

ESTIMATIVA DE EVENTOS FUTUROS COM BASE EM REDES NEURAIS ARTIFICIAIS UM ESTUDO DE CASO DO SETOR DE BENS DE CONSUMO

Raiana Roland Seixas¹-raianars@hotmail.com

Maria Gabriela Marques¹- marquesgabi@hotmail.com

Ubirajara Coutinho Filho¹ -ucfilho@ufu.br

¹ Federal University of Uberlandia, School of Chemical Engineering – Uberlandia, MG, Brazil

Abstract. *A previsão acurada de diferentes eventos é um aspecto chave para o planejamento com sucesso em setores como previsão do tempo, estimativa da poluição do ar e avaliação de estratégias de negócio. Neste contexto, as redes neurais recorrentes do tipo LSTM (long-short term memory) tem sido usadas com sucesso na descrição de séries temporais tanto em situações tradicionais como em casos de dados ausentes, outliers e não-linearidade, o que pode reduzir a eficiência dos modelos tradicionais como os modelos de autorregressão. Neste artigo, o uso de LSTM como modelo de redes neurais recorrentes foi utilizado para descrever a venda de bens duráveis de consumo, utilizando-se um exemplo da venda de móveis para descrever este setor. Seis condições de diferentes usos de variáveis foram estudadas utilizando o coeficiente de determinação e o erro quadrático médio como métricas de ajuste dos modelos. Relativo ao uso do LSTM, foi observado que este modelo foi capaz de descrever a sazonalidade das vendas com coeficiente de determinação superior a 0,7. Relativa à descrição das vendas, foi observado que o uso das vendas nos três meses que antecedem o mês de previsão foi a situação que gerou os melhores resultados. Também foi constatado que o uso adicional de vendas de bens de escritório, considerando-se os três últimos meses, melhorou a qualidade do modelo LSTM que se mostrou promissor, com capacidade de descrever as sazonalidades mesmo quando se considera um período de previsão de vendas de dois anos.*

Keywords: *Séries temporais, redes neurais artificiais, LSTM, vendas de bens de consumo, inteligência artificial, previsão de eventos futuros*

1. INTRODUÇÃO

Embora não tenha consenso de definição de inteligência artificial (AI), pela própria dificuldade de se definir inteligência para seres vivos em geral e humanos em particular, tem-se que o uso de AI cada vez mais se destaca mais como uma das áreas mais promissoras no que diz respeito a possibilidades de uso em diferentes campos do desenvolvimento humano. Tem-se que a mesma se desenvolve de forma tanto prática como acadêmica, com destaque em quatro grandes áreas que são a resolução de problemas inspirados no funcionamento do cérebro, sistemas robóticos que simulam a atuação humana, sistemas que usam a lógica racional e agentes inteligentes (Norvig & Stuart, 2001).

No que se refere a resolução de problemas inspirados no funcionamento do cérebro, tem-se que as redes neurais artificiais (RNAs) se destacam como uma das formas de maior sucesso do uso de AI em diversas áreas como obtenção de modelos que descrevem estudos científicos,

comerciais e aplicações industriais em situações diversas como, por exemplo, na estimativa respostas associadas a eventos futuros ou previsão do que vai acontecer antes mesmo dos primeiros sinais se tornarem evidentes. A estimativa de eventos futuros pode ser feita com o uso de séries temporais e constitui uma área de interesse em diferentes contextos da vida moderna, sendo que são alguns exemplos de interesse, a avaliação de alterações climáticas, de níveis de poluição e a previsão de estoques ou vendas associadas a determinados produtos.

O uso de RNAs na descrição de séries temporais tem-se tornado cada vez mais comum como forma de substituição dos modelos de autorregressão (Box & Jenkins, 1976) e, também, como forma de melhorar o desempenho e a rapidez de cálculos seja em situações tradicionais ou em situações nas quais os modelos de autorregressão não podem ser aplicados de forma satisfatória pela presença de outliers, dados ausentes ou não-linearidade dos problemas. No contexto do uso de RNAs para séries temporais, um dos modelos mais recentes em uso é o modelo LSTM criado por Hochreiter e Schmidhuber em 1997 (Hochreiter & Schmidhuber, 1997) no qual a modulação dos pesos sinápticos (W), associados ao aprendizado da rede neural, passam a ser utilizados na avaliação da resposta no tempo futuro (x_{t+1}), não contendo somente as informação do tempo atual (x_t).

No modelo LSTM a resposta no tempo futuro (x_{t+1}) utiliza a resposta no tempo atual (x_t) e, também, as saídas associadas a função de transferência aplicada a esta informação (h_t), as variáveis de entrada das interações anteriores e as saídas destas variáveis (x_{t-1} e h_{t-1} , respectivamente) de uma forma recorrente na qual uma função logística (σ) modula a importância da informação anterior. A modulação da importância da informação significa que o modelo determina se a memória de curto prazo e longo prazo tem maior ou menor influência na resposta como mostra a Figura 1, que também apresenta as quatro equações que em conjunto descrevem a célula de recorrência LSTM. Na Figura 1, os valores r_t e z_t da função logística (também chamada sigmóide) assumem valores entre 0 e 1, sendo que z_t representa a parte do modelo LSTM que atribui maior ou menor importância a memória de longo prazo ou curto prazo, conforme pode-se observar na equação que descreve h_t , a qual utiliza os valores de z_t com o termo associado a memória mais antiga, h_{t-1} , tendo menor importância quando z_t se aproxima de 1 e maior importância quando se aproxima de 0.

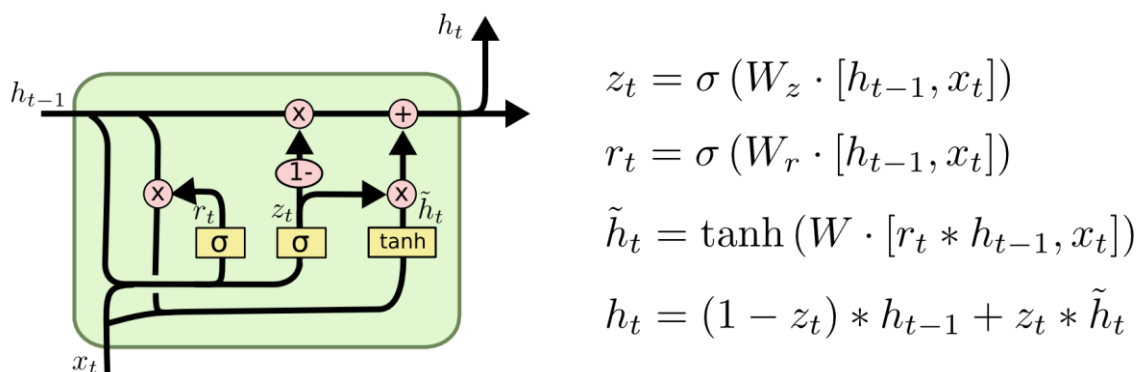


Figura 1: célula LSTM de uma RNA para séries temporais (Olah, 2015)

No presente trabalho é feita a discussão do uso de redes neurais do tipo LSTM na avaliação de séries temporais, sendo utilizando como exemplo de estudo de caso um problema de avaliação de vendas de bens de consumo duráveis.

2. METODOLOGIA

2.1 Problema escolhido para os modelos de séries temporais

No presente estudo foi avaliado a venda de móveis no período de 2014 a 2018 de uma loja de departamentos americana que teve um lucro líquido de 286 milhões neste período associado a venda de móveis e que, também, efetua a vendas de artigos de escritório e tecnologia. A Figura 2 apresenta detalhes da venda de móveis, já nas Figuras 3 e 4 são apresentados, respectivamente, detalhes das vendas de artigos de escritório e tecnologia da referida empresa entre 2014 a 2018. Nestas figuras pode-se observar que estas vendas possuem sazonalidade e variam de forma distinta para os diferentes setores representados no arquivo de dados originais (Superstore.xls) que contém 9994 registros de vendas que descrevem vendas diárias dos diferentes produtos como pode ser observado em detalhe no dados originais que estão disponíveis na plataforma de programação Github que pode ser acessada pelo link https://github.com/ucfilho/Machine_Learning_DataSet.

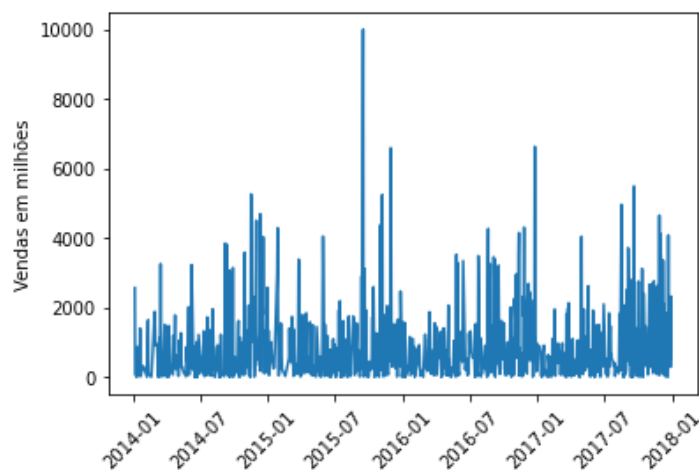


Figura 2: vendas de móveis

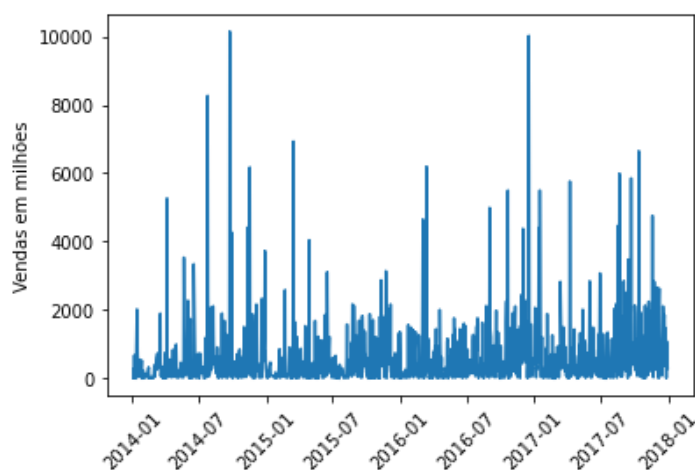


Figura 3: vendas de artigos de escritório

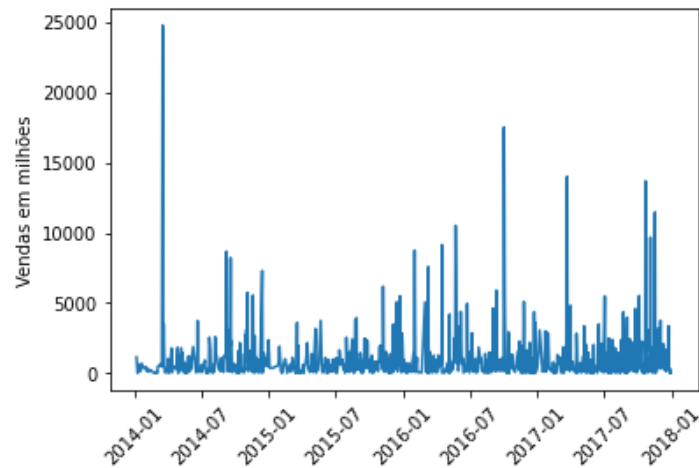


Figura 4: vendas de artigos de tecnologia

2.2 Modelos de séries temporais escolhidas

Após a etapa inicial de análise dos dados na qual os 9994 registros de vendas foram filtrados e agrupados na categoria venda por mês, foram elaborados os diferentes modelos de previsão de séries temporais. Para a construção dos modelos, foi utilizada a metodologia de RNAs do tipo recorrente com memória curta e longa dita LSTM (Hochreiter & Schmidhuber 1997) implementada em *Python* na nuvem utilizando o *TensorFlow 2.5* na plataforma *Google Colab* (<https://colab.research.google.com/>). Os modelos foram elaborados com os dados de 2014 a 2016 e a validação dos modelos foram feitas utilizando as previsões para os anos de 2017 e 2018, sendo que na construção dos modelos foram realizadas variações nas quantidades dos meses anteriores associada a cada entrada considerada na previsão dos resultados futuros e, para garantir a validade dos modelos utilizados, foram avaliados os coeficientes de correlação e erro quadrático médio para os dados de validação e dados de treino.

3. RESULTADOS

3.1. Estudo comparativo dos modelos de RNAs

Na Tabela 1 são apresentados os resultados comparativos da consideração de diferentes números de meses anteriores às vendas como variáveis para previsão da venda futura na série temporal. Pode-se observar, considerando os resultados associados ao teste, que um, dois ou quatro meses anteriores ao mês que se deseja prever a venda não favorece uma previsão tão precisa quanto a que se obtém quando as vendas atuais são estimadas considerando os três últimos meses de vendas.

A Figura 5 apresenta os resultados considerando as vendas em cada um dos três últimos meses como variáveis de entrada da rede LSTM (situação que representa o melhor modelo descrito anteriormente na Tabela 1). Na Figura 5 pode-se observar que os quatro anos são representados na forma de 48 meses e que a previsão futura, meses 25 a 48, utilizando o treinamento com os 24 meses iniciais mostra que o modelo LSTM tem capacidade de prever a sazonalidade adequadamente. O longo período que o modelo se mantém capaz de prever

adequadamente as sazonalidades, entre os meses 25 a 48, mostra que o uso de LSTM é promissor neste complexo problema de prever o futuro de vendas valendo-se apenas das informações das vendas nos meses anteriores sem o conhecimento de detalhes da economia e mesmo outros fatores que alteram o mercado e o funcionamento da empresa.

Tabela 1: avaliação de diferentes arquiteturas de RNAs.

Ensaio	Meses móveis	R ² treino	MSE treino	R ² teste	MSE teste
1	1	0,742	0,0156	0,558	0,0297
2	2	0,726	0,0166	0,512	0,0320
3	3	0,792	0,0122	0,639	0,0240
4	4	0,696	0,0719	0,630	0,0250

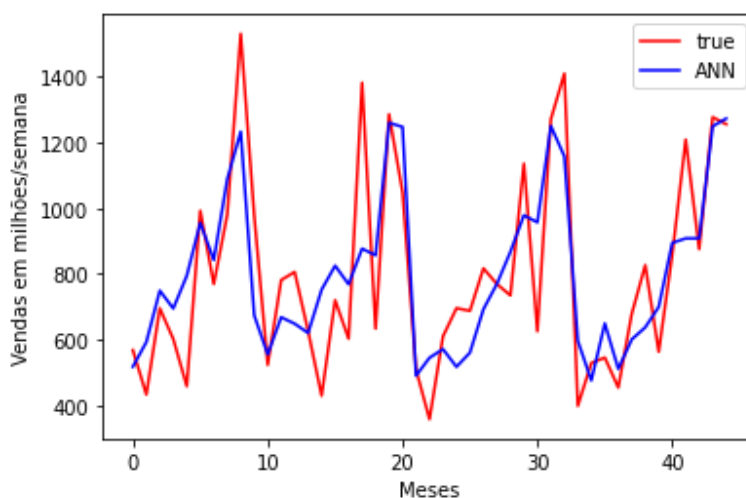


Figura 5: avaliação do modelo LSTM na previsão de vendas de móveis

Na Tabela 2 são apresentados os resultados comparativos da avaliação de diferentes números de meses das vendas anterior ao mês de interesse, considerando-se agora não somente móveis, mas também artigos de escritórios como variáveis para previsão da venda futura. Pode-se observar que a venda de móveis tem uma melhor previsão quando se considera a venda de artigos de escritórios nos meses anteriores como variável de entrada adicional na estimativa das vendas futuras. A situação mais favorável encontrada apresenta como variáveis de entrada as vendas em três meses anteriores, tanto para artigos de escritório, como para vendas de móveis. Nesta condição o modelo LSTM foi capaz de prever a resposta com menor erro quadrático médio e maior coeficiente de correlação, considerando os dados de teste, os quais não foram utilizados no treinamento, e assim representam uma boa referência do aprendizado do modelo sem a ocorrência de *overfitting*.

Tabela 2: avaliação de diferentes arquiteturas de RNAs.

Ensaio	Meses móveis	Meses escritório	R ² treino	MSE treino	R ² teste	MSE teste
1	3	1	0,679	0,0195	0,558	0,0290
2	2	2	0,613	0,0254	0,541	0,0279

3	3	3	0,755	0,0163	0,717	0,0170
4	4	4	0,754	0,0166	0,643	0,0210

4. CONCLUSÕES

A correta previsão de vendas futuras é de fundamental importância para empresas, pois permite a maior segurança na tomada de decisões, incluindo o planejamento de demandas futuras. O uso de modelos de redes neurais para séries temporais do tipo LSTM mostrou-se promissora no estudo de caso apresentado de estimação de vendas de móveis. mesmo na ausência de detalhes da economia, os quais impactam as vendas. A adequação dos modelos pelos critérios MSE e coeficiente de determinação (r^2) mostraram que a previsão de vendas futuras feita considerando a vendas dos últimos três meses como referência foi melhor que a situação com um número de meses inferior ou superior a três meses. Além disso, também foi observado que, quando se considerou as vendas de artigos de escritório em conjunto com a própria venda de móveis como variáveis, o modelo LSTM foi capaz de melhorar a qualidade da previsão de vendas futuras, utilizando os mesmos três meses anteriores de vendas tanto para móveis como para artigos de escritório como referência.

Acknowledgements

Os autores agradecem a CAPES pelas bolsas de estudo.

REFERÊNCIAS

- Box, G. E. P.; Jenkins, G. M. (1976) Time Series Analysis: Forecasting and Control. São Francisco: Holden-Day, p 199-201.
- Hochreiter, S. and Schmidhuber, J. (1997). LSTM can solve hard long time lag problems. In Advances in Neural Information Processing Systems 9. MIT Press, Cambridge MA. Presented at NIPS 96.
- Olah, C (2015) Understanding LSTM <http://colah.github.io/posts/2015-08-Understanding-LSTMs/>
- Russell. S. J. & Norvig, P. (2001) Artificial Intelligence a modern approach Prentice Hall, 932p

EVALUATING OF FUTURE EVENTS BASED IN ARTIFICIAL NEURAL NETWORKS IN A CASE STUDY OF SALES OF FURNITURES

Abstract. *Accurate forecasting of different events is a key aspect for successful management with relevance in different fields including weather forecast, environmental evaluation of air pollution and evaluation of business strategies. In this context, recurrent neural networks, more specifically, LSTM (long-short term memory) have been successfully adopted in many time series or sequential data applications in traditional situations, as well as, in situations where the absence of data, outliers and nonlinearity reduces the applicability of traditional models like the autoregressive models. In this article, the use of LSTM recurrent artificial neural networks is evaluated for sales forecasting. Six different conditions of variables were evaluated to describe the LSTM artificial neural network models using r -scored (R^2) and mean squared error (MSE) as metrics to describe the quality of these models. Relative to LSTM, it was observed that the model evaluated was able to capture the data seasonality, describing the sales with r -scored superior to 0.7 and making a correct description of sales even after 24 months of training. Relative to description of sales, it was observed that the use*

of sales in the last three months was the best period to predict future sales and the use of sales of office items improved the quality of the model.

Keywords: *temporal series, artificial neural network, LSTM, consumer product, artificial intelligence, forecasting, future events.*