



UNIVERSIDADE FEDERAL DO TRIANGULO MINEIRO - UFTM
CURSO DE LICENCIATURA EM MATEMÁTICA

Análise de ativos financeiros por meio de Rede Neural Recorrente do tipo
LSTM

Vinícios Antônio Passos Balduino

Uberaba - MG
Outubro de 2022

Análise de ativos financeiros por meio de Rede Neural Recorrente do tipo LSTM

Vinícios Antônio Passos Balduino

Projeto de Pesquisa apresentado ao Curso de Licenciatura em Matemática da UFTM, como requisito parcial para obter o título de Licenciado em Matemática, sob a orientação do professor Dr. Osmar Aléssio.

Orientador: Prof. Dr. Osmar Aléssio

Uberaba - MG
Outubro de 2022

Resumo

Este projeto de pesquisa tem como proposito, o desenvolvimento de uma rede neural recorrente do tipo LSTM para a predição do preço de ativos do mercado financeiro por meio de séries temporais. Para isso nos apoiamos em teorias de mercado, como a hipótese de mercado adaptativo. Para uma ótima otimização da rede buscaremos entender os melhores parâmetros de configuração desta, bem como o seus desenvolvimentos matemático. Com a conclusão deste trabalho, pretende se que ele contribua para o avanço dos estudos em redes neurais.

Palavras-chave: Programação, Rede Neural Recorrente, Predição do Preço de Ativo.

Sumário

1	INTRODUÇÃO	6
2	JUSTIFICATIVA	8
3	OBJETIVOS GERAIS E ESPECÍFICOS	10
4	METODOLOGIA	11
5	CRONOGRAMA	15
	REFERÊNCIAS	16

Lista de tabelas

Tabela 1 – Cronograma	15
---------------------------------	----

1 Introdução

O mercado de ações de um país é parte importante para o desenvolvimento de sua economia, por meio deste, empresas podem ofertar a um preço que acham justo pequenas partes de seus negócios para pessoas físicas, ditas como investidores.

De acordo com Toledo (TOLEDO, 2006) e Dametto (DAMETTO, 2018), o mercado de ações pode ser dividido em dois segmentos, o mercado primário e o mercado secundário. Por meio do mercado primário as empresas buscam se financiar, através da oferta de valores mobiliários. Já no mercado secundário é o que trás liquides ao mercado primário. Por meio do mercado secundário a pessoa investidora tem acesso títulos de investimento de tais empresas.

A vantagem do desenvolvimento deste mercado se dá tanto para a empresa quanto para a pessoa investidora. Para a empresa a principal vantagem é o levantamento de capital, que a longo prazo é mais vantajoso comparando aos financiamentos bancários, principalmente em épocas de altas taxas de juros, como é o caso de 2022, ano de elaboração deste trabalho.

Já para a pessoa investidora, pela compra destes títulos de investimentos ela tem a oportunidade de se tornar sócio da empresa, o que por consequência trás diversas vantagens, como por exemplo, o recebimento de dividendos, a participação no conselho das empresas, a multiplicação de seus investimentos, dentre outros.

No entanto, não é simples a aplicação de recursos neste mercado, o investidor precisa saber exatamente o que deseja. Precisa alinhar, suas necessidades com suas condições e objetivos. Depois de ter tudo isso em mente, ele chegará em sua primeira conclusão, que é pensar em qual o seu principal plano para esse dinheiro. Depois disso, é sensato considerar uma análise de quais ativos atendem à esse plano para a melhor aplicação de seu recurso.

Após a elaboração do plano de investimento e a escolha dos principais ativos para o comporem, é preciso decidir quando comprar estes ativos. Essa decisão é importante pois, melhor do que comprar bons ativos, é comprá-los a bons preços. Esse momento do bom preço, tanto para comprar quanto para vender, é outro dos motivos que torna o investimento no mercado de ações complexo. Mesmo em um investimento de baixo risco é necessário levar em consideração essa volatilidade do preço.

Frente ao mercado de ações o investidor pode agir segundo alguns princípios, o que implica em uma estratégia de se comportar com o mercado. Uma das mais conservadoras e, se, fundamentada por uma boa análise, mais seguras, formas de investir é seguindo uma estratégia de *buy and hold* (comprar e segurar), que consiste em comprar ações de empresas acreditando em sua valorização no longuíssimo prazo. Ativos comprados seguindo esta estratégia podem ficar na carteira do investidor por anos.

Outra maneira de se portar no mercado e ações é seguindo uma estratégia de *trade* (troca), que consiste em comprar e vender ações em um prazo relativamente curto de tempo. Neste tipo de estratégia os ativos podem permanecer na carteira do investidor desde minutos até dias.

Durante o processo de compra de ações, a pessoa investidora está comprando um determinado número de ações a um determinado preço. Havendo oscilação neste preço, o número de ações que o investidor obteve se mantém o mesmo, no entanto o preço daquelas ações foi alterado. Esse movimento de preço é caracterizado pela volatilidade no preço do ativo.

Levando em consideração as duas maneiras de investir apresentadas anteriormente, podemos observar que considerar a volatilidade no preço do ativo na estratégia de *trade* é essencial, afinal de contas o lucro destas operações se faz por meio desta volatilidade. Já na estratégia de *hold*, de maneira equivocada pode se imaginar que não faça diferença essa volatilidade, basta comprar ações de empresas que ofereçam um serviço bom e perpétuo. No entanto, se pelo motivo que for, o *holder* (pessoa que opera seguindo a estratégia de *hold*) decide liquidar algum investimento, ele também está sujeito a perder dinheiro devido à volatilidade.

Analisar essa volatilidade talvez seja a parte mais complicada de escolher um bom ativo para investir. Por isso contar com técnicas avançadas pode nos auxiliar nessa decisão. Muitos estudos se desenvolveram nos últimos anos buscando aproximar a análise de ativos financeiros e tecnologias de informação, destes estudos cabe destacar o desenvolvimento por meio do uso de redes neurais artificiais (RNA).

O objetivo da RNA é simular o cérebro humano para tomada de decisões complexas, e seu funcionamento varia de acordo com o tipo de rede neural. Pesquisas atuais nesta área trabalham com redes neurais recorrentes (RNR) do tipo *long short-term memory* (*LSTM*).

Tendo em vista a importância de analisar a volatilidade do preço do ativo, este trabalho se desenvolve com o objetivo de estruturar uma rede neural recorrente do tipo LSTM, para a predição do preço de um índice da Bolsa de Valores Brasileira.

2 Justificativa

O mercado de ações pode ser interpretado de maneiras diferentes para cada economista, o que consequentemente influencia a maneira com que o investidor se planeja. Segundo Mussa (MUSSA A.; YANG, 2008) a hipótese de mercado eficiente (HME) surge em meados de 1970 publicada por Eugene F. Fama no Journal of Finance, Cambridge, tal teoria é significativa para o mercado até os dias atuais. E a sua proposta é descrever o mercado como eficiente, tendo sempre refletido em seu preço todas as informações referentes a ele, independente da forma de mercado, seja forte, fraca ou semi-forte.

De acordo com Carvalho (CARVALHO L. F.; GIACHERO, 2012) as formas de mercado se expressariam, sendo o mercado de forte ou eficiente, definido pela característica de que os preços refletem todas as informações existentes no mercado (históricas, públicas e privadas). Assim, nenhum investidor poderia obter retornos extraordinários, mesmo utilizando dados confidenciais ainda não publicados, pois o ajuste dos preços no mercado é instantâneo.

No mercado semiforte nenhum investidor pode obter retornos anormais (em excesso) com base em alguma informação pública disponível, pois os preços se ajustam rapidamente às novas informações divulgadas. Neste caso, teriam lucros os investidores que abusassem de informações privilegiadas, o que é considerado crime. Está previsto na lei nº6.385/76 Art. contando com pena de reclusão de 1 a 5 anos e multa de até 3 vezes o montante da vantagem ilícita.

Já no mercado fraco ou ineficiente, nenhum investidor pode obter retornos anormais (em excesso) baseados na expectativa de que preços passados são bons sinalizadores dos preços futuros. Ou seja, informações de preços passados, não são úteis ou relevantes para atingir retornos em excesso. Nesta definição de eficiência, o lucro somente aconteceria com base nas informações públicas dos ativos e com isso poderia ser descartado o trabalho de analisar preços das ações.

De acordo com Gomes (GOMES L. F. DE A.; SOUZA, 2019) a teoria de hipótese de mercado eficiente trata o ser humano como "Homo Economics" e este tomaria sempre as melhores decisões frente ao mercado financeiro. O que na prática pode nem sempre ser verdade, pois essa forma de pensar ignora no ser humano suas decisões com base na emoção. Em contrapartida, como aponta Dourado (DOURADO G. DE A.; TABAK, 2014), Lo (2004) intersecta a psicologia humana à economia, e propõe a Hipótese de Mercado Adaptativo (HMA), que tem por objetivo verificar a previsibilidade do mercado por ciclos, ou seja, entre eventos semelhantes, poderia haver a possibilidade do mercado repetir a iniciativa tomada anteriormente, o que garantiria uma certa previsibilidade.

Com o avanço tecnológico que temos hoje na sociedade é possível desenvolvermos

sistemas complexos de tomadas de decisão automáticas. Estes sistemas de tomadas de decisão podem ser baseados em redes neurais artificiais. Atualmente há diversos trabalhos relacionando RNAs e análise de ativos do mercado financeiro, e se tem provado que estas são ferramentas interessantes para a tomada de decisão do investidor. Com vários estudos como este de diversas áreas diferentes, há um grande volume de conteúdo a ser consumido, porém por ser de áreas tão abrangentes, tende a ser bem diversificado. Durante as pesquisas para este trabalho, foi observado que são poucos os trabalhos acadêmicos na língua portuguesa que abordam os mesmos conceitos trabalhados aqui com ênfase na matemática por trás das redes neurais, mais especificamente da rede neural do tipo LSTM.

3 Objetivos Gerais e Específicos

O objetivo geral deste trabalho é o desenvolvimento de uma rede neural recorrente do modelo *Long Short Term Memory* (LSTM), para a predição de série temporal financeira com o intuito de prever o preço futuro de uma determinada ação. Com o intuito de encontrar uma melhor configuração possível para o problema, trabalharemos com várias configurações da rede e procuraremos entender qual a função de determinada configuração e a base matemática por trás dela.

Os objetivos específicos podem ser desenvolvidos da seguinte maneira:

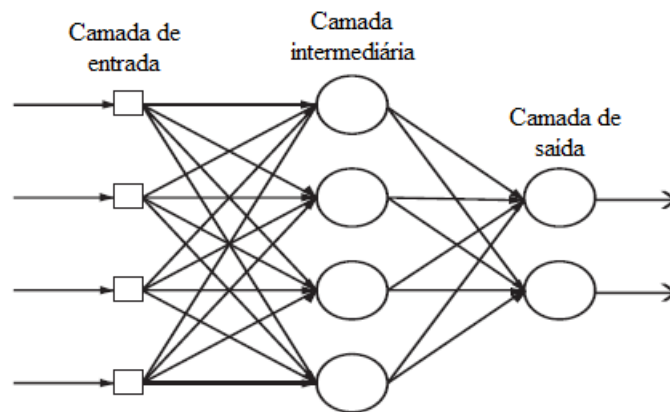
- Estudar e desenvolver modelos de redes neurais recorrentes.
- Estudar teorias do mercado de ações, como hipótese do mercado adaptativo e hipótese do mercado eficiente.
- Estudar a matemática por trás das séries temporais, do tratamento de dados para o Machine Learning e da estrutura das redes neurais, especificamente da rede LSTM.

4 Metodologia

Para o desenvolvimento dos objetivos deste trabalho, será necessário entendermos um pouco do funcionamento das redes neurais artificiais. Uma RNA tem seu funcionamento inspirado pelos neurônios do nosso cérebro, em outras palavras, podemos dizer que esta seria a interpretação matemática do funcionamento do cérebro. Logo, por meio delas seria possível para o computador aprender por exposição a situações problemas e ter suas conclusões de determinada situação.

A maneira mais simples de enxergarmos uma RNA, seria montarmos uma cadeia de neurônios de entrada e uma cadeia de neurônios de saída, da forma que, todos os neurônios da cadeia de entrada estivessem ligados a cada neurônio da cadeia de saída. Estas cadeias são ditas como camadas, e as ligações são sinapses, cada ligação tem a função de atribuir um peso a cada neurônio da camada de entrada. Para além do exemplo, também há a possibilidade de haver mais camadas, que são ditas camadas intermediárias ou camadas ocultas, estas possuem a função de auxiliar a distribuição de pesos para cada informação da entrada.

Figura 1 - Rede neural artificial não recorrente e multicamadas.



Fonte: ResearchGate, 2001

Analisando matematicamente essa operação de atribuição de peso por entrada feita pela rede neural, podemos perceber uma operação de multiplicação de matrizes, na qual a entrada é interpretada como uma matriz $I = (e_{1j})_{1 \times N}$ e os pesos uma matriz $P = (p_{i1})_{N \times 1}$ (cujos dados nós desconhecemos), resultando em uma matriz $R = (r_{11})_{1 \times 1}$.

$$I = \begin{bmatrix} i_{11} & i_{12} & \dots & i_{1n} \end{bmatrix} \times P = \begin{bmatrix} p_{11} \\ p_{21} \\ \vdots \\ p_{n1} \end{bmatrix} = R = \begin{bmatrix} r_{11} \end{bmatrix}$$

E de acordo com a multiplicação de matrizes, definido na forma convencional.

$$r_{11} = i_{11} \times p_{11} + i_{12} \times p_{21} + \dots + i_{1n} \times p_{n1}.$$

Quando se inicia uma rede neural ela precisa passar por um processo de treinamento, visto que ela não conhece os pesos ideais para cada entrada. Para isso deverá existir um conjunto de dados, composto pelas entradas e os resultados. Parte destes dados, geralmente entre 60% e 70%, deverão ser expostos à rede para que ela possa equilibrar seus pesos, para este conjunto de dados é atribuído o nome de conjunto de treinamento ou dados de treino. Os demais entre 30% e 40 % (dependendo do tamanho do seu conjunto de treino), dito conjunto de teste ou dados de teste, servirão de teste para avaliar a precisão da rede. Ambos conjuntos possuem dados de entrada e também dados de saída que são passados à rede.

Com base no que discutimos até o momento, podemos nos perguntar como é formada a nossa matriz de pesos, uma vez que, os únicos dados que o computador recebe são as entradas. A parte responsável dentro da RNA por controlar os pesos da rede, é a chamada função de perda, ou *loss*. Um dos métodos mais usados para a análise de dados, é o método do erro quadrático médio.

Conforme a rede vai fazendo as interações que seria a previsão de todo o seu conjunto de treinamento, no caso do método do erro quadrático médio, ela faz o somatório do quadrado da diferença entre o resultado real e o resultado previsto, termo a termo, ou seja, ela recebe um dado de entrada e a sua saída real e faz uma previsão para aquela saída, após isso a rede pega o valor real, subtrai o valor da previsão e eleva ao quadrado, após fazer isso com todos os termos, a rede faz o somatório destes resultados, quanto mais próximo de zero esse somatório for significa que mais assertiva é a nossa rede. A cada interação a rede alterará algum peso e terá um novo resultado mais próximo de zero da função de perda. A quantidade de interações que a rede fará para obter a menor taxa de perda, é programada pela pessoa que montar a rede, no entanto, uma quantidade excessiva de interações não implica na excelente configuração da rede, isso quer dizer que, em um determinado momento a rede tomará a melhor decisão possível com base no conjunto de dados fornecido, logo não terá mais o que ela aprender com aquele conjunto, então a quantidade excessiva de interações apenas estará gastando capacidade computacional.

A partir deste momento que a rede se estabiliza em uma taxa de erro, com todos os pesos da matriz P bem definidos, com isso poderemos obter bons resultados ou maus resultados.

Caso a rede atinja um bom resultado, isso significa que no decorrer das interações, a rede conseguiu equilibrar a matriz de peso, e as últimas previsões se mostraram satisfatórias, quando fornecermos os dados de teste também teremos resultados satisfatórios, concluindo que a rede aprendeu com os dados informados e está pronta para receber novos dados e apresentar boas previsões.

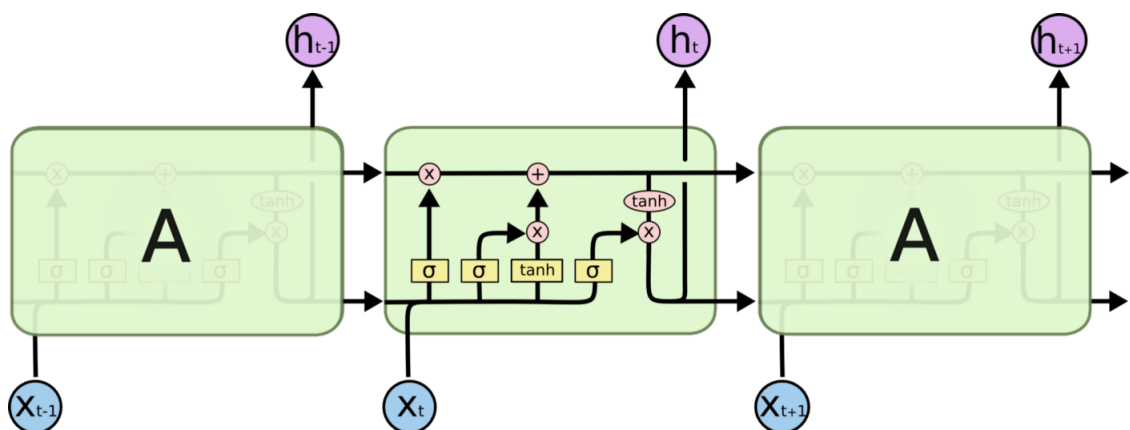
Já no caso da rede não atingir um bom resultado podemos classificar os problemas como sendo de sub ajustamento (*underfitting*) ou sobre ajustamento (*overfitting*).

- Underfitting: Neste caso o modelo não conseguiu aprender com os dados, logo, teremos uma grande taxa de erro tanto no conjunto de treino como no conjunto de teste
- Overfitting: Neste caso o modelo viciou nos dados do conjunto de treino, isso quer dizer que, para este conjunto o modelo apresenta excelente predição, porém quando inseridos os dados de teste ou ainda novos dados, não teremos previsão satisfatória.

Este trabalho se desenvolverá utilizando uma rede neural do tipo LSTM, esta por sua vez é uma rede neural recorrente, logo uma RNR, diferente da RNA onde a única entrada são os dados fornecidos. A RNR pode receber como entrada, dados da camada de saída, possibilitando à rede maior exposição a dados, o que pode contribuir para melhor configuração de pesos.

Graças à sua arquitetura, a rede LSTM se diferencia bastante das redes neurais convencionais, não possuindo camadas de neurônios e sim, células de memória. Abaixo segue um exemplo mais comum de se retratar uma rede LSTM.

Figura 2 - Rede neural recorrente do tipo LSTM



Fonte: Medium, 2019

Cada célula de memória desta está conectada com a sua anterior e posterior. A linha de cima é o estado da memória, fazendo uma analogia às redes neurais convencionais, ali

seria a matriz de pesos que o neurônio atribui a cada variável. Podemos observar que essa linha está conectada com todos os demais processos da célula, logo cada informação que a célula recebe e devolve influencia e é influenciada por ela.

Abaixo da célula temos a entrada de dados, que logo passa pelo primeiro portão, o portão de esquecimento. Neste portão é escolhido a informação que será esquecida do estado da memória. Logo após temos o portão de entrada, neste portão é selecionado a informação que será armazenada no estado da memória e o quanto elas influenciarão neste estado. Por fim temos o portão de saída, como o próprio nome diz, é a conclusão do processo de pensamento da célula de memória, neste momento a informação que entrou lá no começo da célula já sofreu todas as distribuições de peso conforme o estado da célula atual e se estivermos no final dos processos da rede receberemos o valor da predição.

O trabalho se desenvolverá da seguinte maneira, utilizaremos do ambiente de desenvolvimento Python do Google Colaboratory, um serviço gratuito e disponível para quem possui e-mail da google. Para configurar a rede neural utilizaremos das bibliotecas Keras, TensorFlow e SKlearn, para o tratamento de dados e visualização, será utilizada as bibliotecas Pandas, Matplotlib e Numpy. Para a extração dos dados utilizaremos a biblioteca yfinance, cuja função é extrair os dados diretamente do site yahoo finance, site ao qual pertence a biblioteca. Ainda não foi decidido qual ação iremos utilizar para fazermos as predições, no entanto será uma ação da bolsa de valores brasileira. É parte fundamental deste trabalho testar diferentes configurações da rede e propor alterações no que for possível objetivando a melhora significativa da rede, buscando sempre por uma explicação matemática para justificar tais melhorias.

5 Cronograma

2*Atividades	Meses											
	1	2	3	4	5	6	7	8	9	10	11	12
Revisão bibliográfica	x	x	x									
Estudo dos modelos de redes neurais		x	x	x	x							
Estudar dos algoritmos dos modelos			x	x	x	x	x	x				
Programação e teste dos modelos				x	x	x	x	x				
Elaboração do relatório final									x	x	x	x
Apresentação do TCC												x

Tabela 1 – Cronograma

Referências

CARVALHO L. F.; GIACHERO, O. S. R. K. C. S. Uma crítica a hipótese da eficiência de mercado: Análise dos casos ambev e submarino. *Simpósio de Excelência em Gestão e Tecnologia - SEGET*, 2012. Acessado em: 27/09/2022. Disponível em: <https://www.aedb.br/seget/arquivos/artigos07/1149_1149_Artigo3_Karem_Seget_Autores.pdf>.

DAMETTO, R. C. *Estudo da aplicação de redes neurais artificiais para predição de séries temporais financeiras*. Dissertação (Mestrado) — UNESP "Júlio de Mesquita Filho- Programa de Pós-Graduação em Engenharia de Produção, Bauru-SP, 2018.

DOURADO G. DE A.; TABAK, B. M. Teste da hipótese de mercados adaptativos para o brasil. *Revista Brasileira de Finanças*, v. 12, n. 4, p. 517, 2014. Acessado em: 01/08/2022. Disponível em: <<http://www.spell.org.br/documentos/ver/39858/teste-da-hipotese-de-mercados-adaptativos-para-o-brasil->>.

GOMES L. F. DE A.; SOUZA, P. V. S. d. S. C. A. T. Hipótese dos mercados adaptativos: uma análise da eficiência de títulos do mercado acionário brasileiro. *Revista de Contabilidade do Mestrado em Ciências Contábeis da UERJ*, v. 24, n. 3, p. 25, 2019. Acessado em: 01/08/2022. Disponível em: <<https://www.e-publicacoes.uerj.br/index.php/rcmccuerj/article/view/52255/34106>>.

MUSSA A.; YANG, E. T. R. F. R. Hipótese de mercados eficientes e finanças comportamentais - as discussões persistem. *Revista FACEF pesquisa*, v. 11, n. 1, 2008. Acessado em: 01/08/2022. Disponível em: <<https://periodicos.unifacef.com.br/index.php/facefpesquisa/article/view/109>>.

TOLEDO, C. S. D.

A importância do mercado de ações para o crescimento econômico do Brasil — UFSC - Curso de Ciências Econômicas, Florianópolis, 2006.