

## MODELO DE PREDIÇÃO DA TENDÊNCIA DE SÉRIES TEMPORAIS DO MERCADO FOREX

Victor S. Ribeiro<sup>1\*</sup>, Jaasiel L. dos S. de Carvalho<sup>2</sup>, Diego G. F. Suárez<sup>3</sup>, Leandro S. C. de Souza<sup>4</sup>

1. Estudante de IC do Curso de Bacharelado em Sistemas de Informação da UNEB

2. Estudante do Curso de Bacharelado em Sistemas de Informação da UNEB

3. PIMAT-UNEB - Departamento de Ciências Exatas e da Terra – I UNEB

4. PIMAT-UNEB - Departamento de Ciências Exatas e da Terra – I UNEB / Orientador

### Resumo:

Realizar previsões de séries temporais (ST) por meio do uso da inteligência artificial (IA) vem se tornando comum entre especialistas da área financeira. Pressupõe-se então que a IA, através das redes neurais artificiais (RNA), possa auxiliar a predição de tendência em um dos mercados financeiros mais intrigantes: o Forex. Acredita-se que o movimento desse mercado se repita periodicamente e para prevê-lo, especialistas recorrem a métodos de análise dos dados oriundos deste mercado. Nesse sentido, este trabalho tem como objetivo propor um indicador aplicável para predição da tendência de ST, utilizando as RNA, a fim garantir a assertividade de predição. Para alcançar esse objetivo, foi utilizada uma arquitetura de RNA, que quando empregado em conjunto do indicador, resultou em taxas de acerto consideráveis. Resultados alcançados indicam que esse modelo garante a eficácia da predição e seu estudo e desenvolvimento podem trazer benefícios para a área de predição de ST estocásticas.

**Palavras-chave:** Redes Neurais Recorrentes; Indicadores de Tendência; Análise Técnica.

**Apoio financeiro:** Fundação de Amparo à Pesquisa do Estado da Bahia– FAPESB / Universidade do Estado da Bahia - UNEB.

**Trabalho selecionado para a JNIC pela instituição:** Pró-Reitoria de Pesquisa e Ensino de Pós-Graduação– PPG - Universidade do Estado da Bahia - UNEB.

### Introdução:

Desde o seu início, a humanidade tenta prever eventos pelos mais variados motivos. Na área da ciência, pode-se entender uma predição como sendo uma antecipação de acontecimentos futuros de acordo com as condições existentes, passadas e presentes. Assim, desde que o homem passou a estudar diversos fenômenos e questões, a predição vem sendo considerada em diversos aspectos, uma área bastante intrigante. São muitos os eventos a serem estudados com fulcro na tentativa de assertividade de seus comportamentos, em geral estes eventos apresentam complexidade elevada, transitando os mais variados temas de interesses (REINSEL, 2008).

Independente do fator motivacional, uma das técnicas empregadas para realização de predições reside na análise de dados prévios relacionados com o evento em estudo e geralmente são apresentados como séries temporais. Neste sentido, a análise minuciosa destas séries pode viabilizar a previsão do comportamento futuro desejado (PROCHNOW, 2013).

Dentre as séries temporais com maior grau de complexidade encontram-se as séries financeiras. Nestas, essa análise dos dados passados é sempre utilizada para auxiliar nas tomadas de decisões, possibilitando maior grau de assertividade (FILHO, 2015). Contudo, as séries financeiras apresentam, em geral, um comportamento estocástico, ou seja, seus valores futuros podem ser descritos apenas através de probabilidade (OLIVEIRA, 2007), inexistindo métodos e funções matemáticas capazes de determinar o seu comportamento futuro, revelando-se assim um objeto desafiador para cientistas e investidores. Estas séries são ainda mais intrigantes quando estudadas a partir do mercado financeiro Forex.

O Forex é composto por séries temporais que são criadas a partir de cotações, preços dos ativos ao longo do tempo, fornecendo assim dados que são a base de observação para aplicação de diferentes técnicas de análise com fulcro em prever a tendência futura. Estudos apontam para utilização de distintas técnicas para predição das séries do Forex. Alguns destes apontam para utilização de lógica *fuzzy* (CHEN, 2006), outros utilizam técnicas econométricas (WOLLRIDGE, 2000) e dentre os mais referenciados estão as Redes Neurais Artificiais (YAO, 2000) (DHAMIJA, 2010) (PERSIO, 2016).

O objetivo desse trabalho é propor um indicador, através da utilização de uma arquitetura específica de Redes Neurais artificiais com capacidade de predição no Forex.

### Metodologia:

A partir da revisão sistemática realizada definiu-se por utilizar uma Rede Neural Artificial Recorrente, tipo *Long Short Term Memory* (LSTM) como técnica de Inteligência Artificial para predizer tendências das séries temporais estudadas. Esta escolha se justifica devido ao fato da LSTM apresentar memória dos dados para análise entre suas camadas, sendo assim a mais indicada para o estudo proposto.

Foi proposto então um modelo de predição o qual se inicia com a obtenção dos dados do mercado Forex através da utilização de uma plataforma eletrônica de negociação, MetaTrader© (YOUNG, 2015). Esses dados são analisados e discretizados com a finalidade de alimentar a LSTM definida para predição, que tem o objetivo de indicar uma decisão operacional de tendência, sinalizando-a como de alta, baixa ou estacionária.

A técnica proposta para discretização dos dados se baseou inicialmente na utilização das cotações coletadas do mercado, que compõe uma série temporal para o período específico de um mês de dados. Assim sendo, o arquivo coletado fora particionado em arquivos que representassem os dados diários de cotações das séries temporais. Em seguida cada um destes arquivos fora novamente fracionado, desta feita em um total de 24 conjuntos de uma hora cada. Portanto, a análise foi feita em cima desses novos conjuntos contendo uma nova frequência, hora a hora.

Para que a análise técnica pudesse ser viabilizada, fez-se necessário ainda a extração de características destas novas séries temporais, obtendo-se um fragmento importante para a concepção do modelo discretizado. Utilizando-se da estatística descritiva, cujo objetivo é conferir a um conjunto de dados medidas de tendência central, como a mediana e média, ou medidas de variabilidade, como desvio padrão e variância, foi definido o modelo.

Este foi elaborado seguindo os seguintes passos: a) Delimitação de uma janela L, tomando-se como base a frequência dos arquivos, ou seja, foram utilizadas janelas que representam uma hora cada; b) extração da média a partir dos *ticks* que formam cada série temporal; c) quantificação tanto dos *ticks* acima (TicAc) quanto abaixo (TicAb) desta média; d) aplicação da diferença entre os valores (TicAc) e (TicAb).

Empregando o conceito da componente tendência de uma série temporal, foi feita uma classificação de acordo com o resultado da diferença: a) se  $(TicAc - TicAb) > 0$ , então o fragmento da série tem tendência crescente; b) se  $(TicAc - TicAb) < 0$ , então o fragmento da série tem tendência decrescente; c) se  $(TicAc - TicAb) = 0$ , então o fragmento da série não possui uma tendência suficientemente forte para nenhuma das direções. Sabendo-se o resultado da diferença, pode-se então classificar a série analisada atribuindo os valores 1, -1 e 0, seguindo respectivamente a mesma regra acima.

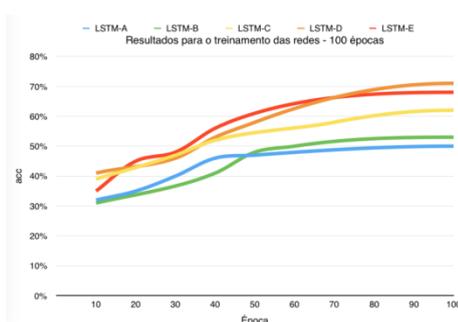
## Resultados e Discussão:

Embora a metodologia para a definição da tendência de uma série (TicAc - TicAb) possa parecer simplória, entende-se que essa métrica é relevante no sentido de minimizar as entradas da LSTM e assim poder viabilizar, em tempo hábil, as operações sob o mercado. Além disso, uma variação sobre o resultado operacional da métrica poderia ser aplicada no intuito de obter-se uma melhor representação de tendência, *i.e.*, poderia ser considerada uma faixa de valores para determinar a tendência (*e.g.*,  $10 < (TicAc - TicAb) \leq -10$ , considerar-se-ia o fragmento de série analisada sem tendência). Não obstante, resultados preliminares mostraram que a utilização destas faixas não altera os resultados, e por isso, não foram considerados no estudo.

Os testes para comprovação dos resultados iniciaram com o propósito de definir os melhores parâmetros para a rede neural artificial, LSTM. Para tanto, foram variadas as quantidades de neurônios na rede neural. Houve ainda alteração na quantia de épocas usada no treinamento, a fim de se encontrar um valor ideal suficiente para que a rede pudesse aprender o mecanismo gerador dos dados, sendo capaz então de realizar previsões satisfatórias.

Assim foram definidas cinco redes LSTM: (A, B, C, D e E) cada uma com 8, 16, 32, 64 e 128 neurônios respectivamente. Todas receberam um universo correspondente a dezoito dias de dados exclusivamente utilizados para treinamento, dos vinte e dois disponíveis para experimentação. Para validação e quantificação da eficácia das redes, foram utilizados os quatro dias restantes.

A figura 1a apresenta a acurácia (acc) das redes neurais durante a fase de treinamento. Os resultados mostram claramente que há uma tendência de estabilidade em relação a métrica acc durante as épocas de treinamento para todas as redes experimentadas. É importante notar que embora a tendência seja comum a todas as arquiteturas, a época em que cada uma delas inicia o período de estabilidade é diferente. Tomando-se como exemplo a rede LSTM-A que, a partir de época 40, começa a deixar de ter um ganho significativo de desempenho, independente da quantidade de épocas a mais empregadas para seu treinamento. Modelos como estes apontam para o problema de *underfitting* (PERSIO, 2016).



(a) Acurácia das Redes em função das épocas de treinamento



(b) Assertividade do modelo em operação no mercado

Figura 1. Resultados experimentais.

De forma sucinta, o número de neurônios afeta a capacidade de aprendizagem da rede. Geralmente, mais neurônios implicam maior possibilidade de aprendizado por parte da rede sobre a estrutura do problema. No entanto, a maior capacidade de aprendizagem vem acompanhada do problema potencial de superadaptação aos dados de treinamento, algo que deve sempre ser levado em consideração quando se tem como objetivo a concepção de um modelo capaz de realizar previsões com um novo conjunto de dados.

Uma vez definida a rede neural, configurada e treinada, passou-se a observar o percentual de acerto das previsões em função das observações das tendências futuras realizadas pela rede em relação ao conjunto de teste. A figura 1b, indica que a rede LSTM-D, com 64 neurônios obteve 62% de acerto em suas previsões, mesmo tendo apenas 71% de assertividade na fase de treinamento (figura 1a).

### **Conclusões:**

Este trabalho teve como objetivo principal a análise dos dados históricos do mercado Forex com intuito de definir um conjunto de dados discretizados para alimentar uma rede neural artificial do tipo recorrente com fulcro na previsão de tendências.

Os resultados preliminares apontaram uma boa relação de indicação de tendência com os dados discretizados tomando por base a abordagem técnica para análise do mercado: a tendência. Já os resultados consolidados demonstram claramente a eficiência do modelo preditivo proposto, onde as redes neurais utilizadas em conjunto com os dados modelados com a discretização apresentaram eficácia de 62% na taxa de acerto. Embora esse resultado seja promissor, não se pode assegurar ganhos financeiros reais baseando-se apenas no modelo proposto. No entanto, uma abordagem mais efetiva que permita reconhecer a direção da tendência principal do mercado pode auxiliar a realizar operações que de fato atinjam um percentual de assertividade aceitável.

### **Referências bibliográficas**

- CHEN, C.-H. L. L. A. L. W.-S. Pattern discovery of fuzzy time series for financial prediction. *IEEE Transactions on Knowledge and Data Engineering*, v. 18, p. 613–625, 2006.
- DHAMIJA, A.; BALLA, V. Financial time series forecasting: Comparison of neural networks and arch models. *International Research Journal of Finance and Economics*, v. 1, n. 1, p. 194–197, 2010.
- FILHO, M. B. Indicadores baseados em diferenças de primeira ordem para orientar robôs no mercado forex. Trabalho de Conclusão de Curso. Universidade do Estado da Bahia. Departamento de Ciências Exatas e da Terra – I. Curso de Sistemas de Informação: Bacharelado. 2015. Disponível em: <<http://www.csi.uneb.br/csitcc/files/ArqTrabCien-4237612.pdf>>
- JIANG, H. G. Q. Pattern matching of time series and its application to trend prediction. *Anti-counterfeiting, Security and Identification*, 2008. ASID 2008. 2nd International Conference, p. 341–344, 2008.
- OLIVEIRA, M. A. Aplicação de redes neurais artificiais na análise de séries temporais econômico-financeiras. 2007. Tese (Doutorado em Administração) - Faculdade de Economia, Administração e Contabilidade, Universidade de São Paulo, São Paulo, 2007. doi:10.11606/T.12.2007.tde-31012008-112504.
- PERSIO, L. D.; HONCHAR, O. Artificial neural networks approach to the forecast of stock market price movements. *International Journal of Economics and Management Systems*, p. 158–162, 2016.
- PROCHNOW, F. A. Programação genética para previsão de séries temporais aplicados a merca dos financeiros. Trabalho de Conclusão de Curso - Universidade Federal do Rio Grande do Sul. Instituto de Informática. Curso de Ciência da Computação: Ênfase em Ciência da Computação: Bacharelado. 2013. Disponível em <<http://hdl.handle.net/10183/66085>>.
- REINSEL, G. E. P. B. G. M. J. G. C. *Time Series Analysis: Forecasting and Control*. 4th. ed.[S.l.]: John Wiley Sons, Inc., 2008. ISBN 9780470272848.
- WOLLRIDGE, J. M. *Introductory econometrics: a modern approach*. South-Western CollegePublishing, a division of Thomson Learning, p. 8, 2000.
- YAO, J.; TAN, C. L. A case study on using neural networks to perform technical forecasting of forex. *Neurocomputing*, v. 34, p. 79–98, 2000.
- YOUNG, A. *Expert Advisor Programming for MetaTrader 4: Creating Automated Trading Systems in the MQL4 Language*. Edgell Publishing, 2015. ISBN 9780982645932.