higgs," *Revista Brasileira de Ensino de Física*, vol. 35, pp. 1 – 14, 06 2013.

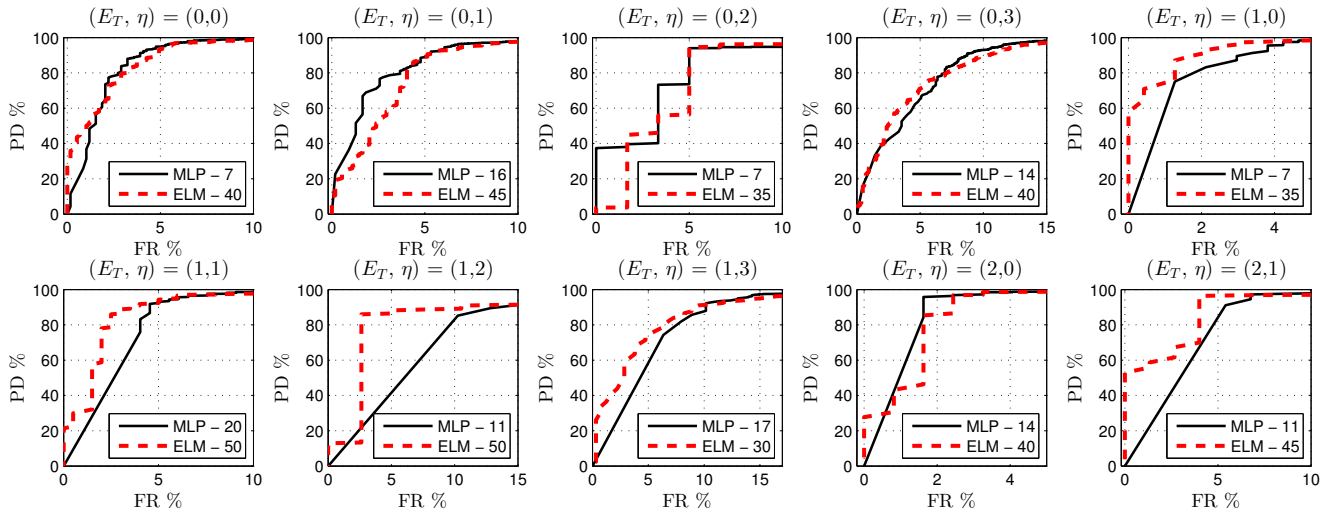


Fig. 5: ROC para índices SP, regiões (0,0) ... (2,1).

 TABELA III: Tempo de treinamento em segundos, para os melhores resultados, ELM \times MLP, em cada região.

Região	(0,0)	(0,1)	(0,2)	(0,3)	(1,0)	(1,1)	(1,2)	(1,3)
t_{ELM}	0,100 \pm 0,008	0,110 \pm 0,009	0,020 \pm 0,007	1,188 \pm 0,090	0,280 \pm 0,027	0,230 \pm 0,016	0,359 \pm 0,055	1,844 \pm 0,135
t_{MLP}	2,240 \pm 0,154	4,890 \pm 0,210	0,410 \pm 0,064	7,610 \pm 0,221	5,760 \pm 0,178	12,230 \pm 0,796	1,230 \pm 0,148	15,030 \pm 0,485
Região	(2,0)	(2,1)	(2,2)	(2,3)	(3,0)	(3,1)	(3,2)	(3,3)
t_{ELM}	1,750 \pm 0,195	1,125 \pm 0,112	0,313 \pm 0,039	0,380 \pm 0,018	1,078 \pm 0,106	0,938 \pm 0,091	0,188 \pm 0,037	1,141 \pm 0,119
t_{MLP}	17,540 \pm 1,144	3,800 \pm 0,389	0,610 \pm 0,144	17,290 \pm 1,683	10,460 \pm 0,127	2,020 \pm 0,124	1,350 \pm 0,230	9,040 \pm 0,105

- [7] S. Haykin, *Neural Networks and Learning Machines*. Nova Jersey, Estados Unidos: Prentice Hall, 2008.
- [8] A. P. Braga, A. C. P. de L. de Carvalho, and T. B. Ludermir, *Redes Neurais Artificiais: Teoria E Aplicações*, 1st ed. RJ: LTC, 2007.
- [9] E. Simas Filho, M. M. Silva Jr., P. C. Farias, M. C. Albuquerque, I. C. Silva, and C. T. Farias, "Flexible decision support system for ultrasound evaluation of fiber-metal laminates implemented in a DSP," *NDT & E International*, vol. 79, pp. 38 – 45, 2016.
- [10] D. Bibicu and L. Moraru, "Cardiac cycle phase estimation in 2-d echocardiographic images using an artificial neural network," *IEEE Transactions on Biomedical Engineering*, vol. 60, no. 5, pp. 1273–1279, May 2013.
- [11] G.-B. Huang, Q.-Y. Zhu, and C.-K. Siew, "Extreme learning machine: Theory and applications," *Neurocomputing*, vol. 70, no. 1/3, pp. 489 – 501, 2006, neural NetworksSelected Papers From The 7th Brazilian Symposium On Neural Networks (sbrn '04)7th Brazilian Symposium On Neural Networks.
- [12] M. Termenon, M. Graña, A. Savio, A. Akusok, Y. Miche, K.-M. Björk, and A. Lendasse, "Brain MRI morphological patterns extraction tool based on extreme learning machine and majority vote classification," *Neurocomputing*, vol. 174, Part A, pp. 344 – 351, 2016.
- [13] P. Horata, S. Chiewchanwattana, and K. Sunat, "Robust extreme learning machine," *Neurocomputing*, vol. 102, pp. 31 – 44, 2013, advances in Extreme Learning Machines (ELM 2011).
- [14] G. A. Barreto and A. L. B. Barros, "A robust extreme learning machine for pattern classification with outliers," *Neurocomputing*, vol. 176, pp. 3 – 13, 2016, recent Advancements in Hybrid Artificial Intelligence Systems and its Application to Real-World ProblemsSelected papers from the HAIS 2013 conference.
- [15] H. Chen, J. Peng, Y. Zhou, L. Li, and Z. Pan, "Extreme learning machine for ranking: Generalization analysis and applications," *Neural Networks*, vol. 53, pp. 119 – 126, 2014.
- [16] B. Qu, B. Lang, J. Liang, A. Qin, and O. Crisalle, "Two-hidden-layer extreme learning machine for regression and classification," *Neurocomputing*, vol. 175, Part A, pp. 826 – 834, 2016.
- [17] K. Facelli, A. C. Lorena, J. Gama, and A. Carvalho, *Inteligência Artificial Uma Abordagem De Aprendizado De Máquina*. Rio De Janeiro: LTC, 2011, vol. 1.

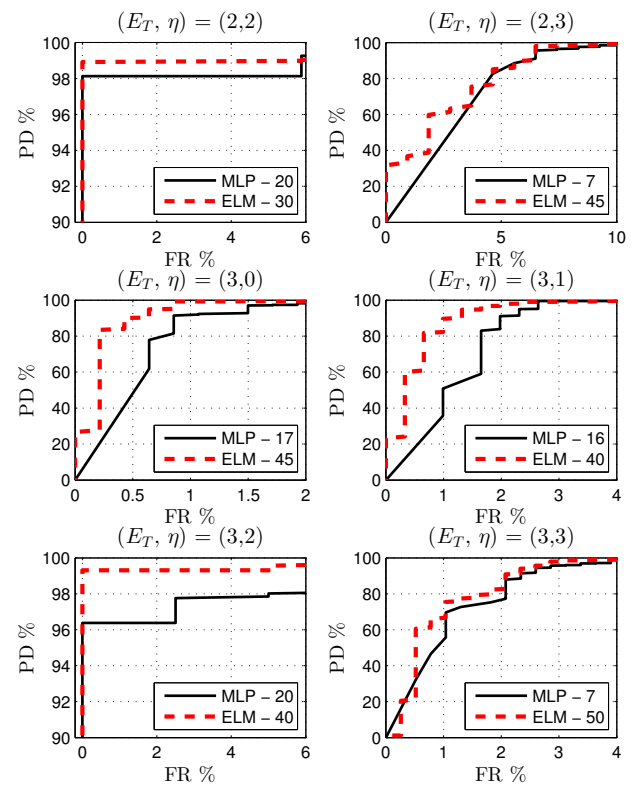


Fig. 6: ROCs para índices SP, regiões (2,2) ... (3,3).