

Previsão de cotação de ações da bolsa de valores usando redes neurais recorrentes - RNN

Stock Market Forecasting Using Recurrent Neural Networks - RNN

Lucas Mrowskovsy Paim*

2019

Resumo

Este trabalho tem por objetivo criar uma rede neural de aprendizagem profunda para prever valores futuros de um papel na bolsa de valores, o papel escolhido foi PETR4 (Petrobras). A biblioteca utilizada para a criação das hipóteses foi a tensorflow, que segundo [Géron \(2019\)](#), é uma ótima biblioteca para cálculo número de código aberto, o modelo atingiu um erro médio de apenas R\$0,68, que apesar de ser baixo o mercado de ações é extremamente volátil, sendo necessário adicionar N outras características ao treinamento, sendo assim, não podendo ser utilizado em produção.

Palavras-chave: deep-learning. tensorflow. rnn. séries temporais.

Abstract

This work aims to create a deep learning neural network to predict future values of a role in the stock exchange, the role chosen was PETR4 (Petrobras). The library used for hypothesis creation was tensorflow, which according to [Géron \(2019\)](#), is a great open source number library. The model reached an average error of only R\$ 0.68, which despite being low the stock market is extremely volatile, being necessary to add N other characteristics to the training, so it cannot be used in production.

Keywords: deep-learning. tensorflow. rnn. temporal series.

*Programa de Pós Graduação em Inteligência Artificial, Curitiba - PR, 80215-901; Especializando em Inteligência Artificial Aplicada na PUCPR; Tecnólogo em Sistemas para Internet pela Universidade Positivo; E-mail: lucasmpaim1@gmail.com

1 Introdução

Segundo Ehlers (2009), uma série temporal é uma coleção de observações ao longo do tempo, isto é, são dados que variam de acordo com o passar do tempo, como: comportamentos cerebrais Ulrich et al. (2014), identificação de ritmos musicais, auxílio na composições musicais Corrêa e Saito (2007), taxa mensal de desemprego, eletrocardiograma, etc. As redes neurais recorrentes tentam identificar estes padrões fazendo conexões com neurônios anteriores. Este trabalho irá utilizar redes neurais com aprendizagem profunda, para tentar prever os valores futuros das cotações do papel PETR4 (Petrobrás), usando dados extraídos do Google Finance, com dados a partir de 2006.

2 Método

A base utilizada para este trabalho, foram os dados históricos das ações da Petrobrás (PETR4) extraídos a partir do google finance, para o período de 05/09/2006 até 07/11/2019, estes dados podem ser visualizados na Figura 1.

Utilizando a biblioteca tensorflow (v. 2.0.0), foi montada uma rede neural recorrente (RNN), utilizando duas camadas de neurônios de memória LSTM (Long Short Term Memory), o modelo gerado pode ser consultado na imagem: Figura 2

Em um primeiro momento, a base foi dividida em teste e validação em que os últimos 6 valores da base foram separados em: 5 para dados e 1 previsão e o resto da base é utilizada para treinamento, em que também é dividida na proporção 5:1 e treinada durante 100 épocas, estas configurações foram escolhidas empiricamente, para a escolha desses hiperparâmetros, uma abordagem melhor seria com o uso de algoritmos de busca como os algoritmos genéticos.

O gráfico de loss durante o treinamento, comparando a base de validação e a de treinamento pode ser visualizada na Figura 3, nele podemos notar que não houve overfitting na rede, as camadas "Dropout" que aparecem na Figura 2 que segundo Maklin (2019) definindo aleatoriamente as arestas de saída das neurônios ocultos como 0, é possível ver graficamente a ação desta célula na Figura 5.

As células LSTM, proposta em 1997 por Sepp Hochreiter e Jurgen Schmidhuber e aperfeiçoada desde então, que segundo Géron (2019), sua ideia-chave é que a rede possa aprender o que armazenar a longo prazo durante o treinamento.

Para a avaliação do erro foi utilizado a métrica MAE (Mean Absoluty Error), que é a média aritmética da diferença entre o valor previsto e o valor real, ela pode ser representada pela fórmula:

$$mae = \frac{1}{n} \sum_{i=0}^M |y_1 - y_i^*| \quad (1)$$

3 Resultados

O modelo atingiu um erro médio de R\$0,68, que é um erro médio aceitável para este trabalho.

O grafico de previsões pode ser visualizado na Figura 4, os valores estão detalhados na tabela a seguir:

Predict	True	Error
-----	-----	-----
27.0385	27.06	0.0215228
27.2117	26.88	0.331698
27.3673	28.06	0.692717
28.3972	27.69	0.707238
28.3212	27.22	1.10122
28.0565	27.29	0.76648
28.0641	27	1.06411
27.818	27.48	0.337986
28.1443	27.27	0.874328
28.0211	27.34	0.681134
28.0867	27.7	0.386704
28.3649	27.66	0.704869
28.3885	27.55	0.838499
28.3401	27.51	0.830096
28.3062	26.72	1.5862
27.7085	26.74	0.968548
27.6006	26.51	1.09064
27.339	26.17	1.16899
27.0153	26.02	0.995316
26.8138	26.52	0.293842
27.13	26.74	0.390012
27.3374	27.26	0.0774144
27.8091	27.31	0.499109
27.953	27.6	0.353045
28.2499	27.93	0.319885
-----	-----	-----

Mesmo a rede tendo resultados satisfatórios, é necessário se considerar que o mercado de ações é extremamente volátil e que não segue um padrão em si, sendo facilmente influenciado por fatos políticos, econômicos etc. que não necessariamente faça com que a ação tenda a seguir um padrão.

Referências

CORRÊA, D. C.; SAITO, J. H. Aplicação de redes neurais para auxílio nas composições musicais utilizando compassos como primitivas e inspiração em relevos naturais. In: TORRES, M. H. C. (Ed.). *Proceedings of the 11th Brazilian Symposium on Computer Music*. São Paulo - SP - Brazil: Sociedade Brasileira de Computação, 2007. p. 251–254. Citado na página 2.

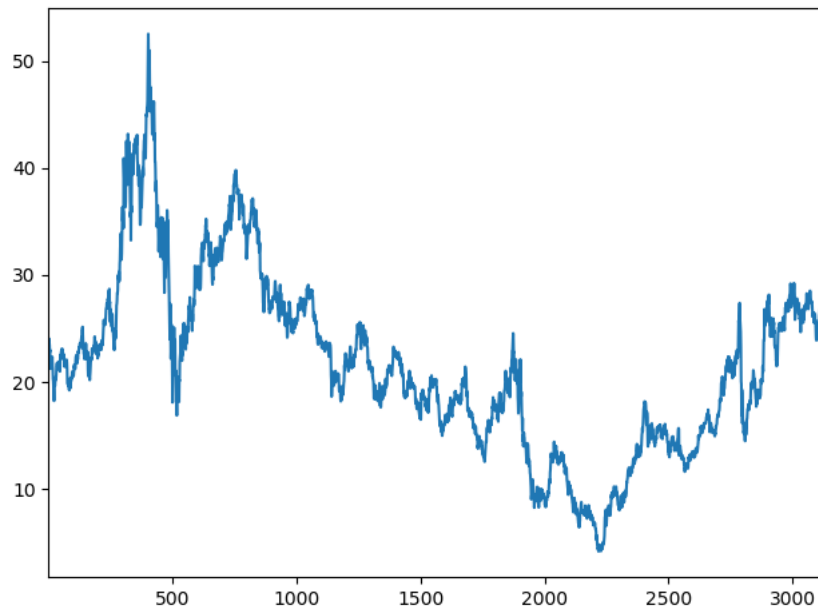
EHLERS, R. Análise de séries temporais. In: . [s.n.], 2009. Disponível em: <http://www.icmc.usp.br/ehlers/stemp/stemp.pdf>. Acesso em: 08 nov 2019. Citado na página 2.

GÉRON, A. *Mãos à obra: Aprendizagem de máquina com Skit-Learn e Tensorflow*. 1. ed. Rio de Janeiro: Alta Books Editora, 2019. Acesso em: 21 ago 2013. Citado 2 vezes nas páginas 1 e 2.

MAKLIN, C. Dropout neural network layer in keras explained. In: . [s.n.], 2019. Disponível em: <<https://towardsdatascience.com/machine-learning-part-20-dropout-keras-layers-explained-8c9f6dc4c9ab>>. Acesso em: 08 nov 2019. Citado na página 2.

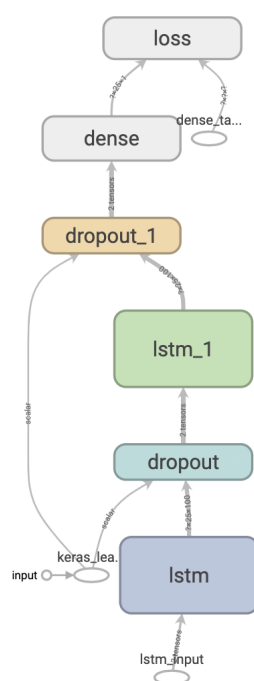
ULRICH, K. R. et al. Analysis of brain states from multi-region lfp time-series. In: GHAHRAMANI, Z. et al. (Ed.). *Advances in Neural Information Processing Systems 27*. Curran Associates, Inc., 2014. p. 2483–2491. Disponível em: <<http://papers.nips.cc/paper/5624-analysis-of-brain-states-from-multi-region-lfp-time-series.pdf>>. Citado na página 2.

Figura 1 – Gráfico histórico PETR4 R\$ x Dias úteis após 05/09/2006



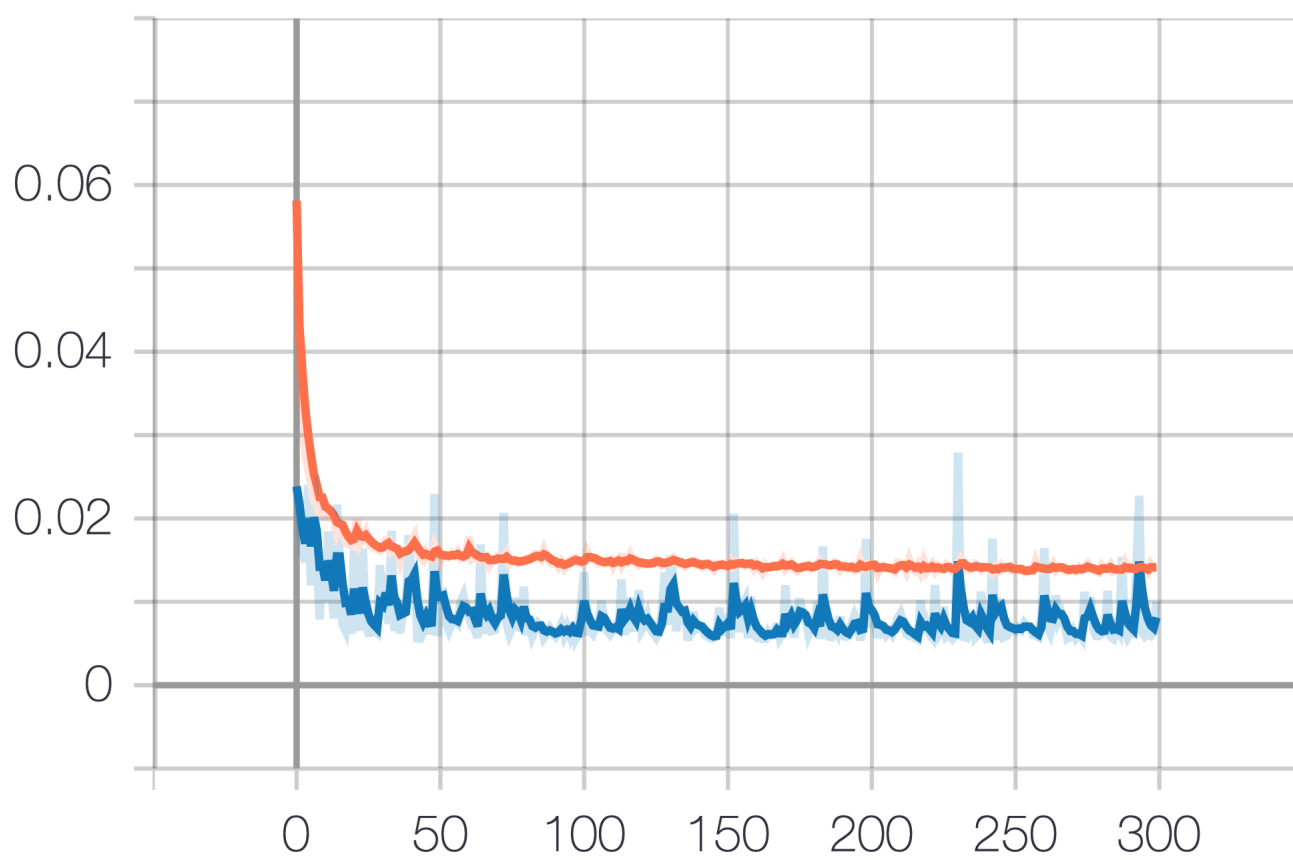
Fonte: O Autor

Figura 2 – Modelo



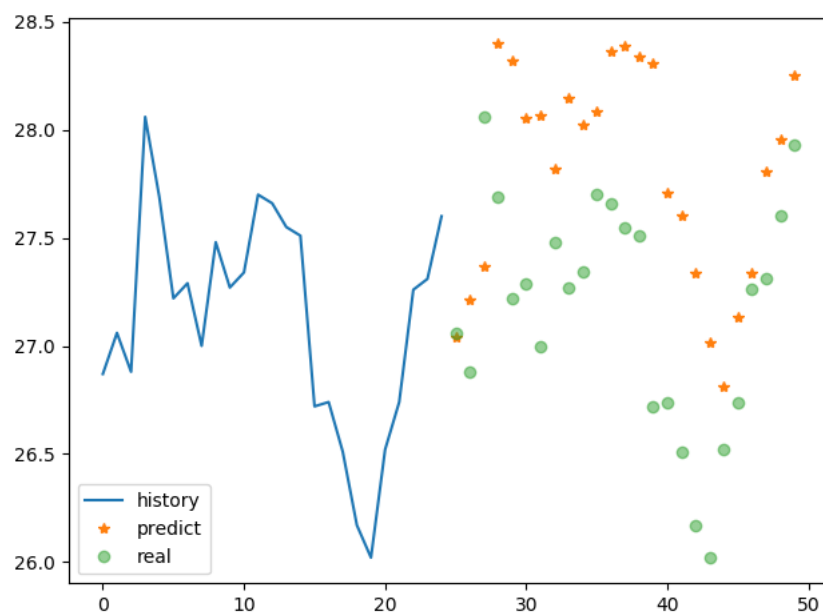
Fonte: O Autor

Figura 3 – Loss



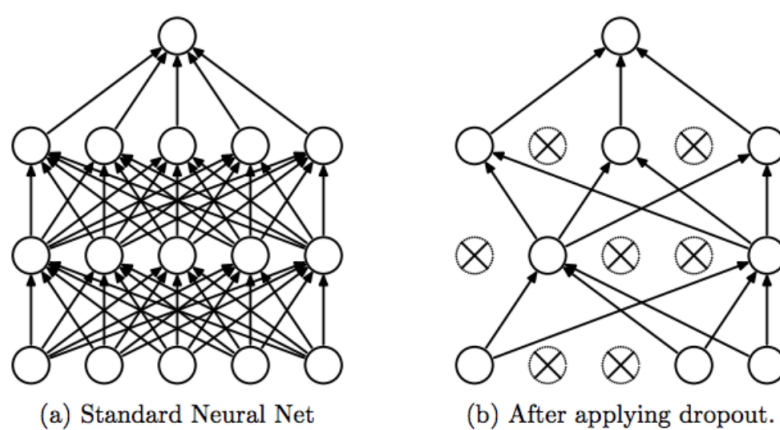
Fonte: O Autor

Figura 4 – Loss



Fonte: O Autor

Figura 5 – Loss



Fonte: [Maklin \(2019\)](#)