



Universidade Estadual de Campinas
Instituto de Matemática, Estatística e Computação
Científica
Departamento de Matemática Aplicada

Métodos de Aprendizado de Máquinas para um Estudo Comparativo do Mercado de Ações Brasileiro

Palavras-chave — Mercado financeiro, aprendizado de máquinas, *deep learning*.

Aluna: Jackeline Leme Gregório

Orientador: João Florindo

Julho de 2023

1 Resumo do Projeto Proposto

O projeto propõe-se a investigar modelos de Aprendizagem de Máquinas [2], [1], como Regressão Logística, Florestas de Decisão Aleatórias, SVM (*Support Vector Machine*), Redes Neurais [11] e LSTM (*Long Short Term Memory*), visando comparar o desempenho destes modelos na tarefa de previsão de preços no mercado de ações (se a ação irá aumentar ou diminuir) [14], [13].

O mercado de ações sofre constantes mudanças [12], [7], o que nos leva a concluir que há grande dificuldade para a escolha de uma ação promissora e mais ainda para uma previsão de qual direção o seu preço tomará. Trata-se de uma das maneiras mais complicadas e sofisticadas de se fazerem negócios e depende de uma grande quantidade de dados. Portanto, prever o mercado de ações (B3) [?], para que assim os investidores usem seu dinheiro para obter lucro é um grande desafio. Nos últimos anos, as técnicas de aprendizado de máquinas têm sido cada vez mais estudadas para tentar prever o mercado com mais exatidão, fugindo das abordagens tradicionais.

Tem-se estudado e pesquisado cada vez mais os modelos de aprendizagem de máquinas aplicado a ações e/ou bolsa de valores. Visto que houve aumento da aplicação de técnicas de *deep learning* em ações, o intuito neste projeto é mostrar que modelos clássicos de aprendizado de máquinas (não *deep*) também podem obter bons resultados e boas previsões.

2 Objetivos

As operações do mercado de ações são complexas e muitas vezes é difícil tomar a decisão correta de qual método utilizar para se obter o melhor desempenho. O objetivo desta pesquisa é fornecer contribuições para a comunidade científica, e para o desenvolvimento de novos procedimentos de Aprendizado de Máquinas que possam ser aplicados ao mercado de ações para auxílio de decisão, analisando alguns métodos, qual teria a melhor previsão e menor risco para investidores e empresas. É esperado que, após o estudo e análise comparativa dos métodos, seja possível comparar os diferentes métodos e sua respectiva precisão, para que assim sejam feitas as análises corretas das informações, que possam ser amplificadas e aperfeiçoadas para uso. Este projeto tem como objetivo realizar um estudo comparativo de métodos de aprendizado de máquinas no mercado de ações, visando especificamente prever a direção seguida por determinado ativo com base em seu histórico passado. Serão estudados em particular regressão logística [5], redes neurais [9], máquinas de vetores de suporte (SVM) [8], florestas aleatórias [5] e *long short term memory* (LSTM).

3 Metodologia Proposta

IBOV B3: O índice Bovespa é o indicador mais importante e o principal da Bolsa Paulista, calcula a média de desempenho de uma carteira com as principais ações negociadas na B3. São calculados em tempo real e seu resultado é baseado na cotação de cada ação que integra o índice, multiplicado pela quantidade teórica dos ativos que compõem a carteira. Analisando seu gráfico e acompanhando sua cotação diariamente, é possível ver como está a economia do país, a saúde das empresas e se vale a pena ou não negociar as ações.

4 Modelos de Aprendizado de Máquinas

O primeiro algoritmo a ser utilizado para prever se temos que comprar ou não uma ação, foi a Regressão Logística, o mais simples a ser implementado. A biblioteca *yfinance* do Python, conhecido como Yahoo Finance, foi usada para obter dados do IBOVESPA (BVSP). Foi escolhida tal biblioteca pois, além da praticidade, é fácil pesquisar a ação que desejamos, é possível colocar um intervalo de tempo e, tudo isso apenas colocando o *ticker*, e então temos como resultado as datas, abertura da ação, fechamento, o menor valor do dia, o maior valor, o preço de fechamento ajustado e o volume. Outro ponto positivo para o uso dessa biblioteca é que os dados já vêm tratados, limpos e atualizados, o que evita que aconteçam erros durante a análise ou que exija maior tempo porque teria que fazer a limpeza dos dados.

A acurácia ficou em 51,1%. Obtivemos alguns resultados preliminares acerca do modelo e na Tabela 1 podemos ver os resultados.

Tabela 1: Resultados do modelo de Regressão Logística para ações do BVSP.

	precision	recall	f1-score	support
-1	0.51	0.62	0.56	826
1	0.49	0.37	0.43	806

A resposta -1 nos dá os resultados em que as ações decresceram e o 1 quando as ações cresceram. A *precision* mede a quantidade de vezes que o modelo acerta em relação ao total de vezes que ele faz uma predição. Já o *recall* nos traz a informação da quantidade de vezes que o modelo acerta em relação ao total de vezes que ele deveria ter acertado. Por último, o *f1-score* é a métrica que combina precisão e *recall* de forma equilibrada, é a média harmônica entre ambos.

Após isso, temos os resultados utilizando o algoritmo Máquina de Vetores de Suporte (SVM), no qual conseguimos obter uma acurácia de 51%, nominalmente maior comparado à Regressão Logística, mas estatisticamente equivalente.

Tabela 2: Primeiro resultado do modelo SVM para ações do BVSP.

	precision	recall	f1-score	support
-1	0.26	0.01	0.03	673
1	0.52	0.96	0.67	738

A fim de melhorar o resultado, desenvolvi outro modelo de SVM. A seguir, o resultado que obtive.

Tabela 3: Primeiro resultado do modelo SVM para ações do BVSP.

	precision	recall	f1-score	support
-1	0.55	0.08	0.14	322
1	0.49	0.93	0.64	308

A acurácia foi de: 49,84%, recall :09935 e F1 score: 0.6595.

Tabela 4: Resultados do modelo de Redes Neurais para ações do BVSP.

	precision	recall	f1-score	support
0	0.57	0.18	0.28	65
1	0.60	0.87	0.72	87

Tivemos uma acurácia de: 49,68%, recall: 09318 e F1 score: 0.6442.

Tabela 5: Resultados do modelo de LSTM para ações do BVSP.

	precision	recall	f1-score	support
0	0.62	0.08	0.14	65
1	0.58	0.97	0.73	87

Para LSTM obtivemos os melhores resultados. Para acurácia temos: 58,55%, recall: 09655 e F1 score: 0.7273.

Tabela 6: Resultados do modelo de Floresta Aleatória para ações do BVSP.

	precision	recall	f1-score	support
0	0.67	0.09	0.16	65
1	0.59	0.97	0.73	87

Por último, temos bons resultados utilizando Floresta Aleatória em um intervalo de 10 anos. A acurácia: 57,89%, recall: 09540 e F1 score: 0.7217.

5 Conclusão

A previsão do movimento de ações é uma tarefa extremamente desafiadora devido à natureza complexa e imprevisível dos mercados financeiros. O comportamento dos preços das ações é influenciado por inúmeras variáveis, muitas das quais podem ser difíceis de capturar com precisão [4].

Após utilizar diferentes modelos para prever o mercado, foi observado que os modelos LSTM e Floresta Aleatória [6] apresentaram os melhores resultados. Isso sugere que modelos de séries temporais, como o LSTM, são mais adequados para capturar padrões temporais e tendências nos dados históricos de preços das ações.

Os resultados mais satisfatórios obtidos pelos modelos LSTM e Floresta Aleatória podem ser atribuídos a suas respectivas capacidades de modelar relações complexas e não lineares entre as variáveis de entrada e a variável de destino (movimento da ação). A LSTM [10] é capaz de capturar padrões temporais de longo prazo nos dados de séries temporais, enquanto a Floresta Aleatória pode lidar bem com características não lineares nos dados.

É importante ressaltar que mesmo com resultados promissores, prever o movimento de ações com alta precisão é uma tarefa extremamente difícil e incerta. Os mercados financeiros são influenciados por eventos imprevisíveis, notícias e fatores econômicos, tornando a tarefa de previsão de preços de ações uma verdadeira incógnita.

Portanto, é fundamental lembrar que o investimento em ações deve ser feito com cuidado, além de ser baseado em análises mais abrangentes e especializadas,

em vez de depender unicamente de modelos de previsão.

Referências

- [1] ALPAYDIN, E. *Introduction to Machine Learning, (Adaptive Computation and Machine Learning)*. almohreraladbi, 2004.
- [2] ALZUBI, J., NAYYAR, A., AND KUMAR, A. Machine learning from theory to algorithms: an overview. In *Journal of physics: conference series* (2018), vol. 1142, IOP Publishing, p. 012012.
- [3] CHASSAGNON, G., VAKALOPOLOU, M., PARAGIOS, N., AND REVEL, M.-P. Deep learning: definition and perspectives for thoracic imaging. *European radiology* 30, 4 (2020), 2021–2030.
- [4] DIXON, M. F., HALPERIN, I., AND BILOKON, P. *Machine learning in Finance*, vol. 1170. Springer, 2020.
- [5] HASTIE, T., TIBSHIRANI, R., FRIEDMAN, J. H., AND FRIEDMAN, J. H. *The elements of statistical learning: data mining, inference, and prediction*, vol. 2. Springer, 2009.
- [6] JANSEN, S. *Machine Learning for Algorithmic Trading: Predictive models to extract signals from market and alternative data for systematic trading strategies with Python*. Packt Publishing Ltd, 2020.
- [7] LEVINE, R. Stock markets: a spur to economic growth. *Finance & Development* 33, 001 (1996).
- [8] NALEPA, J., AND KAWULOK, M. Selecting training sets for support vector machines: a review. *Artificial Intelligence Review* 52, 2 (2019), 857–900.
- [9] NIELSEN, M. A. *Neural networks and deep learning*, vol. 25. Determination press San Francisco, CA, USA, 2015.
- [10] OLAH. Understanding lstm networks. [urlhttps://colah.github.io/posts/2015-08-Understanding-LSTMs/](https://colah.github.io/posts/2015-08-Understanding-LSTMs/), 2015.
- [11] PANG, X., ZHOU, Y., WANG, P., LIN, W., AND CHANG, V. An innovative neural network approach for stock market prediction. *The Journal of Supercomputing* 76, 3 (2020), 2098–2118.
- [12] SHAH, D., ISAH, H., AND ZULKERNINE, F. Stock market analysis: A review and taxonomy of prediction techniques. *International Journal of Financial Studies* 7, 2 (2019), 26.
- [13] SIEGEL, J. J. *Stocks for the long run: The definitive guide to financial market returns & long-term investment strategies*. McGraw-Hill Education, 2021.
- [14] SUNO. Tudo sobre a bolsa de valores: saiba como funciona a b3. [urlhttps://www.suno.com.br/guias/bolsa-de-valores/](https://www.suno.com.br/guias/bolsa-de-valores/), Jan. 2022.