

UNIVERSIDADE DE SÃO PAULO

Instituto de Ciências Matemáticas e de Computação

Experimentos com modelos de aprendizado de máquina para a previsão de preço de ação de empresa de energia elétrica na B3

Luiz Antonio Vieira Candido da Silva

Monografia - MBA em Inteligência Artificial e Big Data

SERVIÇO DE PÓS-GRADUAÇÃO DO ICMC-USP

Data de Depósito:

Assinatura: _____

Luiz Antonio Vieira Candido da Silva

Experimentos com modelos de aprendizado de máquina para a previsão de preço de ação de empresa de energia elétrica na B3

Monografia apresentada ao Departamento de Ciências de Computação do Instituto de Ciências Matemáticas e de Computação, Universidade de São Paulo - ICMC/USP, como parte dos requisitos para obtenção do título de Especialista em Inteligência Artificial e Big Data.

Área de concentração: Inteligência Artificial

Orientadora: Profa. Dra. Marislei Nishijima

Versão original

São Carlos

2024

Ficha catalográfica elaborada pela Biblioteca Prof. Achille Bassi
e Seção Técnica de Informática, ICMC/USP,
com os dados inseridos pelo(a) autor(a)

S586e Silva, Luiz Antonio Vieira Candido da
Experimentos com modelos de aprendizado de
máquina para a previsão de preço de ação de empresa
de energia elétrica na B3 / Luiz Antonio Vieira
Candido da Silva; orientadora Marislei Nishijima. --
São Carlos, 2024.
59 p.

Trabalho de conclusão de curso (MBA em
Inteligência Artificial e Big Data) -- Instituto de
Ciências Matemáticas e de Computação, Universidade
de São Paulo, 2024.

1. Previsão do preço de ações. 2. Aprendizado de
máquina. 3. Support Vector Regression - SVR. 4.
Long Short-Term Memory - LSTM. 5. Convolutional
Neural Network - CNN. I. Nishijima, Marislei,
orient. II. Título.

Luiz Antonio Vieira Candido da Silva

**Experiments with machine learning models for predicting
the share price of an electric power company on B3**

Monograph presented to the Departamento de Ciências de Computação do Instituto de Ciências Matemáticas e de Computação, Universidade de São Paulo - ICMC/USP, as part of the requirements for obtaining the title of Specialist in Artificial Intelligence and Big Data.

Concentration area: Artificial Intelligence

Advisor: Profa. Dra. Marislei Nishijima

Original version

São Carlos

2024

Este trabalho é dedicado à minha família por todo o apoio, amor e inspiração.

AGRADECIMENTOS

Em primeiro lugar, agradeço a Deus por tudo. Em especial pela inspiração que me levou a matricular-me neste curso e pela força que me permitiu concluí-lo.

Agradeço à minha família por estarem presentes em minha vida, especialmente pelo incentivo e pela paciência durante todo o tempo em que estive dedicado a este curso.

Expresso minha sincera gratidão à minha orientadora, Prof^a. Dr^a. Marislei Nishijima, pelo apoio, pelas informações e pelos direcionamentos fundamentais para a elaboração deste trabalho.

Por fim, agradeço a todas as professoras e professores do MBA - Inteligência Artificial e Big Data pelo conhecimento transmitido e pela atenção dedicada às necessidades dos alunos ao longo do curso.

*“O estudo, a busca da verdade e da beleza são domínios
em que nos é consentido sermos crianças por toda a vida.”*

Albert Einstein

RESUMO

SILVA, L. A. V. C. **Experimentos com modelos de aprendizado de máquina para a previsão de preço de ação de empresa de energia elétrica na B3.** 2024. 59 p. Monografia (MBA em Inteligência Artificial e Big Data) - Instituto de Ciências Matemáticas e de Computação, Universidade de São Paulo, São Carlos, 2024.

O setor de geração e distribuição de energia elétrica é um dos mais relevantes e estratégicos de um país, sendo base para o funcionamento de todos os setores produtivos da economia, incluindo a indústria, o comércio, os serviços e os transportes. Por conta disso, as empresas desse setor são muito valorizadas e atraem grande interesse dos investidores do mercado de ações, que fazem análises e buscam ferramentas para antever a evolução do preço desse ativo. Entretanto, essa é uma tarefa muito complexa, devido aos diversos fatores que influenciam esse valor ao longo do tempo, como as condições econômicas do país, as políticas públicas adotadas, o desempenho das empresas e as expectativas dos investidores. Nos últimos anos, técnicas de Inteligência Artificial e Big Data estão sendo elaboradas e experimentadas para auxiliar os investidores nessa difícil tarefa, através da identificação de padrões nos dados históricos e, com base nisso, geração de previsões para o futuro. Nesse trabalho foi elaborado um dataset com os preços da ação ELET3 (ELETROBRAS ON N1) nos pregões da B3 durante um período de 10 anos (preço de abertura, preço mínimo, preço máximo e preço de fechamento), além de atributos econômicos (Inflação – IPCA, PIB do Brasil, cotação do dólar americano frente ao real brasileiro e outros).. Foram empregados quatro modelos de Inteligência Artificial com o objetivo de preverem o preço de abertura da ação nos pregões da B3, a saber, Support Vector Regression – SVR, Long Short-Term Memory – LSTM, Convolutional Neural Networks – CNN e híbrido CNN-LSTM, sendo os três últimos modelos classificados como de Aprendizado Profundo. O dataset da ação ELET3 (ELETROBRAS ON N1) foi aplicado igualmente nos quatro modelos, com o objetivo de identificar qual dos modelos apresentaria a melhor eficiência na tarefa de previsão do preço da ação. Diversos experimentos foram executados em cada um desses modelos. Em cada experimento, os principais parâmetros de cada modelo foram ajustados, até que apresentassem o melhor resultado de previsão. A eficiência dos modelos na tarefa de previsão do preço da ação foi mensurada através dos indicadores Mean Square Error (MSE), Root Mean Square Error (RMSE), Mean Absolute Error (MAE) e Mean Absolute Percentage Error (MAPE). Os resultados mostraram que os quatro modelos apresentaram bom desempenho na tarefa de previsão do preço da ação ELET3 (ELETROBRAS ON N1). Mas observou-se que o modelo Support Vector Regression (SVR), com um resultado do MAPE de 1.0903779, apresentou melhor eficiência que os outros três modelos de Aprendizado Profundo, com o modelo Long Short-Term Memory (LSTM) em segundo lugar (MAPE = 1.2068616), o híbrido CNN-LSTM em terceiro (MAPE =

1.9358492) e o Convolutional Neural Networks (CNN) com o menor desempenho (MAPE de 2.3128867). Em contrapartida, nos sete trabalhos correlatos que foram citados, os modelos de Aprendizado Profundo apresentaram a melhor eficiência nos experimentos de previsão do preço de ações. Isso aponta a complexidade dessa tarefa, que envolve a característica dos dados a serem analisados e o modelo de Inteligência Artificial utilizado, tanto na sua estrutura como no ajuste de parâmetros feitos durante os experimentos.

Palavras-chave: previsão do preço de ações; bolsa de valores; B3; modelo de inteligência artificial; aprendizado de máquina; rede neural artificial; regressão de vetores de suporte; SVR; long short-term memory; LSTM; rede neural convolucional; CNN.

ABSTRACT

SILVA, L. A. V. C. **Experiments with machine learning models for predicting the share price of an electric power company on B3.** 2024. 59 p. Monograph (MBA in Artificial Intelligence and Big Data) - Instituto de Ciências Matemáticas e de Computação, Universidade de São Paulo, São Carlos, 2024.

The electricity generation and distribution sector is one of the most important and strategic in a country, and is the basis for the operation of all productive sectors of the economy, including industry, commerce, services and transportation. As a result, companies in this sector are highly valued and attract great interest from stock market investors, who conduct analyses and seek tools to predict the evolution of the price of this asset. However, this is a very complex task, due to the various factors that influence this value over time, such as the country's economic conditions, the public policies adopted, the performance of companies and the expectations of investors. In recent years, Artificial Intelligence and Big Data techniques have been developed and tested to assist investors in this difficult task, by identifying patterns in historical data and, based on this, generating forecasts for the future. In this work, a dataset was prepared with the prices of the ELET3 (ELETROBRAS ON N1) stock in the B3 trading sessions over a period of 10 years (opening price, minimum price, maximum price and closing price), in addition to economic attributes (Inflation - IPCA, Brazil's GDP, exchange rate of the US dollar against the Brazilian real and others). Four Artificial Intelligence models were used to predict the opening price of the stock in the B3 trading sessions, namely, Support Vector Regression - SVR, Long Short-Term Memory - LSTM, Convolutional Neural Networks - CNN and hybrid CNN-LSTM, the last three models being classified as Deep Learning. The ELET3 (ELETROBRAS ON N1) stock dataset was applied equally in the four models, with the objective of identifying which of the models would present the best efficiency in the task of predicting the stock price. Several experiments were performed on each of these models. In each experiment, the main parameters of each model were adjusted until they presented the best prediction result. The efficiency of the models in the task of predicting the stock price was measured through the indicators Mean Square Error (MSE), Root Mean Square Error (RMSE), Mean Absolute Error (MAE) and Mean Absolute Percentage Error (MAPE). The results showed that the four models performed well in the task of predicting the price of the ELET3 (ELETROBRAS ON N1) stock. However, it was observed that the Support Vector Regression (SVR) model, with a MAPE result of 1.0903779, presented better efficiency than the other three Deep Learning models, with the Long Short-Term Memory (LSTM) model in second place (MAPE = 1.2068616), the hybrid CNN-LSTM in third place (MAPE = 1.9358492) and the Convolutional Neural Networks (CNN) with the lowest performance (MAPE of 2.3128867). On the other hand, in the seven related works that were cited,

the Deep Learning models presented the best efficiency in the stock price prediction experiments. This points to the complexity of this task, which involves the characteristics of the data to be analyzed and the Artificial Intelligence model used, both in its structure and in the parameter adjustment made during the experiments.

Keywords: stock price prediction; stock exchange; B3; artificial intelligence model; machine learning; artificial neural network; support vector regression; SVR; long short-term memory; LSTM; convolutional neural network; CNN.

LISTA DE FIGURAS

Figura 1 – Neurônio artificial. Adaptado de (Silva; Spatti; Flauzino, 2010)	31
Figura 2 – Elemento da Rede Neural. Adaptado de (Olah <i>et al.</i> , 2015)	34
Figura 3 – Diagrama de uma Rede Neural Recorrente (RNR). Adaptado de (Olah <i>et al.</i> , 2015)	34
Figura 4 – Diagrama de uma rede LSTM. Adaptado de (Olah <i>et al.</i> , 2015)	35
Figura 5 – Diagrama de uma Rede Neural Convolucional. Adaptado de (Artasan- chez; Joshi, 2020)	38
Figura 6 – Etapas de execução do trabalho (elaborado pelo autor)	43
Figura 7 – Cotações de abertura (PREABE) da ação ELET3 (em R\$) no período de amostragem (elaborado pelo autor)	44
Figura 8 – Resultados na fase de Teste do modelo SVR (elaborado pelo autor) . .	49
Figura 9 – Resultados na fase de Teste do modelo LSTM (elaborado pelo autor) .	50
Figura 10 – Resultados na fase de Teste do modelo CNN (elaborado pelo autor) . .	51
Figura 11 – Resultados na fase de Teste do modelo CNN-LSTM (elaborado pelo autor)	52

LISTA DE TABELAS

Tabela 1	–	Ações do índice IEE	29
Tabela 2	–	Bibliotecas Python utilizadas nos modelos do experimento.	43
Tabela 3	–	Características dos dados coletados.	44
Tabela 4	–	Fontes dos dados coletados.	45
Tabela 5	–	Resultados obtidos pelos modelos	53

LISTA DE ABREVIATURAS E SIGLAS

2D	Duas dimensões
AM	Aprendizado de máquina
B3	Bolsa de valores brasileira sediada na cidade de São Paulo
BACEN	Banco Central
BM&F	Bolsa de Mercadorias e Futuros
BOVESPA	Bolsa de Valores de São Paulo
CETIP	Central de Custódia e Liquidação Financeira de Títulos Privados
CNN	Convolutional Neural Network
CVM	Comissão de Valores Mobiliários
ELET3	Ação da empresa Eletrobras do tipo ON N1
EPE	Empresa de Pesquisa Energética
IA	Inteligência artificial
IBGE	Instituto Brasileiro de Geografia e Estatística
IBOVESPA	Índice geral do mercado de ações da B3
IEE	Índice do mercado de ações do setor de energia elétrica da B3
LSTM	Long Short-Term Memory
MAE	Indicador de Erro Absoluto Médio
MAPE	Indicador de Erro Percentual Absoluto Médio
MSE	Indicador de Erro Quadrático Médio
ON	Ação do tipo ordinária
PLN	Processamento de linguagem natural
PN	Ação do tipo preferencial
ReLU	Função Rectified Linear Unit ou Unidade Linear Retificada
RMSE	Indicador de Raiz Quadrada do Erro Médio

RNA	Rede Neural Artificial
RNC	Redes Neurais Convolucionais
RNR	Rede Neural Recorrente
SVM	Support Vector Machine
SVR	Support Vector Regression
USP	Universidade de São Paulo
USPSC	Campus USP de São Carlos

SUMÁRIO

1	INTRODUÇÃO	25
1.1	Objetivos do Trabalho	26
1.2	Estrutura do Trabalho	26
2	FUNDAMENTAÇÃO TEÓRICA	27
2.1	O mercado de ações no Brasil	27
2.1.1	Ações de empresas	27
2.1.2	A B3 - Brasil, Bolsa e Balcão S.A.	28
2.2	Aprendizado de máquina para previsão de preços de ações	30
2.2.1	Regressão de Vetores de Suporte (Support Vector Regression - SVR)	30
2.2.2	Redes Neurais Artificiais - RNAs (Artificial Neural Networks – ANNs)	31
2.2.2.1	Neurônio artificial	31
2.2.2.2	Tipos de aprendizado	32
2.2.2.3	Métodos de aprendizado profundo	33
2.2.2.4	Redes Neurais Recorrentes - RNRs (Recurrent Neural Networks - RNNs)	33
2.2.2.5	Redes Neurais Convolucionais - RNCs (Convolutional Neural Networks – CNNs)	36
2.3	Métricas de Avaliação	38
2.3.1	Raiz Quadrada do Erro Médio - RMSE	38
2.3.2	Erro Quadrático Médio - MSE	39
2.3.3	Erro Absoluto Médio - MAE	39
2.3.4	Erro Percentual Absoluto Médio - MAPE	40
2.4	Trabalhos relacionados	40
3	DADOS E ESTRATÉGIA METODOLÓGICA	43
3.1	Coleta e armazenagem dos dados	44
3.2	Tratamento dos dados	45
3.3	Aplicação do modelo SVR de previsão do preço da ação	45
3.4	Aplicação do modelo LSTM de previsão do preço da ação	46
3.5	Aplicação do modelo CNN de previsão do preço da ação	46
3.6	Aplicação do modelo CNN-LSTM de previsão do preço da ação	47
3.7	Avaliação dos resultados	48
4	RESULTADOS: AVALIAÇÃO EXPERIMENTAL E DISCUSSÃO	49
4.1	Modelo Support Vector Regression (SVR)	49
4.2	Modelo Long Short-Term Memory (LSTM)	49

4.3	Modelo Convolutional Neural Network (CNN)	50
4.4	Modelo híbrido CNN-LSTM	52
4.5	Resultados dos experimentos e discussão	53
5	CONCLUSÕES	55
	REFERÊNCIAS	57

1 INTRODUÇÃO

O setor de geração e distribuição de energia elétrica é um dos mais importantes e estratégicos setores de infraestrutura da economia de um país. Ele é responsável por fornecer a energia necessária para o funcionamento de todos os setores produtivos, incluindo a indústria, o comércio, os serviços e os transportes. Adicionalmente, nas últimas décadas se tornou relevante a busca por fontes de energia limpa, que não emitem gases de efeito estufa, por conta das mudanças climáticas que o planeta está passando, fato esse que significa um grande desafio não só para as empresas do setor, como também para toda a sociedade.

No Brasil, o setor de energia elétrica é composto por empresas privadas, públicas e de economia mista. As principais fontes de energia utilizadas no Brasil são a hidrelétrica, a termelétrica, a eólica e a solar. A hidrelétrica é a principal fonte de energia do país, responsável por cerca de 67% da geração de energia elétrica. As fontes de energia limpas para o meio ambiente (hidrelétrica, solar e eólica), responderam por 89% de toda a energia gerada no Brasil em 2023 (ONS - Operador Nacional do Sistema, 2023), fazendo com o Brasil apresente a matriz elétrica mais limpa entre as economias do grupo do G-20 (BloombergNEF, 2023).

As empresas do setor elétrico são, portanto, de grande interesse para investidores, que buscam oportunidades de lucro no mercado de ações. O potencial do Brasil se destaca, pois se encontra entre os 20 mercados mais atraentes para o investimento em energia limpa em todo o mundo (BloombergNEF, 2023). No entanto, a previsão do valor de ações é uma tarefa complexa, que envolve a análise de diversos fatores, como as condições econômicas do país, as políticas públicas, o desempenho das empresas e as expectativas dos investidores.

Nos últimos anos, o mercado financeiro tem testemunhado uma crescente integração entre a Inteligência Artificial (IA) e Big Data, transformando radicalmente a forma como são tomadas decisões no âmbito dos investimentos (Sabbar; Kharrim, 2023). Dentre diversas aplicações, a IA pode ser utilizada para identificar padrões nos dados históricos e, com base nessas informações, gerar previsões para o futuro. Este fenômeno é particularmente evidente no setor elétrico brasileiro, onde a volatilidade das ações apresenta desafios significativos para investidores e gestores de portfólio, dada a importância ambiental da necessidade de fontes limpas de energia.

Este trabalho teve como objetivo desenvolver experimentos com modelos de IA dos tipos Support Vector Regression (SVR), Long Short-Term Memory (LSTM), Convolutional Neural Network (CNN) e híbrido CNN-LSTM para prever o valor diário de abertura ação de empresa do setor elétrico brasileiro, comparando o desempenho desses modelos

nessa tarefa. Os modelos processaram um mesmo dataset que foi elaborado com o histórico de cotações dessa ação nos pregões da Bolsa de Valores B3 e determinados índices socioeconômicos do Brasil, tais como taxa de câmbio em relação ao dólar e consumo da energia elétrica..

1.1 Objetivos do Trabalho

O objetivo geral deste trabalho foi desenvolver modelos de IA para a predição da cotação de abertura da ação ELET3 (ELETROBRAS ON N1) nos pregões da B3 e comparar a eficiência dos mesmos nessa tarefa.

Os objetivos específicos deste trabalho são:

- Coletar e preparar um dataset com os dados históricos de cotações da ação ELET3 (ELETROBRAS ON N1) e índices socioeconômicos do Brasil;
- Desenvolver modelos de IA (SVR, LSTM, CNN e CNN-LSTM) para previsão da cotação de abertura da ação nos pregões da B3;
- Comparar o desempenho dos modelos desenvolvidos na tarefa de previsão do preço da ação.

1.2 Estrutura do Trabalho

A organização do restante deste documento é a seguinte: O Capítulo 2 apresenta a fundamentação teórica na qual se baseia esse trabalho, contendo a exposição dos principais conceitos relacionados ao mercado de ações brasileiro e ao aprendizado de máquina, incluindo suas principais características e técnicas específicas que são aplicadas no problema em questão. Também nesse capítulo, são apresentados trabalhos correlatos. O Capítulo 3 apresenta a metodologia utilizada para o desenvolvimento deste trabalho, bem como as soluções de aprendizado de máquina que serão adotadas para a abordagem do problema proposto. No Capítulo 4 são expostos os resultados alcançados no trabalho. Por fim, no Capítulo 5 são apresentadas a conclusão e as considerações finais do trabalho.

2 FUNDAMENTAÇÃO TEÓRICA

2.1 O mercado de ações no Brasil

2.1.1 Ações de empresas

Ação é a menor parte do capital social de uma empresa. É um título patrimonial que concede ao seu titular, denominado como “acionista”, todos os direitos, bem como todos os deveres de um sócio da empresa, limitado pela quantidade de ações que possua. No Brasil, apenas ações emitidas por empresas registradas na CVM (Comissão de Valores Mobiliários), caracterizadas como empresas de capital aberto, podem ser negociadas no mercado de valores mobiliários nacional de forma pública (Fortuna, 2015).

É por meio da negociação de ações, que envolve a venda direta de participações no capital social, que as empresas de capital aberto (sociedades anônimas – S.A.) arrecadam os recursos necessários para expandir seus negócios e aumentar seu patrimônio, se comprometendo a remunerar os seus investidores (acionistas) conforme o capital aplicado e os resultados futuros da empresa.

Os possuidores de ações de empresas, ou acionistas, fazem jus a receber dividendos, juros sobre o capital próprio e bonificações. Tais pagamentos estão condicionados ao desempenho financeiro de cada empresa envolvida, pois só acontecem quando a empresa apresenta lucro. Além disso, o valor final da parcela do lucro da empresa que a ser distribuída para os acionistas depende de alguns fatores, dentre eles o planejamento de investimentos da empresa, o caixa financeiro necessário para a operação da empresa e o estatuto social da mesma. O estatuto social de uma empresa é um documento que determina as principais características da empresa, incluindo regras a serem obedecidas pelos gestores e acionistas (Fortuna, 2015).

Os acionistas também podem lucrar com a valorização da ação da empresa. Porém, não há garantias nesse sentido. O preço da ação pode aumentar, mas também pode diminuir, já que varia conforme uma série de fatores internos e externos a empresa. Além do desempenho para própria empresa, o preço da ação é influenciado pela perspectiva do setor que atua e da economia nacional e internacional. Inclusive, uma ação pode perder todo o seu valor no caso da liquidação (falência) da empresa.

Existem dois tipos principais de ações:

- Ação Ordinária (ON): Concede ao titular o direito de voto quando da realização de assembleia de acionistas da empresa;
- Ação Preferencial (PN): Em geral sem direito a voto nas assembleias de acionistas da empresa, mas com prioridade no pagamento de dividendos ou no restituição de

capital.

A princípio, as ações de empresas poderiam ser negociadas diretamente entre as pessoas interessadas, porém, essa opção seria ineficiente no que se refere à divulgação das propostas, a determinação dos preços, lisura e comunicação das informações sobre as negociações, garantia dos pagamentos envolvidos, entre demais fatores. Para eliminar esses problemas, a negociação de ações de empresas é feita em bolsas de valores e mercados de balcão (Fortuna, 2015).

O processo de compra e venda de ações pode ser dividido em três fases:

- Pré-negociação (Pré-trading): engloba a avaliação de risco, definição dos ativos e contratos a serem colocados para negociação e a divulgação de informações sobre os mesmos, incluindo os indicadores de cotação;
- Negociação (Trading): é a fase em que os investidores, representados por intermediários financeiros (Corretoras e Distribuidoras), registram ordens para compra ou venda durante o pregão - sistema eletrônico de negociação;
- Pós-negociação (Pós-trading): trata-se das fase de processamento, liquidação e guarda dos ativos.

As empresas de capital aberto, que tem ações negociadas em bolsa, são obrigadas a atender uma série de exigências no que se refere ao fornecimento de informações para ao público. Tais empresas são obrigadas a cumprir um conjunto de determinações na divulgação de informações a seus acionistas - regras de disclosure (Comissão de Valores Mobiliários, 2019).

2.1.2 A B3 - Brasil, Bolsa e Balcão S.A.

Em 1851 foram instaladas no Brasil as primeiras bolsas de valores nas cidades do Rio de Janeiro (RJ) e Salvador (BA). Em 1895 foi criada a Bolsa de Títulos de São Paulo, entidade que passou a reunir os corretores tanto de títulos como de câmbio. Em 1984 foi criada a CETIP (Central de Custódia e Liquidação Financeira de Títulos Privados) com a missão de gerenciar o mercado de títulos de renda fixa, em especial aqueles de emissão bancária. Em 1986 surgiu a BM&F (Bolsa de Mercadorias e Futuros) com a finalidade de ser o principal mercado para contratos de mercadorias e derivativos (Fortuna, 2015).

No começo dos anos 2000, a Bolsa de Valores de São Paulo (BOVESPA) estabeleceu-se como a mais importante bolsa do Brasil. Em 2008, BM&F e Bovespa se uniram, formando a BM&FBOVESPA e em 2017, a BM&FBOVESPA se fundiu com a CETIP, resultando na criação da B3 (Brasil, Bolsa e Balcão). Empresa de capital aberto, se torna uma das maiores bolsas do mundo. A B3 é classificada como uma bolsa vertical, já que os serviços

que oferece cobrem todas as fases de um processo de negociação na bolsa (Pré-negociação, Negociação e Pós-negociação) (Comissão de Valores Mobiliários, 2019).

Atualmente, a negociação das ações na B3 é totalmente informatizada, acarretando agilidade, maior transparência, preços mais competitivos e custos menores. Isso beneficia especialmente os pequenos investidores, pois no sistema de dados as ordens de compra e de venda são tratadas igualmente, não havendo diferença com relação ao porte da transação (Fortuna, 2015).

A B3 divulga índices de desempenho que expressam a evolução do valor de grupos de ativos ou derivativos no decorrer do tempo. Entre eles, o IBOVESPA é o mais significativo, já que expressa a média das cotações das principais ações em termos de negociação e relevância no mercado de ações do Brasil. A cada quatro meses é feita uma reavaliação da carteira de ações que é referência para a formulação do IBOVESPA, de forma que a sua representatividade seja garantida ao longo do tempo (Comissão de Valores Mobiliários, 2019).

Um índice da BOVESPA de especial interesse nesse trabalho é o IEE (índice Setorial de Energia Elétrica). Criado em 1994, é composto pelas ações das maiores empresas de geração e distribuição de energia elétrica negociadas na bolsa, com o intuito de monitorar o desempenho deste segmento (B3, 2024).

Na tabela 1 são apresentadas as ações que compõem o índice IEE.

Tabela 1 – Ações do índice IEE

Código	Ação	Tipo	Qtde. Teórica	Part. (%)
AESB3	AES BRASIL	ON NM	17.300	5,493
ALUP11	ALUPAR	UNT N2	7.000	6,456
AURE3	AUREN	ON ED NM	16.700	6,148
CMIG4	CEMIG	PN N1	19.200	6,963
COCE5	COELCE	PNA	5.100	6,104
CPLE6	COPEL	PNB N2	21.000	6,515
CPFE3	CPFL ENERGIA	ON NM	5.600	5,801
ELET3	ELETROBRAS	ON N1	5.100	6,724
ENGI11	ENERGISA	UNT N2	4.100	6,335
ENEV3	ENEVA	ON NM	15.900	6,027
EGIE3	ENGIE BRASIL	ON NM	4.800	5,922
EQTL3	EQUATORIAL	ON NM	6.100	6,21
NEOE3	NEOENERGIA	ON NM	10.800	7,153
SRNA3	SERENA	ON NM	21.600	5,53
TAEE11	TAESA	UNT N2	5.800	6,218
TRPL4	TRAN PAULIST	PN N1	8.400	6,401
Quantidade Teórica Total			174.500	100

2.2 Aprendizado de máquina para previsão de preços de ações

A previsão do preço das ações é uma área na qual várias abordagens estão sendo feitas há anos, já que podem gerar lucros significativos. Porém, prever o mercado de ações é uma tarefa muito complexa, especialmente pelo comportamento instável das séries temporais envolvidas. As análises fundamentalista e técnica, do âmbito da teoria econômica, foram os primeiros métodos utilizados para prever os preços das ações. Atualmente, as Redes Neurais Artificiais (RNAs) estão sendo mais utilizadas nesse desafio. Entretanto, um ponto de atenção é que, na maioria dos casos, as RNAs sofrem do problema de over-fitting (sobreajuste), devido ao grande número de parâmetros envolvidos e erro na seleção dos mesmos pelo usuário. Nesse caso, o modelo de RNA se ajusta demais aos dados que são inseridos para seu o treinamento do modelo, porém, quando processa novos dados, não consegue capturar os seus padrões com eficiência, comprometendo a qualidade das previsões (Hegazy; Soliman; Salam, 2014).

Algumas das principais soluções de aprendizado de máquina para a previsão de preço de ações são:

2.2.1 Regressão de Vetores de Suporte (Support Vector Regression - SVR)

O SVR é uma variação do algoritmo Support Vector Machine (SVM), uma vez que esse último é normalmente utilizado para a tarefa de classificação. O SVR é aplicado à tarefa de regressão. O SVR pressupõe a determinação de um hiperplano no qual os objetos estejam mais próximos possível, de forma que a distância para a superfície seja a menor possível. Os desvios não devem ultrapassar uma margem especificada. São utilizados vetores de suporte para definir a margem de erro em torno da curva de regressão (Minasny, 2009).

A função de regressão do SVR é definida como:

$$y = f(x) = \sum_{i=1}^n \alpha_i k(x, x_i) + b \quad (2.1)$$

Onde,

- x é o vetor de entrada;
- x_i são os vetores de suporte;
- α_i são os coeficientes de Lagrange;
- $k(x, x_i)$ é uma função de kernel que mede a similaridade entre (x) e (x_i) ;
- b é o termo de viés.

A função de kernel permite que o modelo trabalhe com dados de entrada não linearmente separáveis, mapeando os mesmos para um espaço de maior dimensão. Isso permite que o SVR modele relações mais complexas das variáveis de entrada e saída. A escolha da função kernel depende do tipo de dados de entrada e das características do problema analisado. Alguns tipos de função Kernel são a linear, polinomial, Gaussiana RBF e Exponencial RBF (Gani; Taleb; Limam, 2010).

O SVR é uma das aplicações mais estudadas para a previsão de séries financeiras. Nesse tipo de série, as análises são particularmente feitas visando a previsão do índice do mercado financeiro ou do preço de um ativo como uma ação na bolsa de valores (Sapankevych; Sankar, 2009).

2.2.2 Redes Neurais Artificiais - RNAs (Artificial Neural Networks – ANNs)

São modelos computacionais inspirados no sistema nervoso biológico, que tem como elemento principal uma estrutura denominada neurônio artificial que busca simular o funcionamento de um neurônio humano. As RNAs possuem uma quantidade variável de neurônios artificiais interligados por interconexões representadas por vetores/matrizes de pesos sinápticos (Silva; Spatti; Flauzino, 2010).

2.2.2.1 Neurônio artificial

São modelos bem simplificados dos modelos biológicos. A sua estrutura mais simples foi proposta por McCulloch & Pitts em 1943, conforme ilustrado na figura 1, sendo ainda o modelo mais utilizado nas várias arquiteturas de redes neurais artificiais (Silva; Spatti; Flauzino, 2010).

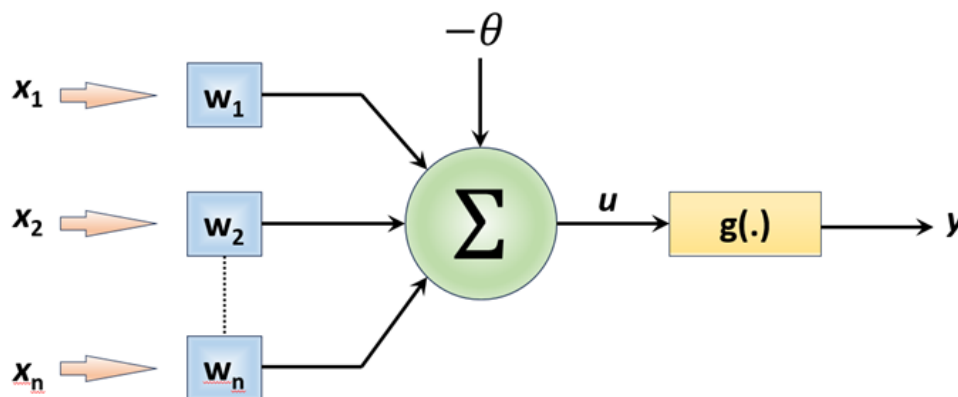


Figura 1 – Neurônio artificial. Adaptado de (Silva; Spatti; Flauzino, 2010)

Os elementos básicos do neurônio artificial são:

- Sinais de entrada (x_1, x_2, \dots, x_n)

São os sinais vindos do meio externo que são geralmente normalizados para melhorar a eficácia dos algoritmos na aprendizagem.

- Pesos sinápticos ($w1, w2_{ewn}$)

São valores que irão multiplicar os sinais de entrada, controlando a influência de cada entrada no resultado final do neurônio.

- Combinador linear (Σ)

Combina todos os sinais de entrada ajustados pelos pesos sinápticos, resultando em um valor de potencial de ativação.

- Limiar de ativação (θ)

É uma variável que determina qual será o nível mínimo para que a saída do combinador linear possa gerar um valor para a saída do neurônio. Ou seja, o limiar de ativação é o ponto de decisão que determina se o neurônio será ativado ou não com base na entrada recebida.

- Potencial de ativação (μ)

Corresponde à diferença resultante entre o valor gerado pelo combinador linear e o limiar de ativação. É usado como entrada para uma função de ativação, que determinará a saída final do neurônio.

- Função de ativação (g)

Determina a saída do neurônio dentro de um limiar de valores admissíveis conforme a sua própria imagem funcional. Ela introduz não-linearidades na resposta do neurônio. Existem várias funções de ativação diferentes, sendo algumas das mais utilizadas as funções degrau, sigmoide e ReLU.

- Sinal de saída (y)

Refere-se ao valor final emitido pelo neurônio considerando os sinais de entrada, podendo ser utilizado como sinal de entrada para outro neurônio artificial.

2.2.2.2 Tipos de aprendizado

Existem 3 tipos principais de aprendizado de máquina:

- Aprendizado supervisionado:

O treinamento do algoritmo é realizado com um conjunto de dados de entrada que está associado a um rótulo, ou seja, aprende a associar cada entrada a uma saída específica. Dessa forma, quando novas entradas forem tratadas pelo algoritmo, ele pode prever as novas saídas com base no treinamento anterior (Heaton, 2018).

- Aprendizado não supervisionado:

Ao contrário do aprendizado supervisionado, o conjunto de dados não está associado a rótulos, onde as saídas desejadas não são fornecidas. Assim, o algoritmo aprende as propriedades úteis da estrutura dos dados, envolvendo toda a distribuição de probabilidade ou algumas propriedades interessantes dessa distribuição inerente ao conjunto de dados (Heaton, 2018).

- Aprendizado por reforço:

Em algumas aplicações, a saída do sistema é uma sequência de ações. Cada sequência de ações está associada a recompensas ou penalidades ao final. O algoritmo deve aprender com as boas sequências de ações anteriores, de forma a gerar as melhores recompensas ao longo do tempo. Esse tipo de aprendizado é geralmente utilizado no campo dos jogos, robótica, controle de processos e navegação autônoma (Heaton, 2018).

2.2.2.3 Métodos de aprendizado profundo

As técnicas de aprendizado profundo tem sido cada vez mais pesquisadas devido a sua capacidade de extrair características complexas dos dados de entradas, sem a necessidade de pré-processamento desses dados, como no caso dos modelos de Aprendizado de Máquina tradicionais (Liu *et al.*, 2017). Apesar de suas vantagens, cabe registrar que essas novas soluções de aprendizado profundo apresentam a desvantagem do elevado tempo necessário para o treinamento dos modelos devido a complexidade de ajuste de seus hiperparâmetros (Derbentsev *et al.*, 2020).

2.2.2.4 Redes Neurais Recorrentes - RNRs (Recurrent Neural Networks - RNNs)

São uma família de redes neurais para processamento de dados sequenciais. Conseguem tratar sequências muito mais longas do que outras redes sem especialização. São amplamente utilizadas em tarefas de processamento de linguagem natural (PLN), previsão de séries temporais, reconhecimento de fala, entre outras. Utilizam a realimentação de parâmetros em diferentes partes da topologia do modelo. Cada elemento da saída é uma função dos elementos anteriores da saída. Cada membro de saída decorre da utilização da mesma regra de atualização aplicada às anteriores, o que resulta em um compartilhamento profundo de parâmetros (Heaton, 2018).

Conforme ilustrado pelo diagrama na figura 2, um pedaço da rede neural, A, recebe um sinal da entrada x_t e gera um valor h_t . Um loop faz com que os dados sejam avançados de um ponto da rede para o próximo.

Uma rede neural recorrente pode ser entendida como múltiplas cópias dessa estrutura, cada qual avançado um sinal para um sucessor, conforme ilustrado na figura 3. Esta

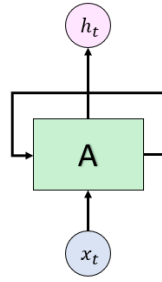


Figura 2 – Elemento da Rede Neural. Adaptado de (Olah *et al.*, 2015)

estrutura em cadeia indica que as redes neurais recorrentes se relacionam fortemente com sequências e listas. Recentemente, as RNRs tem sido aplicadas com sucesso para tratar diversos problemas como reconhecimento de fala, modelagem de linguagem, tradução, legendagem de imagens e outros (Olah *et al.*, 2015).

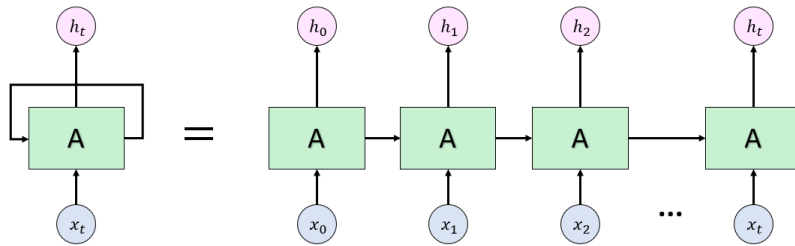


Figura 3 – Diagrama de uma Rede Neural Recorrente (RNR). Adaptado de (Olah *et al.*, 2015)

- Long Short-Term Memory (LSTM):

A rede LSTM é uma variante de RNR capaz de aprender dependências de longo prazo. Ela foi projetada em 1997 por Hochreiter e Schmidhuber para superar algumas das limitações das RNRs padrão, como a questão do gradiente que desaparece ou explode, que pode dificultar o aprendizado de dependências de longo prazo em sequências de dados. Uma LSTM é adequada para classificar e prever dados de séries temporais. O LSTM foi aplicado com muito sucesso no reconhecimento irrestrito de caligrafia, reconhecimento de fala, geração de caligrafia, tradução automática, legendagem de imagens e outras aplicações (Heaton, 2018). Uma representação gráfica da LSTM é ilustrada na figura 4.

A primeira etapa executada dentro da célula LSTM é a remoção das memórias que deveriam ser omitidas da célula. Esta ação é feita por uma camada sigmoide conhecida como "camada de esquecimento". Ela analisa $h(t-1)$ e x_t e gera um valor entre 0 e 1 para cada valor no estado da célula $c(t-1)$. Essa etapa é representada

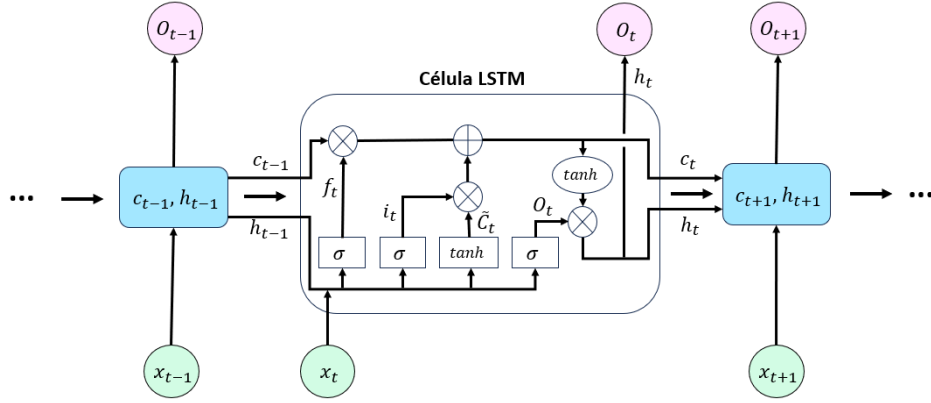


Figura 4 – Diagrama de uma rede LSTM. Adaptado de (Olah *et al.*, 2015)

pela equação 2.10 (Olah *et al.*, 2015).

$$f_t = \sigma(W_f[h_{t-1}, x_t] + b_f) \quad (2.2)$$

Onde,

- f_t representa o gate de esquecimento;
- σ é a função de ativação sigmóide;
- x_t é input de entrada no tempo t ;
- W_f e b_f representam o peso e parâmetro de vetores de viés que precisam ser aprendidos durante o treinamento, respectivamente.

A próxima etapa é definir quais novos dados serão carregados no estado da célula LSTM. Essa etapa é dividida em duas partes. Na primeira parte, uma camada sigmóide chamada “camada de porta de entrada” define quais valores serão utilizados. Na sequência, uma camada tanh gera um vetor de possíveis novos valores, C_t , que podem ser adicionados ao estado (Olah *et al.*, 2015). Essa etapa é representada pelas equações 2.11 e 2.12.

$$i_t = \sigma(W_i[h_{t-1}, x_t] + b_i) \quad (2.3)$$

Onde i_t representa o gate de entrada e atualização.

$$\tilde{C}_t = \tanh(W_c[h_{t-1}, x_t] + b_c) \quad (2.4)$$

Sendo que \tilde{C}_t é a célula de entrada.

O próximo passo é atualizar o estado anterior da célula, c_{t-1} , adicionando as novas informações definidas nas etapas anteriores. Atualizando o instante no tempo t , conforme descrito na equação 2.13.

$$C_t = f_t * C_{t-1} + i_t * \tilde{C}_t \quad (2.5)$$

Onde C_t representa o estado da célula.

O gate de saída será definido pela célula de estado filtrada. Primeiro, executamos uma camada sigmóide que define as partes do estado da célula que serão geradas. Isso é representado pela equação 2.14.

$$O_t = \sigma(W_0[h_{t-1}, x_t] + b_0) \quad (2.6)$$

Onde O_t representa o gate de saída.

Finalmente, o estado da célula em \tanh é multiplicado pela saída da porta sigmóide, de modo a se obter a saída desejada no instante de tempo t (Olah *et al.*, 2015), conforme a equação 2.15.

$$h_t = O_t * \tanh(C_t) \quad (2.7)$$

Sendo h_t a saída da unidade LSTM.

2.2.2.5 Redes Neurais Convolucionais - RNCs (Convolutional Neural Networks – CNNs)

Do inglês Convolutional Neural Networks (CNNs), foi idealizada pelo pesquisador francês Yann LeCun (LeCun *et al.*, 1998). Inspirada na visão humana, foi desenvolvida para lidar com dados representados por múltiplos vetores, como imagens coloridas que consistem em três vetores 2D correspondentes aos pixels dos canais de cores RGB (vermelho, verde e azul) (LeCun; Bengio; Hinton, 2015).

Preserva as relações entre pixels próximos na imagem, atuando nos campos receptivos locais executando um processo denominado convolução. Assim, a sensibilidade a translações, rotações e distorções das imagens é diminuída (Lecun *et al.*, 1993). Outras redes neurais não são capazes de captar essas correlações, uma vez que processam as imagens de forma unidimensional.

Nos últimos anos, o uso das redes neurais convolucionais não está restrito a tarefa de classificação de imagens. Vem sendo utilizada com sucesso nas tarefas de reconhecimento de voz (Dua *et al.*, 2022), processamento de linguagem natural (Geng *et al.*, 2022) e outras aplicações.

- Convolução.

Trata-se de uma operação matemática que combina duas funções para produzir uma terceira função. Na rede RNC, a convolução é executada entre a matriz de uma imagem de entrada e um filtro (também chamado de kernel ou máscara), que é uma matriz menor a da imagem. Os elementos do filtro são multiplicados e somados com os elementos da imagem, produzindo uma nova matriz denominada mapa de características, que representa as regiões da imagem de entrada que ativam determinados padrões aprendidos pelo filtro. Quanto mais convoluções forem realizadas com filtros

diferentes, mais informações sobre as características da imagem original são extraídas (Heaton, 2018).

A convolução em 2 dimensões é expressa através da fórmula 2.16.

$$I * k[i, j] = \sum_m \sum_n I(i - m, j - n)k(m, n) \quad (2.8)$$

Onde,

- $I * k[i, j]$ representa o valor resultante da convolução na posição (i, j) do mapa de características;
- $I(i, j)$ é o pixel da posição (i, j) da imagem de entrada;
- $k(m, n)$ é o coeficiente do filtro na posição (m, n) ;
- m e n são os índices de iteração sobre as dimensões do filtro.

- Arquitetura da RNC

As RNCs se baseiam explicitamente na estrutura da imagem ao processar os dados. Os neurônios nas RNCs são estruturados em 3 dimensões – largura, altura e profundidade. Cada neurônio de uma determinada camada está conectado a um pequeno trecho da saída da camada anterior, como se um filtro de tamanho $N \times N$ fosse usado para processar a imagem de entrada. Esse filtro atua como um extrator das características da imagem. Porém, apenas um filtro não consegue captar todas as características da imagem, por isso esse processo de filtragem é aplicado M vezes para que se garanta que todos os detalhes da imagem sejam capturados. Os filtros das camadas iniciais da RNC capturam as características principais da imagem, como arestas, cantos e assim por diante. As últimas camada da rede extraem características de nível superior da imagem (Artasanchez; Joshi, 2020).

Em sua forma mais básica, conforme ilustrado na figura 5, a RNC é composta pelas seguintes camadas:

- Camada de Entrada (em inglês, Input Layer): recebe os dados brutos da imagem original;
- Camada Convolutiva (em inglês, Convolutional Layer): esta camada calcula as convoluções entre os neurônios e os vários patches na entrada. Basicamente, essa camada calcula o produto escalar entre os pesos e uma região específica da saída da camada anterior;
- Camada de Unidade Linear Retificada (em inglês, Rectified Linear Unit Layer): executa uma função de ativação (ReLU) à saída da camada anterior. Tal função geralmente é do tipo $\max(0, x)$. É responsável por introduzir não-linearidade na rede, possibilitando a captura de relações mais complexas nos dados de entrada;

- Camada de Pooling (em inglês, Pooling layer): reduz a dimensionalidade da saída da camada anterior para uma estrutura menor. Ajuda a preservar as características importantes da imagem ao longo da rede. O pooling máximo é geralmente usado nessa camada, optando-se pelo valor máximo em uma determinada janela $K \times K$;
- Camada Totalmente Conectada (em inglês, Fully Connected Layer): calcula as pontuações de saída na última camada. Cada neurônio dessa camada está interligado a todos os neurônios na camada anterior. A saída final possui dimensões $1 \times 1 \times L$, sendo L o número de classes que estão rotuladas no conjunto de imagens utilizado para treinamento da rede.

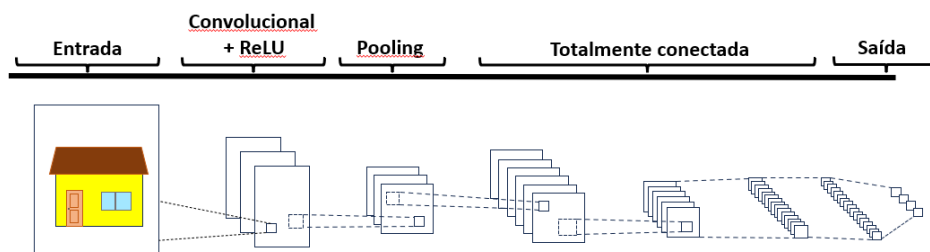


Figura 5 – Diagrama de uma Rede Neural Convolucional. Adaptado de (Artasanchez; Joshi, 2020)

Desde a camada de entrada até a camada de saída na rede, a imagem original é transformada dos valores de pixel nas pontuações finais da classe. Diversas arquiteturas foram propostas para RNCs foram propostas e a pesquisa com esse tipo de rede continua. A acurácia e a estabilidade de um modelo são decorrentes de vários fatores, tais como o tipo, a disposição e a quantidade de camadas da rede, as funções designadas para cada camada, os conjuntos de dados de treinamento e outros fatores relacionados (Artasanchez; Joshi, 2020).

2.3 Métricas de Avaliação

Para avaliar a efetividade das soluções de aprendizado de máquina para a previsão de preço de ações pode ser feita ao se comparar os valores reais com os valores que foram previstos, medindo-se o erro ocorrido. Quanto menor esse erro, melhor é a eficiência da previsão. Nesse trabalho, serão utilizadas as seguintes métricas (Carmo; Silva, 2023):

2.3.1 Raiz Quadrada do Erro Médio - RMSE

Do inglês "Root Mean Square Error", o RMSE é a raiz quadrada da média dos quadrados dos erros. Expressa a dispersão dos erros entre os valores preditos pelo modelo e os valores reais. Ele amplifica os erros grandes, mas está na mesma unidade que os dados

originais, facilitando a interpretação. É menos sensível a outliers. Quanto menor o valor do RMSE, melhor é o desempenho do modelo

É expresso através da fórmula 2.17.

$$RMSE = \sqrt{\frac{\sum_{i=1}^n (y_i - y_p)^2}{n}} \quad (2.9)$$

Onde,

- y_i é o valor real;
- y_p é o valor predito pelo modelo;
- n é o número total de observações no conjunto de dados de teste.

2.3.2 Erro Quadrático Médio - MSE

Do inglês "Mean Square Error", o MSE é a soma das diferenças entre o valor predito e o valor real, elevada ao quadrado. Destaca as medições nas quais a diferença entre esses valores é muito grande, o que pode gerar distorção na avaliação desse indicador. Quanto menor o valor do MSE, melhor é o desempenho do modelo.

É expresso através da fórmula 2.18.

$$MSE = \frac{\sum_{i=1}^n (y_i - y_p)^2}{n} \quad (2.10)$$

Onde,

- y_i é o valor real;
- y_p é o valor predito pelo modelo;
- n é o número total de observações no conjunto de dados de teste.

2.3.3 Erro Absoluto Médio - MAE

Do inglês "Mean Absolute Error", o MAE é o módulo do erro médio. Expressa a média das diferenças absolutas entre os valores preditos pelo modelo e os valores reais. Sendo assim, sofre a influência da mesma forma dos erros grandes e pequenos, tornando-se mais robusto a valores atípicos. Quanto menor o valor do MAE, melhor é o desempenho do modelo.

É expresso através da fórmula 2.19.

$$MAE = \frac{1}{n} \sum_{i=1}^n |y_i - y_p| \quad (2.11)$$

Onde,

- y_i é o valor real;
- y_p é o valor predito pelo modelo;
- n é o número total de observações no conjunto de dados de teste.

2.3.4 Erro Percentual Absoluto Médio - MAPE

Do inglês "Mean Absolute Percentage Error", o MAPE representa a média das diferenças percentuais absolutas entre os valores preditos pelo modelo e os valores reais, fornecendo uma medida da precisão média das previsões em relação aos valores reais. É especialmente útil nos casos em que o valor absoluto do erro não é tão relevante quanto a proporção entre o erro e o valor real. Quanto menor o valor do MAPE, melhor é o desempenho do modelo.

É expresso através da fórmula y.12.

$$MAPE = \frac{100}{n} \sum_{i=1}^n \left| \frac{y_i - y_p}{y_i} \right| \quad (2.12)$$

Onde,

- y_i é o valor real;
- y_p é o valor predito pelo modelo;
- n é o número total de observações no conjunto de dados de teste.

2.4 Trabalhos relacionados

São apresentados a seguir alguns trabalhos que aplicaram os modelos de Aprendizado de Máquina que serão experimentados nesse trabalho para predição de preços do mercado de ações. Entretanto, cabe registrar que os trabalhos citados, embora utilizando os mesmos modelos de AM desse trabalho, utilizaram atributos, abordagens e técnicas adicionais diferentes, como decorrência do amplo espectro de soluções que a área de Aprendizado de Máquina, em constante evolução, possibilita para a execução de uma mesma tarefa.

Sim, Kim e Ahn compararam a aplicabilidade do modelo CNN para previsão do mercado de ações, comparando com modelos RNA e SVM. Utilizaram, também, 9 indicadores técnicos (preço de fechamento, média móvel simples (SMA), média móvel exponencial (EMA), taxa de variação (ROC), convergência/divergência de média móvel (MACD), Fast%K, Slow%D, banda superior e banda inferior) para verificar o uso do método CNN no mercado de ações. Segundo a conclusão desse trabalho, o modelo CNN apresentou a melhor eficiência na previsão (Sim *et al.*, 2019).

Eapen, Bein e Verma apresentaram dois modelos de RNC com um único e com vários pipelines e unidades LSTM bidirecionais, e compararam o desempenho desses com um modelo SVR. Constatam, através dos resultados do indicador MSE, que os dois modelos RNC tiveram um desempenho em torno de 4,7 vezes melhor que o modelo SVR (Eapen; Bein; Verma, 2019).

Nikou, Mansourfar e Bagherzadeh propuseram um modelo LSTM e comparou-o com o modelo ANN, modelo de Regressão de Vetores de Suporte (SVR) e modelo RF. Os resultados indicaram que o modelo LSTM teve melhor desempenho na previsão dos preços de fechamento do iShares MSCI United Kingdom do que os demais modelos constantes do trabalho (Nikou; Mansourfar; Bagherzadeh, 2019).

Lakshminarayanan e McCrae fizeram simulações com quatro modelos usando LSTM e quatro modelos SVM. Utilizaram quatro conjunto de dados de preços de ações do Dow Jones Index (DJI), alterando entre os valores dos preços, preços de petróleo bruto e do ouro, e valores da média móvel desses parâmetros. Com os resultados obtidos, constataram que, no geral, o LSTM apresentou o melhor desempenho que o SVM (Lakshminarayanan; McCrae, 2019).

Rana, Uddin e Hoque propuseram modelos de Regressão Logística, LSTM e SVR, utilizando um conjunto de dados com o preço de fechamento da ação da empresa espanhola Acciona entre 2008 e 2018. Fizeram experimentos alterando parâmetros desses modelos. Os resultados apontaram que o modelo LSTM apresentou o melhor desempenho na previsão do preço da ação, particularmente utilizando a função de ativação linear com o otimizador amax e com a função de ativação tanh e o otimizador adam (Rana; Uddin; Hoque, 2019).

Bathla comparou um modelo LSTM com um modelo SVR na previsão de valores dos índices de ações, como S&P 500, NYSE, NSE, BSE, Dow Jones Industrial Average e NASDAQ. Nesse experimento, constatou que o LSTM supera o SVR e fornece melhor precisão de previsão (Bathla, 2020).

Sharaf, Hemdan, El-Sayed e El-Bahnasawy compararam modelos de aprendizado profundo e de máquina como SVR, LSTM, CNN e outros, para a previsão do preço de ações em diferentes cenários. Segundo os experimentos realizados, o modelo CNN superou os demais modelos de forma consistente nos diferentes casos (Sharaf *et al.*, 2021).

3 DADOS E ESTRATÉGIA METODOLÓGICA

Nessa seção são descritas as sete etapas da execução do experimento desse trabalho, que estão representadas no diagrama da figura 6. Das técnicas de aprendizado de máquina descritas no item 2.2, são aplicadas as de SVR, LSTM, CNN e a híbrida CNN-LSTM no experimento realizado nesse trabalho. Esses modelos, que se baseiam no aprendizado supervisionado, foram desenvolvidos na linguagem Python, utilizando as bibliotecas relacionadas na tabela 2.

Tabela 2 – Bibliotecas Python utilizadas nos modelos do experimento.

Biblioteca	Documentação
datetime	https://docs.python.org/3/library/datetime.html
importlib	https://docs.python.org/3/library/importlib.html
matplotlib.pyplot	https://matplotlib.org/stable/api/pyplot_
numpy	https://numpy.org/
pandas	https://pandas.pydata.org/docs/
pytz	https://pypi.org/project/pytz/
scikit-learn - SVM	https://scikit-learn.org/stable/api/sklearn.svm.html
scikit-learn - metrics	https://scikit-learn.org/stable/api/sklearn.metrics.html
scikit-learn - preprocessing	https://scikit-learn.org/stable/modules/preprocessing
tensorflow - keras	https://www.tensorflow.org/guide/keras?hl=pt-br
subprocess	https://docs.python.org/3/library/subprocess.html
sys	https://docs.python.org/3/library/sys.html

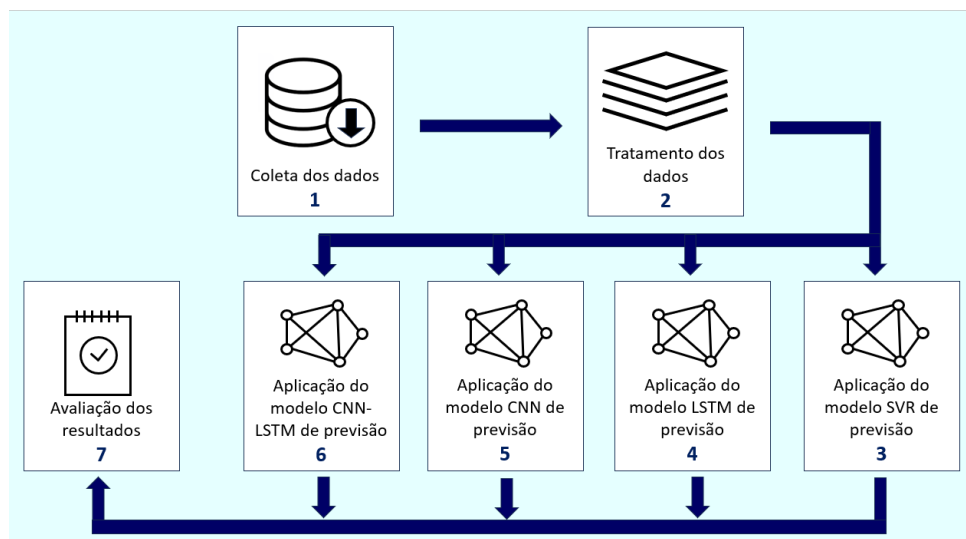


Figura 6 – Etapas de execução do trabalho (elaborado pelo autor)

As explicações detalhadas das etapas do experimento são apresentadas nas seções seguintes.

3.1 Coleta e armazenagem dos dados

São coletadas as cotações diárias do preço de abertura da ação ELET3 (ELETROBRAS ON N1) do site da B3, referentes a um período de 10 anos, de 02/01/2014 à 28/12/2023, conforme ilustrado na figura 7.

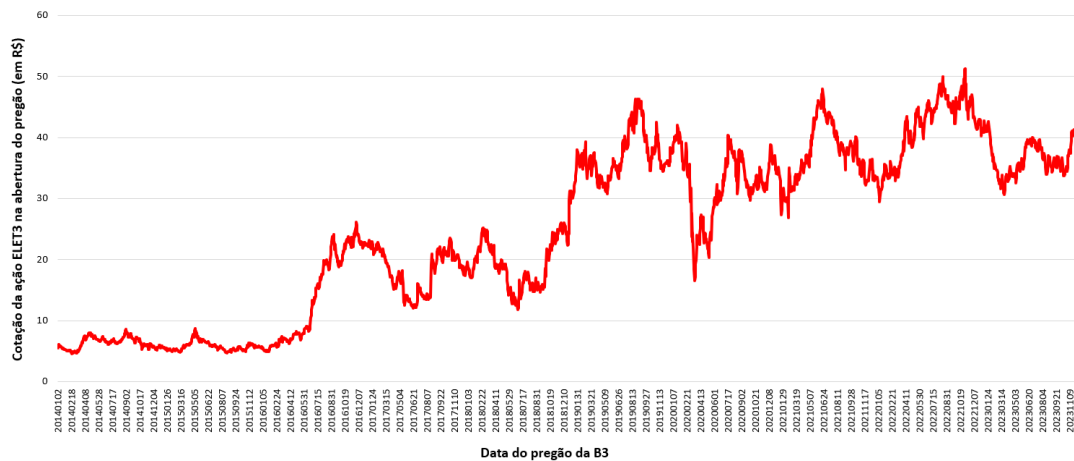


Figura 7 – Cotações de abertura (PREABE) da ação ELET3 (em R\$) no período de amostragem (elaborado pelo autor)

Também são coletados indicadores econômicos referentes tanto ao setor de energia elétrica no Brasil como ao cenário macro do país. Os dados coletados são armazenados em arquivos independentes no micro local do autor. Todos os dados coletados e suas fontes estão descritos nas tabelas 3 e 4.

Tabela 3 – Características dos dados coletados.

Código	Descrição	Período	Unidade	Fonte
DTPREG	Data do pregão	Diária	aaaammdd	B3
CODNEG	Código de negociação do papel	Diário	Alfanumérico	B3
PREABE	Preço de abertura do papel no pregão	Diário	Moeda - Real (R\$)	B3
PREMAX	Preço de máximo do papel no pregão	Diário	Moeda - Real (R\$)	B3
PREMIN	Preço de mínimo do papel no pregão	Diário	Moeda - Real (R\$)	B3
PREULT	Preço do último negócio do papel no pregão	Diário	Moeda - Real (R\$)	B3
IPCABR	Índice Nacional de Preços ao Consumidor Amplo - IPCA	Mensal	Percentual (%)	IBGE
CEELBR	Consumo de energia elétrica no Brasil	Mensal	Megawatt/hora (MWh)	EPE
NCEEBR	Número de consumidores de energia elétrica no Brasil	Mensal	Unidade	EPE
PIBBRA	Produto Interno Bruto do Brasil - PIB	Trimestral	Moeda - Real (R\$)	IBGE
DOLARA	Cotação do Dólar Americano	Diário	Moeda - Real (R\$)	BACEN

Tabela 4 – Fontes dos dados coletados.

Código	Fonte	Link para os dados	Acesso em
DTPREG	B3	https://www.b3.com.br/pt_br/market-data-e	10/04/2024
CODNEG	B3	https://www.b3.com.br/pt_br/market-data-e	10/04/2024
PREABE	B3	https://www.b3.com.br/pt_br/market-data-e	10/04/2024
PREMAX	B3	https://www.b3.com.br/pt_br/market-data-e	10/04/2024
PREMIN	B3	https://www.b3.com.br/pt_br/market-data-e	10/04/2024
PREULT	B3	https://www.b3.com.br/pt_br/market-data-e	10/04/2024
IPCABR	IBGE	https://ftp.ibge.gov.br/Precos_Indices_de	13/04/2024
CEELBR	EPE	https://www.epe.gov.br/pt/publicacoes-dados	27/03/2024
NCEEBR	EPE	https://www.epe.gov.br/pt/publicacoes-dados	27/03/2024
PIBBRA	IBGE	https://www.ibge.gov.br/estatisticas/economicas	13/04/2024
DOLARA	BACEN	https://www.bcb.gov.br/estabilidadefinanceira/	20/04/2024

3.2 Tratamento dos dados

Inicialmente, é feita uma verificação em busca de dados ausentes ou nulos. Nessa etapa, os dados dos onze atributos são agregados em um único arquivo. Para os quatro atributos de periodicidade mensal e trimestral, são repetidos os valores para os dias faltantes conforme o mês ou trimestre mais próximo da data sem valor. Em seguida, é feita uma verificação em busca de dados ausentes ou nulos no arquivo final.

Na sequência, o arquivo é dividido em dois arquivos. Isso é feito para evitar que o modelo tenha acesso aos dados de teste durante a fase de treinamento, visando diminuir a probabilidade de ocorrência de overfitting (sobreajuste) durante as fases de treinamento e teste dos modelos. Um dos arquivos com 80% dos dados é utilizado para treinar os modelos. O outro arquivo com 20% dos dados é utilizado para testar as previsões dos preços da ação feitas pelos modelos.

Durante a execução dos modelos, os dados de teste apenas são carregados após o término da etapa de treinamento. Isso é feito para que não exista a possibilidade dos modelos apreenderem as características dos dados de teste durante a fase de treinamento.

3.3 Aplicação do modelo SVR de previsão do preço da ação

Nessa etapa é aplicado um modelo de Support Vector Regression (SVR) nos arquivos de dados de treino e teste. São feitos experimentos alterando os valores dos parâmetros do modelo LSTM relacionados abaixo, sendo escolhida a configuração que apresente a melhor eficiência na previsão do preço da ação com base nos indicadores RMSE, MSE, MAE e MAPE descritos no item 2.3.

- Kernel: o tipo de kernel utilizado (linear, polinomial, rbf e sigmoid);
- C: controla o equilíbrio entre o erro durante o treinamento e a eficácia do modelo em generalizar;

- Epsilon: Define uma faixa nas previsões em que os erros não são penalizados. Valor numérico.

Cabe registrar que uma ação adicional que poderia ser aplicada na tentativa de melhorar o desempenho do modelo de Support Vector Regression (SVR) seria a seleção dos atributos do dataset com base na sua correlação e cointegração com o atributo alvo ("PREABE"). Porém, como o objetivo desse trabalho é comparar a eficiência entre os modelos, essa ação específica não foi executada, sendo processado o mesmo dataset em todos os modelos.

3.4 Aplicação do modelo LSTM de previsão do preço da ação

Nessa etapa é aplicado um modelo de Long Short-Term Memory (LSTM) nos arquivos de dados de treino e teste. O modelo executa a previsão do preço de abertura da ação (PREABE) do dia seguinte, utilizando uma janela deslizante com os valores dos 9 (nove) atributos relevantes do dataset nos 60 (sessenta) dias anteriores, tanto na fase de treinamento do modelo como na fase de teste.

São feitos experimentos alterando os valores dos parâmetros do modelo LSTM relacionados abaixo, sendo escolhida a configuração que apresente a melhor eficiência na previsão do preço da ação com base nos indicadores RMSE, MSE, MAE e MAPE descritos no item 2.3.

- Unidades LSTM: aumentar ou diminuir o número de células LSTM em cada camada;
- Número de Camadas LSTM: aumentar as camadas LSTM;
- Taxa de Dropout: variar esse parâmetro que tem como valores padrões 0.2 e 0.5;
- Função de Otimização: alterar os otimizadores entre otimizadores RMSprop, SGD e Adam;
- Número de Épocas: quantidade de vezes que o modelo processa o conjunto de dados de treinamento durante o treinamento;
- Batch Size: número de amostras de dados de treinamento processadas por vez.

3.5 Aplicação do modelo CNN de previsão do preço da ação

Nessa etapa é aplicado um modelo de Convolutional Neural Network (CNN) nos arquivos de dados de treino e teste. O modelo executa a previsão do preço de abertura da ação (PREABE) do dia seguinte, utilizando uma janela deslizante com os valores dos 9 (nove) atributos relevantes do dataset nos 60 (sessenta) dias anteriores, tanto na fase de treinamento do modelo como na fase de teste.

São feitos experimentos alterando os valores dos parâmetros do modelo CNN relacionados abaixo, sendo escolhida a configuração que apresente a melhor eficiência na previsão do preço da ação com base nos indicadores RMSE, MSE, MAE e MAPE descritos no item 2.3.

- Número de filtros: impacta na capacidade do modelo de extrair características dos dados;
- Tamanho do kernel: afeta a eficiência do modelo em extrair características dos dados;
- Tamanho da janela de pooling: impacta na capacidade do modelo de aprender pequenas alterações nos dados;
- Número de unidades nas camadas densas: afeta a eficiência do modelo em aprender configurações mais complexas dos dados;
- Função de ativação: experimentar as funções ReLU, tanh ou sigmoid.

3.6 Aplicação do modelo CNN-LSTM de previsão do preço da ação

Nessa etapa é aplicado um modelo híbrido CNN-LSTM nos arquivos de dados de treino e teste. O modelo executa a previsão do preço de abertura da ação (PREABE) do dia seguinte, utilizando uma janela deslizante com os valores dos 9 (nove) atributos relevantes do dataset nos 60 (sessenta) dias anteriores, tanto na fase de treinamento do modelo como na fase de teste.

São feitos experimentos alterando os valores dos parâmetros do modelo CNN-LSTM relacionados abaixo, sendo escolhida a configuração que apresente a melhor eficiência na previsão do preço da ação com base nos indicadores RMSE, MSE, MAE e MAPE descritos no item 2.3.

- Número de filtros: impacta na capacidade do modelo de extrair características dos dados;
- Tamanho do kernel: afeta a eficiência do modelo em extrair características dos dados;
- Função de ativação: experimentar as funções ReLU, tanh ou sigmoid;
- Número de Camadas LSTM: aumentar as camadas LSTM.
- Taxa de Dropout: variar esse parâmetro que tem como valores padrões 0.2 e 0.5;
- Número de Épocas: quantidade de vezes que o modelo processa o conjunto de dados de treinamento durante o treinamento;

- Batch Size: número de amostras de dados de treinamento processadas por vez;
- Função de Otimização: alterar os otimizadores entre otimizadores RMSprop, SGD e Adam.

3.7 Avaliação dos resultados

Nessa etapa são comparados os melhores resultados dos indicadores de desempenho RMSE, MSE, MAE e MAPE alcançados por cada um dos quatro modelos nos experimentos realizados. O objetivo é elencar o desempenho dos quatro modelos na tarefa de prever o preço da ação e verificar se há algum modelo que se destaque em relação aos demais.

4 RESULTADOS: AVALIAÇÃO EXPERIMENTAL E DISCUSSÃO

4.1 Modelo Support Vector Regression (SVR)

Após executados os experimentos, o melhor resultado com o modelo Support Vector Regression (SVR) foi obtido com os seguintes dos principais parâmetros:

- Kernel = linear;
- $C = 1$;
- Epsilon = 0.01.

O gráfico com os resultados na fase de Teste do modelo SVR é apresentado na figura abaixo:

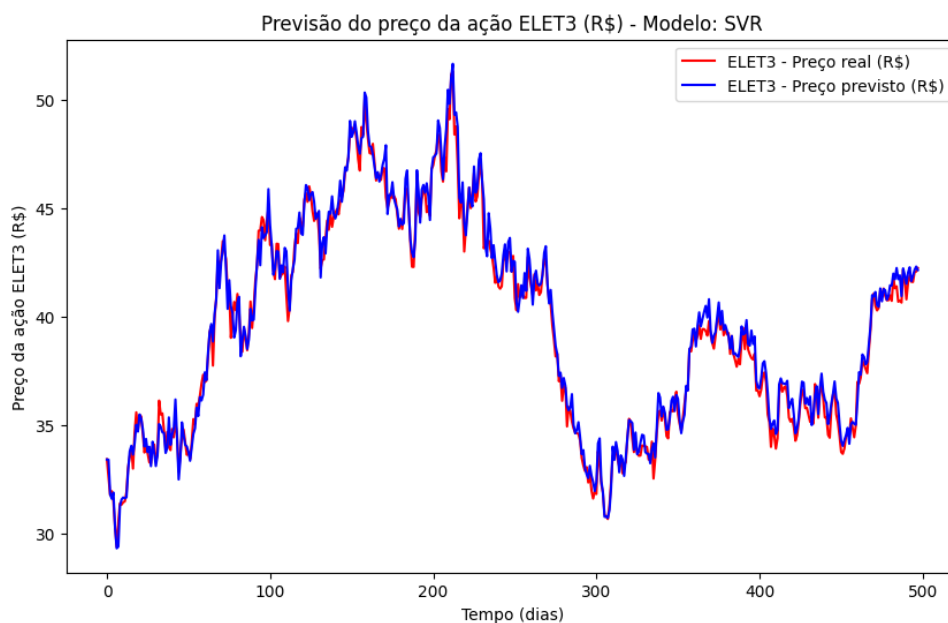


Figura 8 – Resultados na fase de Teste do modelo SVR (elaborado pelo autor)

4.2 Modelo Long Short-Term Memory (LSTM)

Após executados os experimentos, o melhor resultado com o modelo Long Short-Term Memory (LSTM) foi obtido com os seguintes dos principais parâmetros:

- Número de camadas LSTM = 3;
- Unidades LSTM = 50;

- Taxa de Dropout = 0.2;
- Função de Otimização = Adam;
- Número de Épocas = 150;
- Batch Size = 32.

O gráfico com os resultados na fase de Teste do modelo LSTM é apresentado na figura abaixo:

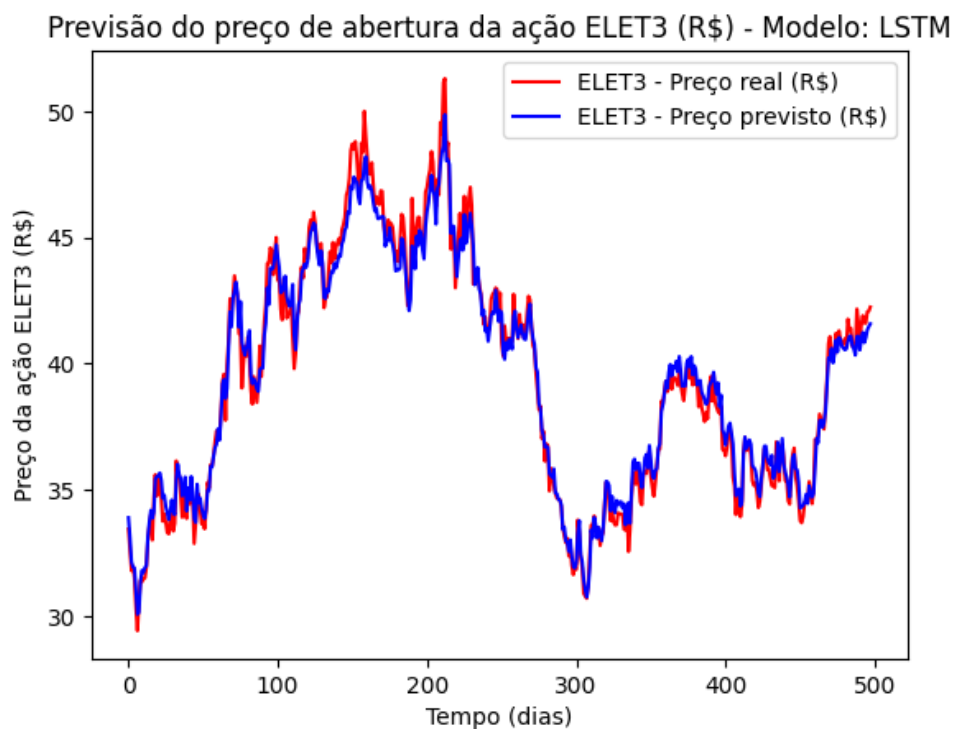


Figura 9 – Resultados na fase de Teste do modelo LSTM (elaborado pelo autor)

4.3 Modelo Convolutional Neural Network (CNN)

Após executados os experimentos, o melhor resultado com o modelo Convolutional Neural Network (CNN) foi obtido com os seguintes dos principais parâmetros:

- Número de filtros:
 - Primeira camada Conv1D = 256;
 - Segunda camada Conv1D = 256;
 - Terceira camada Conv1D = 128;
 - Quarta camada Conv1D = 128.

- Tamanho do Kernel:
 - Primeira camada Conv1D = 7;
 - Segunda camada Conv1D = 5;
 - Terceira camada Conv1D = 3;
 - Quarta camada Conv1D = 3.
- Tamanho da janela de pooling = 2;
- Número de unidades nas camadas densas:
 - Primeira camada Dense = 128;
 - Segunda camada Dense = 64;
 - Camada de saída = 1.
- Função de ativação:
 - Todas as camadas Conv1D = LeakyReLU;
 - Primeira camada Dense = LeakyReLU;
 - Segunda camada Dense = LeakyReLU;
 - Camada de saída = Linear.

O gráfico com os resultados na fase de Teste do modelo CNN é apresentado na figura abaixo:

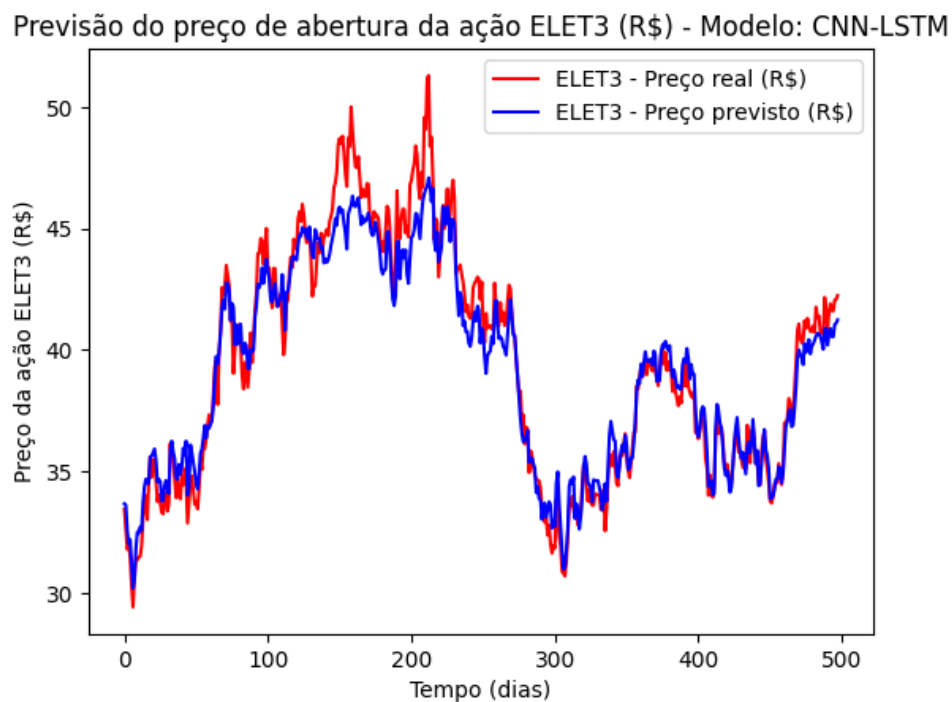


Figura 10 – Resultados na fase de Teste do modelo CNN (elaborado pelo autor)

4.4 Modelo híbrido CNN-LSTM

Após executados os experimentos, o melhor resultado com o modelo híbrido CNN-LSTM foi obtido com os seguintes dos principais parâmetros:

- Número de filtros Conv1D = 64;
- Tamanho do Kernel = 3;
- Função de ativação = ReLU;
- Número de camadas LSTM = 3;
- Unidades LSTM = 50;
- Taxa de Dropout = 0.2;
- Função de Otimização = Adam;
- Número de Épocas = 800;
- Batch Size = 64.

O gráfico com os resultados na fase de Teste do modelo CNN-LSTM é apresentado na figura abaixo:

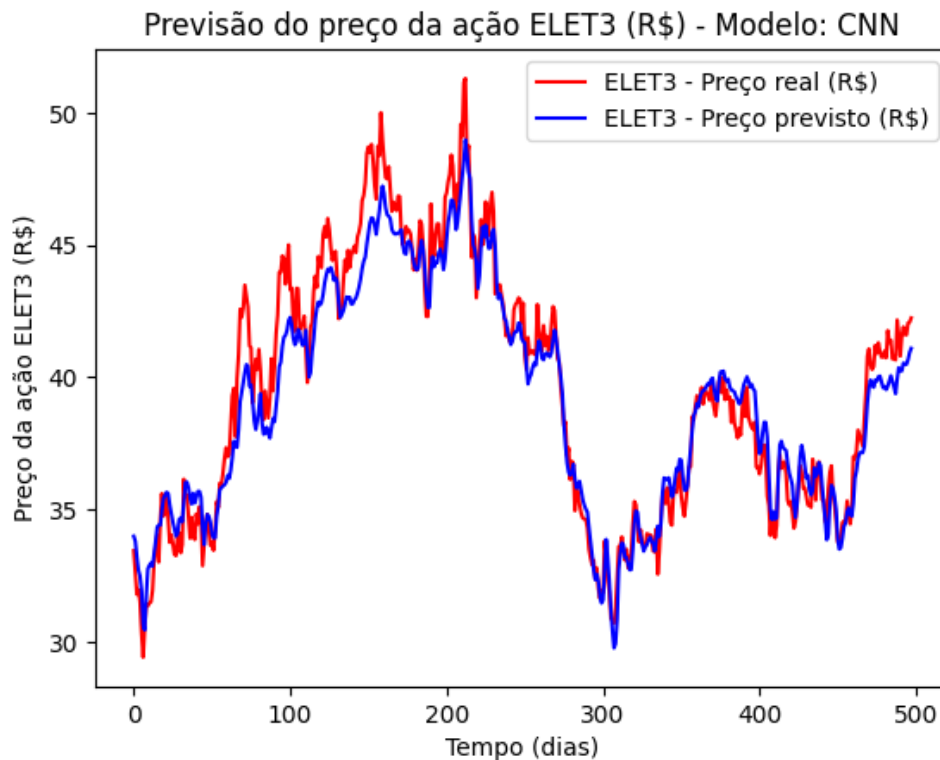


Figura 11 – Resultados na fase de Teste do modelo CNN-LSTM (elaborado pelo autor)

4.5 Resultados dos experimentos e discussão

Os resultados obtidos nos experimentos, expressos através dos indicadores de desempenho MSE, RMSE, MAE e MAPE, estão informados na tabela 8:

Tabela 5 – Resultados obtidos pelos modelos

Modelo	MSE	RMSE	MAE	MAPE
SVR	0.319892837	0.565590698	0.428316575	0.010903779
LSTM	0.384973745	0.620462525	0.483386130	0.012068616
CNN-LSTM	1.107077647	1.052177574	0.788157109	0.019358492
CNN	1.429498396	1.195616325	0.933586650	0.023128867

Observa-se que os resultados dos quatro modelos retratam bom desempenho na previsão do preço da ação. O modelo SVR foi o que apresentou o melhor desempenho, seguido pelo modelo LSTM em segundo lugar. O modelo CNN-LSTM ficou na terceiro posição e o modelo CNN foi o que apresentou o menor desempenho entre os quatro modelos.

Estudos como os de (Sim *et al.*, 2019), (Eapen; Bein; Verma, 2019) e (Nikou; Mansourfar; Bagherzadeh, 2019), entretanto, que fizeram aplicações destes métodos para diferentes preços de ações, indicam uma melhor performance dos modelos de aprendizado profundo. Estas diferenças de resultados podem em parte serem explicadas pelos diferentes tipos de dados utilizados nas pesquisas, ou mesmo por diferenças potenciais em ajustes de parâmetros. Isto aponta para a potencial complexidade do fenômeno a ser estudado e para as ferramentas disponíveis. Por exemplo, dado que o número de observações deste estudo não é relativamente muito grande, isso pode ter reduzido a performance dos classificadores de aprendizagem profunda.

5 CONCLUSÕES

Os experimentos realizados nesse trabalho com os quatro modelos de Inteligência Artificial apresentaram bons resultados na tarefa de previsão do preço da ação ELET3 (ELETROBRAS ON N1) nos pregões da B3 (tabela 5). Porém, o modelo Support Vector Regression (SVR) se mostrou mais eficiente nessa tarefa do que os três modelos de Aprendizado Aprofundo (LSTM, CNN e híbrido CNN-LSTM). Por outro lado, nos sete trabalhos correlatos citados (item 2.4), os modelos de Aprendizado Profundo apresentaram melhor desempenho nessa tarefa que modelos SVM e SVR.

Isso aponta a complexidade da tarefa de previsão do preço de ações, mesmo utilizando de soluções de Inteligência Artificial. Desde a escolha das variáveis do arquivo de dados, a escolha do modelo de Inteligência Artificial a ser utilizado, a estrutura desse modelo e o ajuste de seus parâmetros, afetam significativamente o resultado da tarefa. Daí porque observamos os vários estudos e experimentos realizados nessa área, buscando a melhor eficiência possível.

Como trabalhos futuros sugere-se a utilização de outros modelos de Inteligência Artificial nessa tarefa, como os de Regressão Avançada (ex.: XGBoost e LightGBM), de Redes Neurais Recorrentes Avançadas (ex.: GRU) e outros modelos de Aprendizado Profundo (ex.: Transformer).

REFERÊNCIAS

ARTASANCHEZ, A.; JOSHI, P. **Artificial Intelligence with Python: Your complete guide to building intelligent apps using Python 3**. x. 2. ed. Birmingham B3 2PB, UK: Packt Publishing Ltd, 2020.

B3. **IEE B3**: Índice de energia elétrica. 2024. Disponível em: https://www.b3.com.br/pt_br/market-data-e-indices/indices/indices-de-segmentos-e-setoriais/indice-de-energia-eletrica-iee-b3.htm. Acesso em: 03 fev. 2024.

BATHLA, G. Stock price prediction using lstm and svr. *In*: IEEE. **2020 Sixth International Conference on Parallel, Distributed and Grid Computing (PDGC)**. [S.l.: s.n.], 2020. p. 211–214.

BLOOMBERGNEF. **CLIMATESCOPE 2023**: Power transition factbook. 2023. Disponível em: <https://www.global-climatescope.org/downloads/climatescope-2023-report-en.pdf>. Acesso em: 11 jan. 2024.

CARMO, C. R. S.; SILVA, J. R. de M. Aprendizado de máquina e prestação de serviços de armazenamento de dados: Métricas para análise e validação de algoritmos previsores. **Revista GeTeC**, v. 12, n. 38, 2023.

COMISSÃO DE VALORES MOBILIÁRIOS. **Mercado de valores mobiliários brasileiro**. 4. ed. Rio de Janeiro: Comissão de Valores Mobiliários, 2019.

DERBENTSEV, V. *et al.* Machine learning approaches for financial time series forecasting. *In*: CEUR WORKSHOP PROCEEDINGS. 2020. Disponível em: <http://ceur-ws.org/Vol-2713/paper47.pdf>. Acesso em: 05 dez. 2023.

DUA, S. *et al.* Developing a speech recognition system for recognizing tonal speech signals using a convolutional neural network. **Applied Sciences**, MDPI, v. 12, n. 12, p. 6223, 2022.

EAPEN, J.; BEIN, D.; VERMA, A. Novel deep learning model with cnn and bi-directional lstm for improved stock market index prediction. *In*: IEEE. **2019 IEEE 9th annual computing and communication workshop and conference (CCWC)**. Las Vegas, NV, USA, 2019. p. 0264–0270.

FORTUNA, E. **Mercado financeiro: produtos e serviços**. 20. ed. Rio de Janeiro: Qualitymark Editora Ltda, 2015.

GANI, W.; TALEB, H.; LIMAM, M. Support vector regression based residual control charts. **Journal of Applied Statistics**, Taylor & Francis, v. 37, n. 2, p. 309–324, 2010.

GENG, Z. *et al.* Novel target attention convolutional neural network for relation classification. **Information sciences**, Elsevier, v. 597, p. 24–37, 2022.

HEATON, J. Ian goodfellow, yoshua bengio, and aaron courville: Deep learning: The mit press, 2016, 800 pp, isbn: 0262035618. **Genetic programming and evolvable machines**, Springer, v. 19, n. 1, p. 305–307, 2018.

HEGAZY, O.; SOLIMAN, O. S.; SALAM, M. A. A machine learning model for stock market prediction. **arXiv preprint arXiv:1402.7351**, 2014.

LAKSHMINARAYANAN, S. K.; MCCRAE, J. P. A comparative study of svm and lstm deep learning algorithms for stock market prediction. *In*: NUI GALWAY. **AICS**. Galway, Ireland, 2019. p. 446–457.

LECUN, Y. *et al.* On-line handwriting recognition with neural networks: Spatial representation versus temporal representation. *In*: ECOLE NATIONALE SUPERIEURE DES TELECOMMUNICATIONS. **Proceedings of the International Conference on Handwriting and Drawing**. Paris, France, 1993.

LECUN, Y.; BENGIO, Y.; HINTON, G. Deep learning. **nature**, Nature Publishing Group UK London, v. 521, n. 7553, p. 436–444, 2015.

LECUN, Y. *et al.* Gradient-based learning applied to document recognition. **Proceedings of the IEEE**, Ieee, v. 86, n. 11, p. 2278–2324, 1998.

LIU, W. *et al.* A survey of deep neural network architectures and their applications. **Neurocomputing**, Elsevier, v. 234, p. 11–26, 2017.

MINASNY, B. **The Elements of Statistical Learning, Trevor Hastie, Robert Tibshirani, Jerome Friedman, (2009), Springer Series in Statistics, ISBN 0172-7397, 745 pp.** [*S.l.: s.n.*]: Elsevier, 2009.

NIKOU, M.; MANSOURFAR, G.; BAGHERZADEH, J. Stock price prediction using deep learning algorithm and its comparison with machine learning algorithms. **Intelligent Systems in Accounting, Finance and Management**, Wiley Online Library, v. 26, n. 4, p. 164–174, 2019.

OLAH, C. *et al.* Understanding lstm networks. 2015. Disponível em: colah.github.io/posts/2015-08-Understanding-LSTMs/. Acesso em: 14 mar. 2024.

ONS - OPERADOR NACIONAL DO SISTEMA. **Histórico da Operação**: Composição, geração de energia por tipo de usina. 2023. Disponível em: https://www.ons.org.br/Paginas/resultados-da-operacao/historico-da-operacao/geracao_energia.aspx. Acesso em: 11 jan. 2024.

RANA, M.; UDDIN, M. M.; HOQUE, M. M. Effects of activation functions and optimizers on stock price prediction using lstm recurrent networks. *In*: ASSOCIATION FOR COMPUTING MACHINERY. **Proceedings of the 2019 3rd International Conference on Computer Science and Artificial Intelligence**. Normal, IL, USA, 2019. p. 354–358.

SABBAR, K.; KHARRIM, M. E. Average variance portfolio optimization using machine learning-based stock price prediction case of renewable energy investments. *In*: EDP SCIENCES. **E3S Web of Conferences**. France, 2023. v. 412, p. 01077.

SAPANKEVYCH, N. I.; SANKAR, R. Time series prediction using support vector machines: a survey. **IEEE computational intelligence magazine**, IEEE, USA, v. 4, n. 2, p. 24–38, 2009.

SHARAF, M. *et al.* Stockpred: a framework for stock price prediction. **Multimedia Tools and Applications**, Springer, v. 80, n. 12, p. 17923–17954, 2021.

SILVA, I. N. d.; SPATTI, D. H.; FLAUZINO, R. A. Redes neurais artificiais para engenharia e ciências aplicadas. Artliber, São Paulo, 2010.

SIM, H. S. *et al.* Is deep learning for image recognition applicable to stock market prediction? **Complexity**, Hindawi, v. 2019, 2019. Disponível em: <https://www.hindawi.com/journals/complexity/2019/4324878/>. Acesso em: 15 jan. 2024.