

MÁQUINAS DE APRENDIZADO EXTREMO E REDES COM ESTADOS DE ECO PARA CLASSIFICAÇÃO *ONLINE* DE EVENTOS NO DETECTOR ATLAS

MARTON S. DOS SANTOS*, EDUARDO F. SIMAS FILHO*, PAULO C. M. A. FARIAS*, JOSÉ M. DE SEIXAS†

*Laboratório de Sistemas Digitais - PPGEE
Universidade Federal da Bahia
Salvador, Brasil

†Laboratório de Processamento de Sinais - COPPE/Poli
Universidade Federal do Rio de Janeiro
Rio de Janeiro, Brasil

Emails: mssandes@gmail.com, eduardo.simas@ufba.br, paulo.farias@ufba.br, seixas@lps.ufrj.br

Abstract— ATLAS is a detector of LHC particle accelerator, and its cylindrical structures, which comprises several layers of sensors, was designed to characterize the events of interest generated in the particle bunches collisions. ATLAS energy measurement system (calorimeter) comprises more than 180,000 sensors, which provide valuable information for online particle selection. In this context, Neural Ringer is an electromagnetic particles (electrons and photons) discriminator which uses a multi-layer perceptron (MLP) neural network for particle classification using as inputs the ring-formatted energy deposition profile measured at the calorimeters. This work proposes the use of low computational requirements learning machines to replace the MLP in the Neural Ringer discriminator. Extreme Learning Machines (ELM) and Echo State Networks (ESN) were used and the obtained results indicate that it is possible to achieve similar discrimination efficiency and a considerable reduction in system development time.

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

Resumo— O ATLAS é um dos detectores do acelerador de partículas LHC. Com sua estrutura cilíndrica, o ATLAS é capaz de caracterizar os fenômenos de interesse que ocorrem após as colisões dos feixes de partículas. O sistema de medição de energia (calorímetro) do ATLAS é composto por mais de 180.000 sensores e fornece informações importantes para a seleção *online* dos eventos de interesse. Neste contexto, o *Neural Ringer* é um discriminador de partículas eletromagnéticas (elétrons e fótons) que opera no sistema *online* de filtragem (*trigger*) do ATLAS. O referido discriminador de partículas utiliza uma rede neural tipo Perceptron de múltiplas camadas (MLP - *Multi-layer Perceptron*) para realizar a classificação das partículas, a partir do perfil de deposição de energia medido nos calorímetros e formatado em anéis. Neste trabalho, é proposta a substituição dos classificadores MLP do *Neural Ringer* por máquinas de aprendizado, com reduzido custo computacional de treinamento. Utilizam-se máquinas de aprendizado extremo (ELM - *Extreme Learning Machines*) e redes com estados de eco (ESN - *Echo State Networks*) e os resultados apontam que é possível obter eficiência de classificação semelhante ao sistema original com uma considerável redução do tempo de treinamento.

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

1 Introdução

A compreensão a respeito da constituição fundamental da matéria obteve evolução significativa nos últimos anos, devido à comprovações resultantes de experimentos de física de altas energias. O Grande Colisor de Hadrons (*Large Hadron Collider* - LHC) (Evans e Bryant, 2008) é o maior acelerador de partículas em operação atualmente e está situado no Centro Europeu para Pesquisa Nuclear (CERN) (CERN-Colaboration, 2016). O LHC foi construído com o objetivo de analisar a estrutura fundamental da matéria, investigar as propriedades das partículas fundamentais propostas pelo Modelo Padrão (*Standard Model*) (Moreira, 2009) e também buscar por fenômenos desconhecidos.

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*), conforme mostrado na Figura 1. O túnel do acelerador tem cerca de 27 km de comprimento e as colisões podem ocorrer numa taxa de até 40×10^6 vezes por segundo (Evans e Bryant, 2008).

Para os próximos anos, existe uma agenda de atualizações que envolvem elevação gradativa dos níveis de energia e da quantidade de partículas nas colisões, além de aprimoramentos nos sistemas de seleção, tratamento e armazenamento da informação de interesse (Heuer, 2013).

O ATLAS, ilustrado na Figura 2, é um detector de propósito geral, ou seja, possui capacidade de detectar diversos tipos de partículas. É composto pelos seguintes subdetectores: 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

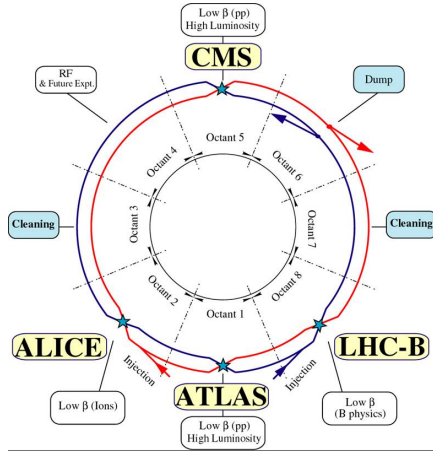


Figura 1: Ilustração da estrutura do LHC e seus detectores. Fonte: CERN

altamente segmentada e da alta taxa de colisões, é produzido um volume de dados da ordem de 60 TB/s, o que requer a detecção (ou *trigger*) *online* das assinaturas de interesse.

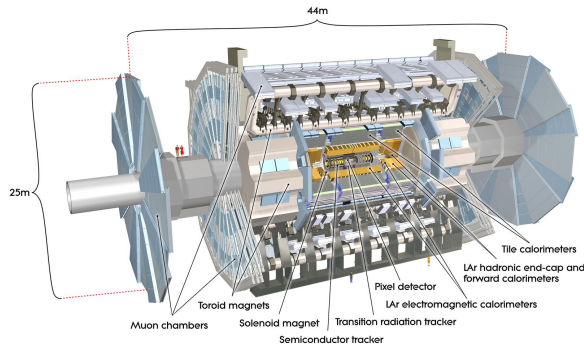


Figura 2: Ilustração do Detector ATLAS, com destaque para seus sub-detectores. Fonte: CERN

Este trabalho está organizado conforme descrito a seguir. As principais características do discriminador *Neural Ringer* são abordadas na Seção 2. Na Seção 3, são apresentados os aspectos teóricos, relacionados às técnicas de processamento de sinais utilizadas. A Seção 4 apresenta a metodologia utilizada no trabalho. Os resultados obtidos com bases de dados simuladas e experimentais são discutidos na Seção 5. As conclusões são apresentadas na Seção 6.

2 Sistema de Filtragem

O sistema de seleção ou filtragem (*trigger*) *online* do ATLAS (Martínez e ATLAS Collaboration, 2016) é responsável pela seleção dos eventos interessantes para o experimento e, também, pela redução do ruído de fundo (assinaturas não relevantes), produzido nas colisões. Sua estrutura é composta por três níveis, o L1 (*Level 1*), o RoIB

(*Region of Interest Builder*) e o terceiro é o HLT (*High Level Trigger*).

O L1 é desenvolvido em *hardware* dedicado e deve reduzir a taxa de eventos na ordem de 40 MHz para 100 kHz; no RoIB toda a informação vinda do L1 é utilizada pelo nível de *software*, criando as regiões de interesse, RoI, na qual se opera o *Neural Ringer*, e em seguida, é enviada ao terceiro nível, HLT, que reduz a taxa de eventos para ~ 1 kHz, ver Figura 3 (Galster, 2015; Kilby e ATLAS Collaboration, 2016; Martínez e ATLAS Collaboration, 2016; Vazquez, 2016).

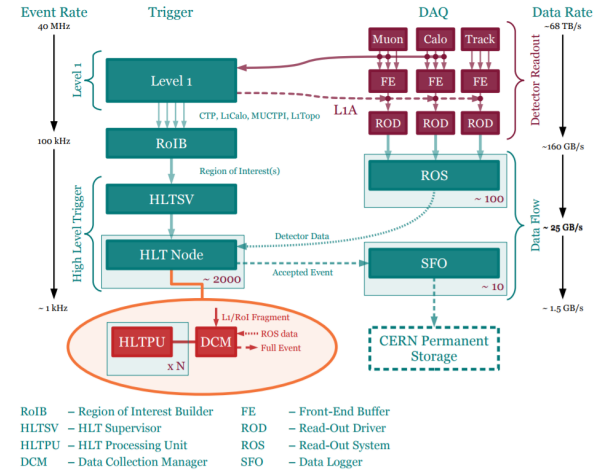


Figura 3: Esquema do *trigger online* do ATLAS atual. Fonte: (Galster, 2015)

A identificação de elétrons é muito importante para o desempenho do detector, pois a busca por assinaturas de interesse¹ pode estar relacionada aos elétrons, e, para isso, são utilizadas informações dos calorímetros. Um dos discriminadores utilizados atualmente, no ATLAS, para a identificação online de elétrons, é o *Neural Ringer* (dos Anjos et al., 2006), no qual o perfil de deposição de energia é utilizado como entrada para uma rede neural tipo *perceptron* de múltiplas camadas, que opera como classificador no *trigger online*.

Para a filtragem *online*, o ATLAS utiliza o discriminador *Neural Ringer*, que organiza a região de interesse em anéis concêntricos de deposição de energia por camada do calorímetro, conforme ilustrado na Figura 4. A energia medida nos sensores de cada anel é somada, e essa informação é utilizada para alimentar um classificador neural do tipo *perceptron* de múltiplas camadas (*multi-layer perceptron* - MLP). Na Tabela 1, é possível visualizar a organização do número de anéis por camadas.

Um desafio associado ao *Neural Ringer* é o elevado tempo de treinamento do sistema, que precisa ser repetido para as diferentes configurações

¹A descoberta do Bóson de Higgs (2012) ocorreu com o auxílio de informações contidas em canais de elétrons, múons e fótons isolados (Freund et al., 2016).

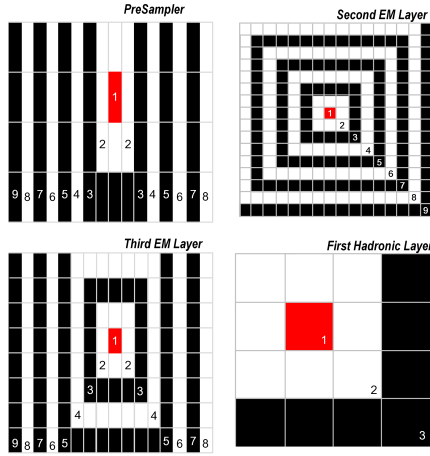


Figura 4: Diagrama dos anéis do calorímetro nas RoIs. Fonte: Colaboração ATLAS.

Tabela 1: Número de anéis por camada. PS - *Presampler*, E1 – E3 - Camadas Eletromagnéticas, H0 – H2 - Camadas Hadrônicas.

Camadas	PS	E1	E2	E3	H0	H1	H2
Anéis	08	64	08	08	04	04	04

de operação do detector. Este trabalho propõe a avaliação de dois outros classificadores, a saber, a máquina de aprendizado extremo (*Extreme Learning Machine* - ELM) e a rede com estados de eco (*Echo State Networks* - ESN), em substituição ao classificador MLP no discriminador *Neural Ringger*, com o objetivo de reduzir o tempo de treinamento. As redes ELM apresentam estrutura semelhante à de uma rede MLP, porém seu tempo de treinamento é, comparativamente, muito menor. Já as redes ESN apresentam uma estrutura baseada em redes recorrentes, explorando sua capacidade de generalização, devido às conexões de realimentação, apresentando tempo de treinamento significativamente inferior ao das redes MLP.

3 Técnicas de Processamento de Sinais Utilizadas

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

3.1 Redes Neurais Artificiais

Sistemas baseados em RNA vêm sendo utilizados com sucesso em diferentes aplicações. Em de Souza et al. (2014) e Freund et al. (2016), uma rede neural foi utilizada para realizar a detecção eficiente de elétrons no segundo nível de filtragem *online* de eventos do detector ATLAS.

Em Faria et al. (2017), foi projetado um filtro FIR, baseado em rede neural desenvolvida em hardware dedicado (FPGA), utilizada para estimação da energia deposita no calorímetro de telhas do ATLAS, o TileCal.

3.2 Máquinas de aprendizado extremo

As máquinas de aprendizado extremo (*Extreme Learning Machines* - ELM) (Huang et al., 2006) utilizam 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 1.

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

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.

A Equação 1 pode ser reescrita na forma matricial como $\mathbf{H}\boldsymbol{\beta} = \mathbf{Y}$, sendo,

$$\mathbf{H} = \begin{pmatrix} \phi(w_1 x_1 + b_1) & \dots & \phi(w_N x_1 + b_N) \\ \vdots & \ddots & \vdots \\ \phi(w_1 x_M + b_1) & \dots & \phi(w_N x_M + b_N) \end{pmatrix}, \quad (2)$$

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

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

A ELM vem sendo utilizada em diferentes aplicações, como, por exemplo, em Termenon et al. (2016), com o objetivo de desenvolver uma ferramenta de auxílio na extração de características de imagens de ressonância magnética no diagnóstico de mal de Alzheimer. Foi provado que uma ELM, assim como uma rede MLP, é um aproximador universal. Entretanto, em alguns casos, a ELM requer um número maior de neurônios na camada oculta para resolver, com desempenho equivalente, o mesmo problema. Em Qu et al. (2016), uma estrutura com duas camadas foi avaliada e comparada em problemas de regressão e classificação, sendo observado que a estrutura se torna interessante para problemas complexos, na presença

² d_1 e d_2 são as dimensões dos sinais da entrada e saída, respectivamente.

de recursos computacionais de armazenamento limitado.

3.3 Redes com Estados de Eco

As redes com estados de eco (ESN) são redes neurais compostas por: a) uma camada de entradas; b) uma camada interna, denominada reservatório de dinâmicas (RD), constituída de neurônios organizados numa estrutura recorrente totalmente conectados, utilizando funções de ativação não-linear, conforme Figura 5, e; c) uma camada de saídas de característica linear, cujo treinamento e obtido de maneira semelhante ao que ocorre com a ELM, por meio da inversa generalizada de Moore Penrose, ou método de regressão linear dos mínimos quadrados, como se observa em ?.

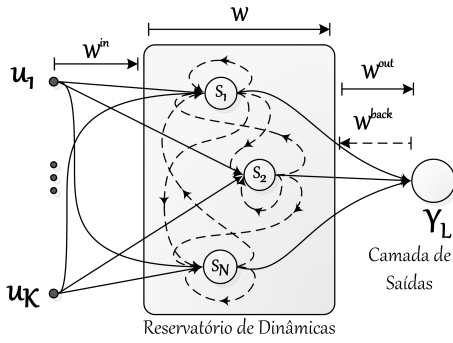


Figura 5: Diagrama de uma rede ESN que possui estados de eco.

Com base nos padrões disponíveis para o treinamento e na resposta esperada, $\mathbf{Y}$, é possível determinar os coeficientes da matriz $\mathbf{W}^{out}$, Equação 5, por meio da inversa generalizada expressa na Equação 6, a matriz de pesos do reservatório, $\mathbf{W}$ é esparsa e gerada aleatoriamente, tendo distribuição com média 0 e variância 1.

$$x(n+1) = \mathbf{f}(\mathbf{W}^{in}\mathbf{u}(n+1) + \mathbf{W}\mathbf{x}(n)) \quad (3)$$

$$y(n+1) = \mathbf{W}^{out}\mathbf{x}(n+1) \quad (4)$$

$$\mathbf{W}^{out} = \mathbf{X}^\dagger \mathbf{Y} \quad (5)$$

$$\mathbf{X}^\dagger = (\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T \quad (6)$$

Para que uma ESN possua estados de eco, é necessário que possua algumas características: $|\sigma_{max}(\mathbf{W})| < 1^3$, no qual σ é o valor singular de $\mathbf{W}$; $|\lambda_{max}(\mathbf{W})| < 1$, sendo λ o autovalor de $\mathbf{W}$, que é chamado de raio espectral da ESN (Jaeger, 2010).

Adicionalmente, pode ser acrescido o parâmetro α (*leak rate*) na Equação 3, o que resulta na Equação 7, e a escolha adequada do valor parâmetro permite a melhora no ajuste da dinâmica

do reservatório da ESN (Simeón, 2015). O valor ótimo para o parâmetro α pode ser definido empiricamente, ou por busca num conjunto de valores por uma função de otimização. No trabalho de Antonelo et al. (2008), um pequeno robô é treinado no contexto de computação de reservatórios, utilizando-se redes ESN e abordando métodos de busca do valor adequado para o parâmetro α .

$$x(n+1) = \mathbf{f}((1-\alpha)\mathbf{x}(n) + \alpha(\mathbf{W}^{in}\mathbf{u}(n+1) + \mathbf{W}\mathbf{x}(n))) \quad (7)$$

A ESN tem sido utilizada em problemas de classificação de padrões, como em Wen et al. (2015), em que um conjunto de redes ESN (*Ensemble Convolutional Echo State Networks* - EC-ESN) foi utilizado para o reconhecimento de padrões de expressões faciais, utilizando-se imagens de duas bases de dados sem nenhuma técnica de extração de características. As imagens foram apresentadas às redes SVM, SRC⁴, Softmax, ESN e EC-ESN. Os resultados indicaram a que a ESN tem capacidade de separação de classes em problemas de reconhecimento de expressões faciais. Já em Trentin et al. (2015), uma variação da ESN, π -ESN (*Probabilistic ESN*) foi aplicada num problema de reconhecimento de cinco expressões de fala de mulheres.

4 Método

Nos treinos das redes com ELM, ESN e MLP, foi utilizada uma base de dados simulados por meio da técnica Monte Carlo, disponibilizada pela colaboração ATLAS. Nesta base, os dados estão organizados em 16 regiões, ver Tabela 2, sendo 4 intervalos para o nível de energia transversa (E_T) e 4 intervalos para a posição no interior do detector, (η - pseudo-rapidez⁵), os quais foram representados neste trabalho em 16 pares (E_T, η), desde (0,0) até (3,3), um para cada região e energia.

A organização da base de dados 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 passam. Outro ponto relevante é a interação entre as partículas eletromagnéticas e hadrônicas com o detector, que dependem do nível de energia das colisões, justificando a utilização de um conjunto de classificadores para diferentes posições e níveis de energia.

Para o treinamento das redes neurais, a base de dados foi igualmente dividida em 10 subconjuntos. Sendo, 60% utilizados para treino e 40% para teste. Em cada conjunto de treino e teste são realizados 50 sorteios disjuntos e cada sorteio

³Condição demonstrada quando não há realimentação da saída para o RD com uma rede, utilizando como função de ativação a tangente hiperbólica.

⁴*Sparse representation classifier*

⁵definida em função do ângulo de incidência (θ) das partículas no interior do experimento $\eta = -\ln(\tan \frac{\theta}{2})$.

é treinado 100 vezes, com o objetivo de evitar mínimos locais. Para a ELM e ESN, que possuem etapas com pesos gerados aleatoriamente, esse método contribui para a obtenção da melhor rede, reduzindo tendências causadas pelo processo aleatório.

Tabela 2: Segmentação da 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 desempenho foi avaliado com a curva ROC (Marzban, 2004) e o índice SP (Torres et al., 2009), definido na Equação 8:

$$SP = \sqrt{\frac{(Ef_e + Ef_j)}{2}} \times \sqrt{Ef_e \times Ef_j} \quad (8)$$

na qual $Ef_e = PD$ 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).

Em todas as regiões da base de dados, os pesos da camada oculta da ELM foram gerados aleatoriamente com distribuição normal, porque ensaios preliminares resultaram em melhores índices SP ao se gerar pesos que possuíssem distribuição normal para a ELM. Para a ESN, os pesos do reservatório de dinâmicas foram gerados também com distribuição normal (Farkaš et al., 2016), 0,3 de grau de esparsidade, raio espectral igual a 0,8 e não foi utilizada a taxa de aprendizagem nos ensaios.

Para os ensaios com o MLP foram utilizados os seguintes número de neurônios na camada oculta: {7, 16, 7, 14, 7, 20, 11, 17, 14, 11, 9, 20, 20, 7, 17, 16}. Para a ELM foram testados desde 5 até 100 e para ESN testaram-se desde 5 até 50. As redes que compõem os resultados são as que obtiveram o melhor índice SP para cada região.

Todas as redes foram treinadas, na nuvem, em máquinas virtuais, com processadores de arquitetura Intel[®] Xeon E5 v4 2,2 GHz com 8 núcleos e 24 GB de RAM.

5 Resultados

Na Figura 6, são exibidas as *boxplot* de cada um dos classificadores, MLP, ELM e ESN, nessa ordem, para o melhor dos 100 treinos realizados em cada uma das 16 regiões.

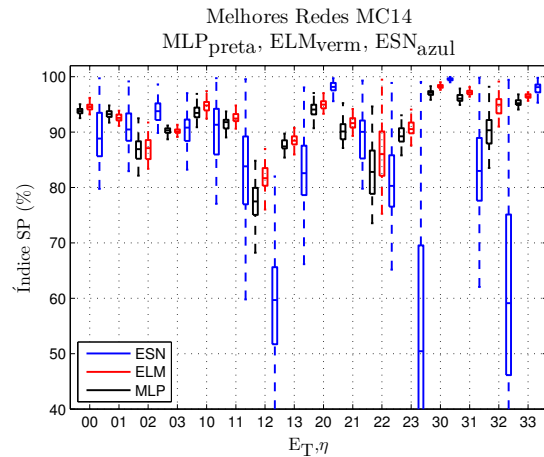


Figura 6: Boxplot ELM $\times$ ESN $\times$ MLP para cada região.

Na cor azul, observam-se os resultados obtidos para as redes ESN, e é possível visualizar que o classificador, apesar de alcançar valores semelhantes ao do MLP, no que se refere ao máximo índice SP obtido, apresenta uma grande dispersão entre o melhor e o pior índice obtido, em algumas regiões a dispersão é superior a 50%, como nas regiões (2,0), (2,3) e (3,2). Em três regiões, o desempenho alcançado pela ESN foi equivalente e com pequeno espalhamento em relação ao MLP, (2,0), (3,0) e (3,3).

Já em vermelho, observam-se os resultados obtidos com a ELM. Nestes, nota-se uma similaridade, tanto nos máximos SPs alcançados com o MLP, quanto no grau de dispersão entre o valor máximo e mínimo.

Nas Figuras 7 e 8, são apresentadas as curvas ROC para os melhores índices SP dos três classificadores. Das 16 regiões (E_T, η), o desempenho da ELM foi equivalente ao desempenho do MLP, sendo levemente inferior nas regiões (0,1), (0,2), (0,3) e (2,2). Destaca-se ainda a baixa taxa de falso alarme apresentado, exceto na região (2,2).

Em relação ao desempenho da ESN, apesar do grande espalhamento visto na Figura 6, as curvas ROC indicam que a capacidade de separação de classes da ESN é bastante significativa quando comparada ao MLP e mesmo a ELM, sendo superior em quase todas as regiões, exceção feita à região (2,2), a qual a ESN teve o pior desempenho entre os classificadores.

Tendo em consideração os parâmetros de ajuste disponíveis para a configuração de uma rede ESN, é possível que o desempenho da técnica possa ser melhorado, com redução do espalhamento apresentado, utilizando-se parâmetros de otimização, ou estruturas híbridas de outras redes com ESN.

Para uma avaliação da qualidade da separação realizada pelos classificadores avaliados, ELM e ESN, é apresentada a Tabela 3, na qual o resultado

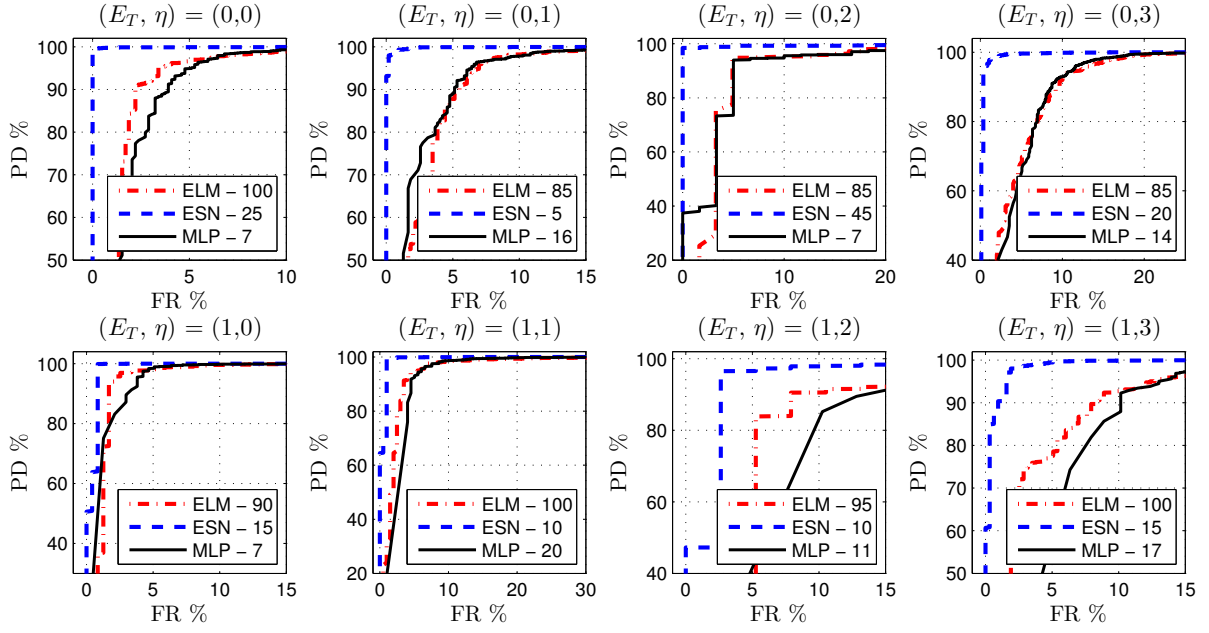


Figura 7: ROCs para os melhores índices SP, regiões (0,0) ... (1,3).

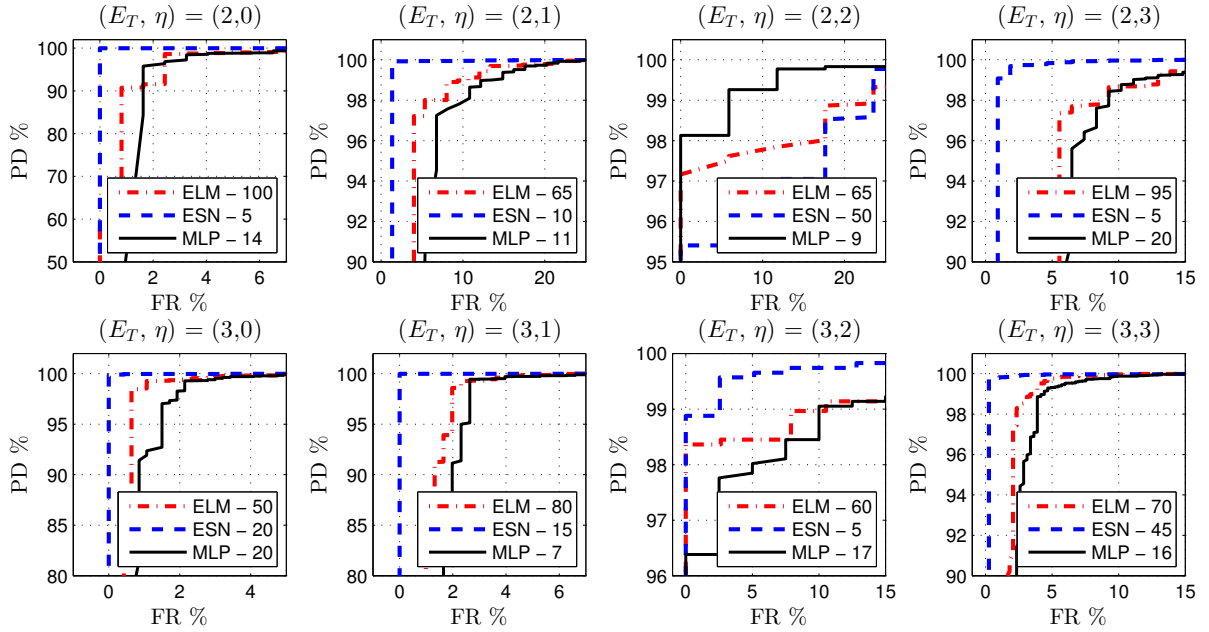


Figura 8: ROCs para os melhores índices SP, regiões (2,0) ... (3,3).

de desempenho da ELM e ESN são comparados com os resultados obtidos com o *Neural Ringer* (MLP) em cada uma das regiões.

Os resultados observados nas Figuras 7 e 8, quanto à similaridade de desempenho da ELM em relação ao MLP, são reforçados na Tabela 3. Em todas as regiões, o desempenho da ELM foi superior, restando analisar o grau de incerteza para os valores alcançados, que também ficaram muito próximos, sendo a maior incerteza registrada na região (2,2).

Analisando os resultados para a ESN,

percebe-se que o desempenho máximo SP foi superior tanto ao *Ringer*, quanto à ELM. Porém, somente em quatro regiões, (0,2), (1,0), (3,0) e (3,3), a incerteza alcançada fica próxima à incerteza alcançada pelo *Neural Ringer*. Em outras, (1,3), (2,1), (3,1), (3,2), alcançam valores proibitivos, com a incerteza superior em pelo menos 10 vezes, mesmo com elevadas taxas de acerto.

Também foi possível observar que o aumento do número de neurônios do reservatório de dinâmicas não produziu alteração significativa no valor do índice SP, mas reduziu a incerteza para os

valores obtidos. Já para a ELM, o aumento no número de neurônios reduziu tanto a dispersão entre os valores máximos e mínimos para o índice SP, quanto a incerteza.

Na Tabela 4, são exibidos os tempos de treinamento para cada um dos classificadores, em cada uma das regiões, na mesma plataforma computacional. Na comparação entre os três classificadores, observa-se que para a ELM o tempo médio total de treinamento das redes foi de 11,1 s, para as redes ESN foi de 21,1 s, enquanto que para as redes com MLP, 111,51 s.

6 Conclusões

A eficiência da detecção *online* de eventos de interesse é muito importante para a adequada operação de experimentos de física de altas energias. Considerando-se a agenda de atualizações do ATLAS, a raridade dos eventos procurados, e os níveis de energia envolvidos, fazem com que a identificação de elétrons, fundamental para o estudo dos fenômenos físicos de interesse, torne-se um desafio para os discriminadores atualmente utilizados.

Dentre os discriminadores está o *Neural Ring*, que utiliza uma rede neural MLP alimentada por informações do sistema de calorimetria. Uma das limitações do *Neural Ring* está relacionada com o longo tempo de treinamento decorrente da utilização de grandes bases de dados simulados e experimentais.

Neste trabalho, foi proposta a utilização de máquinas de aprendizado extremo e redes em estado de eco para substituir as redes MLP no *Neural Ring*. Os resultados obtidos indicam que a ELM consegue alcançar desempenho equivalente ao MLP e tempo de treinamento inferior. Já com a ESN é possível atingir taxas de acerto elevadas quando comparada com o MLP em tempo de treinamento reduzido, porém os resultados registraram uma incerteza superior à alcançada pelo MLP operando no *Neural Ring*. Mas a ESN possui parâmetros ajustáveis, de tal modo que seja possível manter o desempenho alcançado nesse trabalho com redução da incerteza, o que habilitaria a ESN como uma alternativa de classificador a ser utilizado pela Colaboração ATLAS.

Agradecimentos

Os autores agradecem à FAPESB, FAPERJ, CNPq, CAPES e RENAFAE (MCTI) pelo apoio financeiro e aos colegas da Colaboração ATLAS pelas relevantes contribuições para esse trabalho.

Referências

Antonelo, E., Schrauwen, B. e Stroobandt, D. (2008). Event detection and localization

for small mobile robots using reservoir computing, *Neural Networks* **21**(6): 862 – 871. Computational and Biological Inspired Neural Networks, selected papers from ICANN 2007.

CERN-Colaboration (2016). Cern homepage, Disponível Em :<<http://www.cern.ch/>>.

de Souza, E. E. P., SIMAS FILHO, E. F., Farias, P. C. M. A. e de Seixas, J. M. (2014). Sistema neural para seleção online de eventos com pré-processamento através de DWT, *Anais Do XX Congresso Brasileiro de Automática*, CBA - Congresso Brasileiro De Automática, pp. 1072 – 1079.

dos Anjos, A., C., T. R., de Seixas, J. M., B.C., F. e C., X. T. (2006). Neural triggering system operating on high resolution calorimetry information, *Nuclear Instruments and Methods In Physics Research A* **559**: 134–138.

Evans, L. e Bryant, P. (2008). LHC machine, *Journal Of Instrumentation* **3**(Ss08001): 1–158.

Faria, M. H. M., Filho, L. M. A., da S. Duarte, J. P. B. e de Seixas, J. M. (2017). Redes neurais para filtragem inversa com aplicação em calorímetros operando a alta taxa de eventos, *XXXV Simpósio Brasileiro de Telecomunicações e Processamento de Sinais (SBrT 2017) (SBrT 2017)*, São Pedro, Brazil, pp. 403–407.

Farkaš, I., Bosák, R. e Gergeľ, P. (2016). Computational analysis of memory capacity in echo state networks, *Neural Networks* **83**: 109 – 120.

Freund, W. S., da Fonseca Pinto, J. V. e de Seixas, J. M. (2016). Redes neurais aplicadas na identificação de partículas em um detector finamente segmentado, *Anais do XXI Congresso Brasileiro de Automática*, CBA - Congresso Brasileiro de Automática, pp. 1072 – 1079.

Galster, G. (2015). *ATLAS trigger: Preparations for run ii*, Dissertação de mestrado, University of Copenhagen.

Heuer, R. D. (2013). General meeting january 2013.

Huang, G.-B., Zhu, Q.-Y. e Siew, C.-K. (2006). Extreme learning machine: Theory and applications, *Neurocomputing* **70**(1?3): 489 – 501. Neural Networksselected Papers From The 7th Brazilian Symposium On Neural Networks (sbrn '04)7th Brazilian Symposium On Neural Networks.

Tabela 3: Índice SP para os melhores resultados, ELM e ESN comparados com os valores obtidos no *Neural Ringer* com o MLP, em cada região.

Região	(0,0)	(0,1)	(0,2)	(0,3)	(1,0)	(1,1)	(1,2)	(1,3)
ELM	96,187 ± 0,628	95,105 ± 0,742	95,052 ± 2,350	91,982 ± 0,592	97,397 ± 1,024	95,882 ± 1,004	91,436 ± 2,990	91,995 ± 1,175
<i>Ringer</i>	95,229 ± 0,410	93,926 ± 0,555	90,661 ± 1,954	91,079 ± 0,385	95,044 ± 0,829	93,780 ± 0,938	88,889 ± 2,728	89,970 ± 0,823
ESN	99,725 ± 4,782	99,155 ± 3,526	98,608 ± 1,953	98,377 ± 3,388	99,730 ± 0,868	99,305 ± 5,010	98,892 ± 7,455	99,077 ± 20,910

Região	(2,0)	(2,1)	(2,2)	(2,3)	(3,0)	(3,1)	(3,2)	(3,3)
ELM	98,184 ± 0,999	96,888 ± 1,452	99,460 ± 5,169	96,211 ± 1,598	99,271 ± 0,365	98,557 ± 0,505	99,914 ± 1,966	98,289 ± 0,437
<i>Ringer</i>	95,518 ± 1,096	93,426 ± 1,615	91,050 ± 3,021	92,508 ± 1,084	99,045 ± 0,169	97,718 ± 0,454	97,197 ± 1,383	97,322 ± 0,406
ESN	99,731 ± 6,018	99,555 ± 10,120	96,978 ± 11,245	98,073 ± 6,657	99,960 ± 0,345	99,827 ± 13,416	99,439 ± 17,398	99,809 ± 1,123

Tabela 4: Tempo de treinamento em segundos, para os melhores resultados, ELM × ESN × 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 ± 0,008	0,110 ± 0,009	0,020 ± 0,007	1,188 ± 0,090	0,280 ± 0,027	0,230 ± 0,016	0,359 ± 0,055	1,844 ± 0,135
t_{ESN}	1,246 ± 0,080	0,728 ± 0,057	0,170 ± 0,012	1,279 ± 0,100	2,302 ± 0,582	1,740 ± 0,150	0,413 ± 0,039	2,280 ± 0,637
t_{MLP}	2,240 ± 0,154	4,890 ± 0,210	0,410 ± 0,064	7,610 ± 0,221	5,760 ± 0,178	12,230 ± 0,796	1,230 ± 0,148	15,030 ± 0,485

Região	(2,0)	(2,1)	(2,2)	(2,3)	(3,0)	(3,1)	(3,2)	(3,3)
t_{ELM}	1,750 ± 0,195	1,125 ± 0,112	0,313 ± 0,039	0,380 ± 0,018	1,078 ± 0,106	0,938 ± 0,091	0,188 ± 0,037	1,141 ± 0,119
t_{ESN}	1,810 ± 0,507	1,673 ± 0,120	0,361 ± 0,026	2,423 ± 0,241	1,683 ± 0,091	1,105 ± 0,101	0,187 ± 0,011	1,700 ± 0,127
t_{MLP}	17,540 ± 1,144	3,800 ± 0,389	0,610 ± 0,144	17,290 ± 1,683	10,460 ± 0,127	2,020 ± 0,124	1,350 ± 0,230	9,040 ± 0,105

- Jaeger, H. (2010). The “echo state” approach to analysing and training recurrent neural networks - with an Erratum note, *Technical report*, German National Research Center for Information Technology.
- Jaeger, H. e Haas, H. (2004). Harnessing nonlinearity: Predicting chaotic systems and saving energy in wireless communication, *Science* **304**(5667): 78–80.
- Kilby, C. e ATLAS Collaboration (2016). The design and performance of the atlas inner detector trigger for run 2 lhc collisions at 13 tev, *Journal of Physics: Conference Series* **762**(1): 012029.
- Martínez, A. R. e ATLAS Collaboration (2016). The run-2 atlas trigger system, *Journal of Physics: Conference Series* **762**(1): 012003.
- Marzban, C. (2004). The roc curve and the area under it as performance measures, *Weather and Forecasting* **24**: 1106–1114.
- Moreira, M. A. (2009). O modelo padrão da física de partículas, *Revista Brasileira de Ensino de Física* **31**(1): 1–158.
- Qu, B., Lang, B., Liang, J., Qin, A. e Crisalle, O. (2016). Two-hidden-layer extreme learning machine for regression and classification, *Neurocomputing* **175**, Part A: 826 – 834.
- Simeón, E. J. A. (2015). Prognóstico de falhas baseado em redes neurais com estados de eco.
- Termenon, M., Graña, M., Savio, A., Akusok, A., Miche, Y., Björk, K.-M. e Lendasse, A. (2016). Brain MRI morphological patterns extraction tool based on extreme learning machine and majority vote classification, *Neurocomputing* **174**, Part A: 344 – 351.
- Torres, R. C., de Lima, D. E. F., de Simas Filho, E. F. e de Seixas, J. M. (2009). Neural online filtering based on preprocessed calorimeter data, *2009 IEEE Nuclear Science Symposium Conference Record (NSS/MIC)*, pp. 530–536.
- Trentin, E., Scherer, S. e Schwenker, F. (2015). Emotion recognition from speech signals via a probabilistic echo-state network, *Pattern Recognition Letters* **66**: 4 – 12. Pattern Recognition in Human Computer Interaction.
- Vazquez, W. P. (2016). The atlas data acquisition system: from run 1 to run 2, *Nuclear and Particle Physics Proceedings* **273-275**: 939 – 944. 37th International Conference on High Energy Physics (ICHEP).
- Wen, G., Li, H. e Li, D. (2015). An ensemble convolutional echo state networks for facial expression recognition, *2015 International Conference on Affective Computing and Intelligent Interaction (ACII)*, pp. 873–878.