

## COMPARAÇÃO DE REDES NEURAIS ARTIFICIAIS NA CLASSIFICAÇÃO DE DADOS

Matheus Henrique do Amaral Prates, [matheusprates@alunos.utfpr.edu.br](mailto:matheusprates@alunos.utfpr.edu.br)

Thiago Antonini Alves, [antonini@utfpr.edu.br](mailto:antonini@utfpr.edu.br)

Hugo Valadares Siqueira, [hugosiqueira@utfpr.edu.br](mailto:hugosiqueira@utfpr.edu.br)

Yara de Souza Tadano, [yaratadano@utfpr.edu.br](mailto:yaratadano@utfpr.edu.br)

Universidade Tecnológica Federal do Paraná (UTFPR), *Câmpus* Ponta Grossa, Av. Doutor Washington Subtil Chueire, 330, Jardim Carvalho, Ponta Grossa/PR, 84.017-220

**Resumo.** O conceito de inteligência artificial já existe a algum tempo, e com o avanço das pesquisas e das novas tecnologias a inteligência artificial está cada vez mais presente em nosso cotidiano. Uma das técnicas que voltou a ter grande destaque nesta década foram as Redes Neurais Artificiais (RNAs). Nesse sentido, o presente trabalho buscou comparar o desempenho de três arquiteturas diferentes de RNAs: Perceptron de Múltiplas Camadas (MLP), Máquina de Aprendizado Extremo (ELM) e Memória Longa de Curto Prazo (LSTM). A base de dados utilizada foi o ativo do mini índice (WINFUT) da bolsa de valores, coletados no site oficial da B3 - Brasil, Bolsa, Balcão. Após o tratamento dos dados, foi utilizado do Método Wrapper para definir os melhores atrasos para serem feitas as previsões. Os resultados computacionais mostraram que a rede LSTM apesar do custo computacional mais elevado, apresentou desempenho superior aos demais.

**Palavras chave:** Previsão de Séries Temporais. Redes Neurais Artificiais. Bolsa de Valores.

**Abstract.** The concept of artificial intelligence has been around for some time, and with the advancement of research and new technologies, artificial intelligence is increasingly present in our daily lives. One of the techniques that has returned to great prominence in this decade were the Artificial Neural Networks (ANNs). In this sense, the present work sought to compare the performance of three different ANN architectures: Multilayer Perceptron (MLP), Extreme Learning Machine (ELM), and Long Short-Term Memory (LSTM). The database used was the mini index asset (WINFUT) of the stock exchange, collected on the official website of B3 - Brasil, Bolsa, Balcão. After processing the data, the Wrapper Method was used to define the best delays for making predictions. The computational results showed that the LSTM network, despite the higher computational cost, performed better than the others.

**Keywords:** Forecasting Time Series. Artificial Neural Network. Stock Exchange.

### 1. INTRODUÇÃO

A cada ano que se passa, a evolução das tecnologias tem sido mais rápida do que no ano anterior. Isto tem causado grandes impactos, modificando a forma como vivemos. Tal evolução acontece em todas as áreas, sendo de grande destaque as tecnologias relacionadas a inteligência artificial, principalmente áreas de aprendizagem de máquina (*machine learning*). Apesar de terem surgido da década de 1950, recentemente vem apresentado grandes feitos pelo avanço de tecnologias como GPUs, o que permite o algoritmo ser executado de forma paralela, diminuindo seu tempo de execução (Géron, 2017).

Dentro da grande área de aprendizagem de máquina, há diversas técnicas e algoritmos que, a partir de uma base de dados, irá encontrar padrões e conseguirá replicá-los para dados ainda desconhecidos (Kelleher *et al.*, 2015). Duas grandes áreas são os algoritmos de aprendizagem não supervisionada e os supervisionados. No primeiro caso, apenas com os dados de entrada, será utilizado algum método para encontrar os padrões (Géron, 2017). Enquanto que no segundo tipo, há dados de entrada e de saída, os quais serão utilizados para ajustar os parâmetros do algoritmo (Braga *et al.*, 2007). Uma grande área da inteligência de máquina que trabalha com aprendizagem supervisionada é a das Redes Neurais Artificiais (RNAs). Dentro dessas, há também diversas arquiteturas que buscam diferente formas de encontrar os padrões presentes nos dados, tais como: Perceptron de Múltiplas Camadas (MLP – *Multilayer Perceptron*), Máquina de Aprendizado Extremo (ELM – *Extreme Learning Machine*) e a Memória Longa de Curto Prazo (LSTM – *Long Short-Term Memory*).

Como as redes neurais conseguem realizar previsões de séries históricas mesmo que sejam caóticas (Wang *et al.*, 2011), uma aplicação que obteve resultados melhores utilizando essas técnicas foi a previsão do mercado financeiro. Diversos estudos relacionados a este tema podem ser vistos nos trabalhos de Rêgo & Mussa (2008) e Roque (2009), os quais realizam a predição de índice de bolsas de valores brasileira, o IBOVESPA.

Desta forma, este trabalho tem como objetivo comparar o desempenho de 3 arquiteturas diferentes de Redes Neurais Artificiais: *Perceptron* de Múltiplas Camadas (MLP – *Multilayer Perceptron*), Máquina de Aprendizado Extremo (ELM – *Extreme Learning Machine*) e a Memória Longa de Curto Prazo (LSTM – *Long Short-Term Memory*) para previsão do ativo do mini índice (WINFUT), que é o contrato futuro derivado do índice Bovespa, o qual reflete a expectativa do preço futuro do mercado de ações.

## 2. METODOLOGIA

Nesta seção será abordado como funcionam as Redes Neurais Artificiais MLP, ELM e LSTM, bem como as etapas pré-processamento dos dados e alguns parâmetros que precisam ser definidos nos modelos. Primeiramente, antes dos dados serem treinados por uma rede neural artificial, precisam ser analisados e realizados alguns tratamentos para aumentar a chance da rede neural convergir e também aumentar a taxa de acerto. Isto significa diminuir o erro entre os dados previstos e os dados reais (Santos & Von Zuben, 1999).

Para esse trabalho, o único tratamento de dado feito foi a normalização dos mesmos. Considerando duas variáveis quaisquer, há uma grande chance de elas apresentarem escalas diferentes (Santos & Von Zuben, 1999). Por exemplo: idade e salário das pessoas. Enquanto a idade raramente passa dos 100 anos, o salário pode ser inicialmente 10x maior, ou até muito mais. Para evitar que a rede neural entenda de forma equivocada que o salário é mais relevante que a idade, todas as variáveis precisam ser normalizadas, geralmente transformando todos os valores de cada variável entre 0 e 1 (Géron, 2017).

Há duas grandes etapas em uma rede neural artificial: o treinamento, em que a rede busca encontrar quais os melhores valores de seus parâmetros livres; e o teste, em que será avaliado se a rede neural conseguiu convergir e obter um resultado satisfatório (Da Silva *et al.*, 2010). Esse é medido através do erro entre os valores previstos e os valores reais. Nesse trabalho foi utilizado a Raiz do Erro Quadrático Médio (RMSE – *Root Mean Squared Error*), definido pela Equação (1).

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - \tilde{y}_i)^2} \quad (1)$$

O treinamento de uma RNA pode ser supervisionado ou não-supervisionado. No primeiro caso, é utilizado a resposta da rede neural como referência, enquanto para o segundo caso, o modelo usa de outros métodos para reproduzir uma saída semelhante para exemplos semelhantes. As três redes neurais artificiais utilizadas nesse trabalho, apresentam treinamento supervisionado.

Uma Rede Neural *Perceptron* de múltiplas Camadas (MLP), é uma rede neural artificial densa (Géron, 2017), onde todos os neurônios de uma camada são conectados com todos os neurônios da próxima camada, partindo da camada de entrada em direção a camada de saída (Géron, 2017) – Figura 1. Como essa arquitetura de RNA não apresenta realimentação, é considerada uma rede neural *feedforward*, e há pelo menos uma camada intermediária (Da Silva *et al.*, 2010). Para o treinamento do modelo, o algoritmo mais difundido na literatura é o *backpropagation* (ou gradiente descendente), o qual utiliza a derivada da função custo baseada no erro quadrático médio para ajustar os pesos sinápticos (Da Silva *et al.*, 2010).

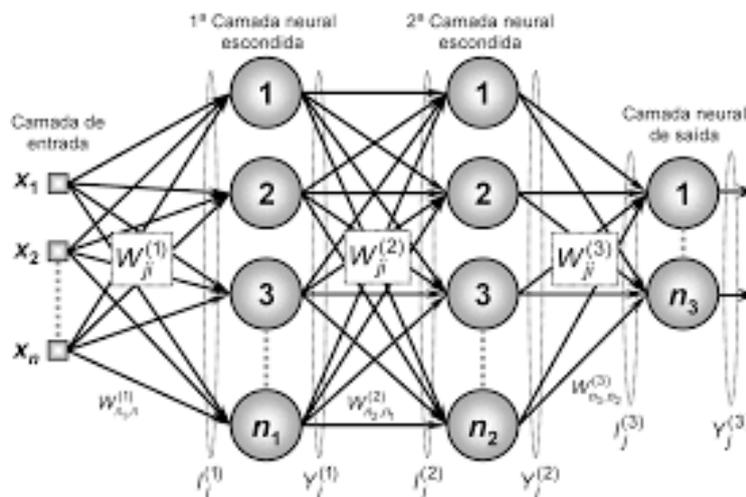


Figura 1. Arquitetura de uma rede neural MLP

A Rede Neural ELM também é *feedforward*, não apresentando realimentação (Figura 2), porém diferente da MLP, não utiliza algoritmos iterativos para ajustar seus pesos sinápticos. Há uma camada intermediária, em que os pesos são atribuídos aleatoriamente, enquanto os pesos referentes a camada de saída são ajustados de forma analítica, sendo necessário encontrar a solução de um sistema de equações lineares, baseada, por exemplo, em uma solução de mínimos quadrados (Huang *et al.*, 2006). Uma forma muito comum para resolver este problema é usando a pseudo inversa *Moore-Penrose*. Dessa forma, essa rede apresenta custo computacional menor do que as outras topologias e resultados eficazes (Huang *et al.*, 2006).

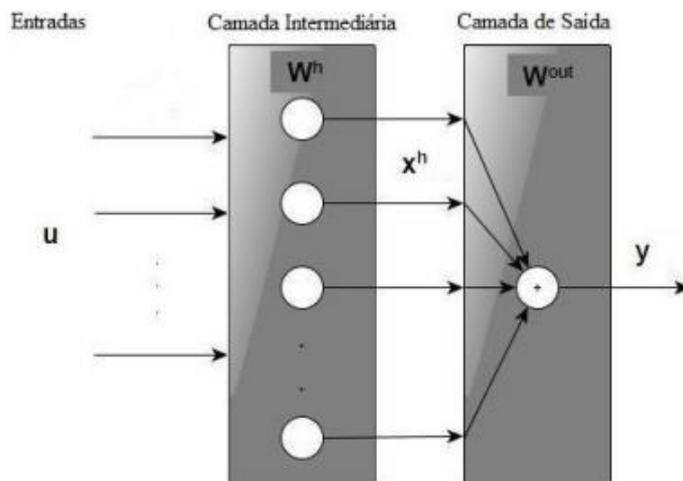


Figura 2. Arquitetura de uma rede neural ELM

Diferente das duas redes neurais artificiais apresentadas, a Rede Neural LSTM é uma rede recorrente (Figura 3), com seu uso normalmente direcionado para previsão de séries temporais (Schäfer & Zimmermann, 2006), uma vez que apresenta em sua arquitetura uma memória das amostras anteriores, ou seja, cada amostra não é tratada individualmente como nas RNAs MLP e ELM. Cada neurônio, geralmente chamado de célula em uma LSTM, recebe três conjuntos de dados: as duas saídas produzidas pela célula no tempo anterior, a saída referente a memória de longo prazo, a saída referente a memória de curto prazo e também recebe os dados de entrada deslocados no tempo  $t$ .

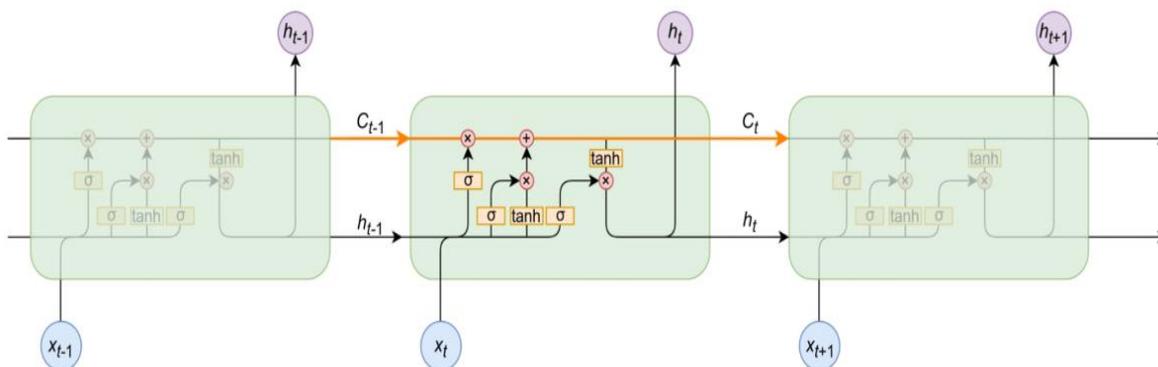


Figura 3. Célula LSTM através do tempo

A célula LSTM apresenta três etapas: portão de esquecimento (*forget gate*), portão de entrada (*input gate*) e o portão de saída (*output gate*) (Chen, 2017). O portão de esquecimento define a porção de  $C(t)$  que será esquecido, portão de entrada define a porção de entrada que será agregado em  $C(t)$  e o portão de saída define a porção do resultado da tangente hiperbólica ( $\tanh$ ) de  $C(t)$  que irá compor o estado intermediário  $H(t)$  (Olah, 2015).

### 3. RESULTADOS & DISCUSSÕES

A base de dados utilizada foi o ativo do mini índice (WINFUT) da bolsa de valores referente ao período de 14/05/2018 até 12/05/2020, em um total de 493 amostras (B3, 2020). Destas 80% foi utilizado para treinar e 20% foi utilizado para o testar os modelos. Essa base de dados é composta por cinco variáveis: valor mínimo, máximo, abertura, fechamento e volume financeiro referente a cada dia. Como entrada dos modelos foi utilizado um *Wrapper* com 2 atrasos para cada uma das variáveis de entrada. A saída foi o fechamento com apenas um passo à frente. Para cada ajuste feito nos modelos, foram realizadas trinta execuções. A Tabela 1 apresenta o menor erro encontrado referente ao modelo com os melhores atrasos e a melhor quantidade de neurônios na camada intermediária.

Tabela 1. Alguns resultados das previsões

Parâmetros	Rede Neural Artificial		
	MLP	ELM	LSTM
RMSE	3.754	11.847	3.065
Atrasos	Mínimo: 1	Mínimo: 1	Fechamento: 1 Mínimo: 1, 2
Neurônios	90	25	15

Como pode ser observado, a Rede Neural LSTM apresentou o menor erro dentre as 3 arquiteturas apesar do custo computacional mais elevado. Este erro representa 5,11% em relação a amplitude referente ao período vigente das amostras, sendo que este ativo apresentou uma mínima e máxima de 62.050 e 121.998, respectivamente.

#### 4. CONCLUSÕES

O presente trabalho comparou o desempenho de três arquiteturas diferentes de Redes Neurais Artificiais (RNAs): *Multilayer Perceptron* (MLP), *Extreme Learning Machine* (ELM) e *Long Short-Term Memory* (LSTM). A base de dados utilizada foi o ativo do mini índice da bolsa de valores referente ao período de 14/05/2018 até 12/05/2020, em um total de 493 amostras. Destas 20% foram utilizadas para o testar e comparar o desempenho das três arquiteturas de redes neurais artificiais mencionadas anteriormente. Para o pré-processamento dos dados foi utilizado a normalização. Os atrasos foram definidos pelo Método *Wrapper* e a quantidade de neurônios da camada intermediária foi definido empiricamente. Como resultado, a Rede Neural LSTM apresentou melhor resultado comparado com as demais.

#### 5. AGRADECIMENTOS

Agradecimentos são externados ao CNPq (Conselho Nacional de Desenvolvimento Científico e Tecnológico) pelo apoio financeiro fornecido ao primeiro autor em 2020 através do Programa de Bolsas de Iniciação em Desenvolvimento Tecnológico e Inovação (PIBITI) da Universidade Tecnológica Federal do Paraná (UTFPR).

#### 6. REFERÊNCIAS

- Braga, A.P., Carvalho, A.P.L.F., Ludermir, T.B., 2007. *Redes Neurais Artificiais: Teoria e Aplicações*. LTC, Rio de Janeiro.
- Brasil, Bolsa, Balcão (3B), 2020. 15 de maio de 2020. <[http://www.b3.com.br/en\\_us/market-data-and-indices/data-services/market-data/historical-data/equities/historical-quotes/](http://www.b3.com.br/en_us/market-data-and-indices/data-services/market-data/historical-data/equities/historical-quotes/)>.
- Chen, E., 2017. *Exploring LSTMs*. 21 de maio de 2020. <<http://blog.echen.me/2017/05/30/exploring-lstms/>>.
- Da Silva, I.N., Spatti, D., Flauzino, R., 2010. *Redes Neurais Artificiais para Engenharia e Ciências Aplicadas: Curso Prático*. Artliber Editora, São Paulo.
- Géron, A., 2017. *Hands-On Machine Learning with Scikit-Learn and TensorFlow*. O'Reilly, Sebastopol.
- Huang, G.-B., Zhu, Q.-Y., Siew, C.-K., 2006. "Extreme learning machine: theory and applications". *Neurocomputing*, Vol. 70, p. 489-501.
- Kelleher, J.D., Mac Namee, B., D'arcy, A., 2015. *Fundamentals of Machine Learning for Predictive Data Analytics: Algorithms, Worked Examples, and Case Studies*. MIT Press, Cambridge.
- Olah, C., 2015. *Understanding LSTM Networks*. 20 de maio de 2020. <<http://colah.github.io/posts/2015-08-Understanding-LSTMs/>>.
- Rêgo, R.H.T., Mussa, A., "Anomalias do mercado acionário: A verificação do efeito feriado no IBOVESPA e IBX-100 no período de 2002 a 2007". In *Anais do 8º Congresso USP de Controladoria e Contabilidade*. São Paulo/SP, Brasil.
- Roque, R.C., 2009. *Estudos sobre a empregabilidade da previsão do índice BOVESPA usando Redes Neurais Artificiais*. Trabalho de Conclusão de Curso, Escola Politécnica, Departamento de Eletrônica e de Computação, Universidade do Rio de Janeiro, Rio de Janeiro/RJ, Brasil.
- Santos, E.P., Von Zuben, F.J., 1999. "Improved second-order training algorithms for globally and partially recurrent neural networks". In *Proceedings of the International Joint Conference on Neural Networks - IJCNN'99*. Washington/DC, United States of America.
- Schäfer, A.M., Zimmermann, H.G., 2006. "Recurrent neural networks are universal approximators". In *Proceedings of the International Conference on Artificial Neural Networks – ICANN 2006*. Berlin, Germany.
- Wang, J.Z., Wang, J.J., Zhang, Z.G., Guo, S.P., 2011. "Forecasting stock indices with back propagation neural network". *Expert System with Applications*, Vol. 38, p. 14346-14355.

## **7. RESPONSABILIDADE PELAS INFORMAÇÕES**

Os autores são os únicos responsáveis pelas informações incluídas nesse trabalho.