UNIVERSIDADE ESTADUAL PAULISTA "JÚLIO DE MESQUITA FILHO"
FACULDADE DE CIÊNCIAS - CAMPUS BAURU
DEPARTAMENTO DE COMPUTAÇÃO
BACHARELADO EM CIÊNCIA DA COMPUTAÇÃO

LUCIANO EIJI TANAKA

VISUALIZAÇÃO DE DADOS E ANÁLISE DO MERCADO DE AÇÕES BRASILEIRO

BAURU
2022

LUCIANO EIJI TANAKA

VISUALIZAÇÃO DE DADOS E ANÁLISE DO MERCADO DE AÇÕES BRASILEIRO

Trabalho de Conclusão de Curso do Curso de Ciência da Computação da Universidade Estadual Paulista "Júlio de Mesquita Filho", Faculdade de Ciências, Campus Bauru.
Orientador: Prof. Dr. João Pedro Albino

BAURU
2022

T161v

Tanaka, Luciano Eiji
Visualização de dados e análise do mercado de ações brasileiro / Luciano Eiji Tanaka. -- Bauru, 2022
46 f. : il.

Trabalho de conclusão de curso (Bacharelado - Ciência da Computação) - Universidade Estadual Paulista (Unesp), Faculdade de Ciências, Bauru
Orientador: João Pedro Albino

1. Mercado de ações. 2. Máquinas de vetor de suporte. 3. Redes neurais artificiais. 4. Long short-term memory (LSTM). 5. Visualização de dados. I. Título.

Sistema de geração automática de fichas catalográficas da Unesp. Biblioteca da Faculdade de Ciências, Bauru. Dados fornecidos pelo autor(a).

Essa ficha não pode ser modificada.

Luciano Eiji Tanaka

Visualização de dados e análise do mercado de ações brasileiro

Trabalho de Conclusão de Curso do Curso de Ciência da Computação da Universidade Estadual Paulista "Júlio de Mesquita Filho", Faculdade de Ciências, Campus Bauru.

Banca Examinadora

Prof. Dr. João Pedro Albino
Orientador
Universidade Estadual Paulista "Júlio de Mesquita Filho"
Faculdade de Ciências
Departamento de Ciência da Computação

Profa. Dra Simone das Graças Domingues Prado
Universidade Estadual Paulista "Júlio de Mesquita Filho"
Faculdade de Ciências
Departamento de Ciência da Computação

Prof. Dr Kelton Augusto Pontara da Costa
Universidade Estadual Paulista "Júlio de Mesquita Filho"
Faculdade de Ciências
Departamento de Ciência da Computação

Bauru, _____ de _____ de _____.

Agradecimentos

Agradeço a minha família, que possibilitou toda essa jornada, me dando suporte, conselhos e foi meu porto seguro para todos os momentos. Agradeço também aos meus amigos e colegas que estiveram sempre ao meu lado, seja como companhia, diversão ou crescimento pessoal, deixando tudo mais leve. Agradeço aos meus professores e a todos outros profissionais de educação e infraestrutura, direta ou indiretamente, que me ensinaram e guiaram por essa etapa. E por fim, agradeço a todas as pessoas que conheci durante esse período, em momentos de paz e de conflito, sem elas não poderia ter crescido tanto e aprendido sobre o mundo.

"De nada tenho certeza, mas a visão das estrelas me faz sonhar."
Van Gogh

Resumo

O mercado de ações é uma das áreas mais populares dentro do mercado financeiro, hoje em dia, através da popularização da internet, da mídia e da democratização da informação, tornou-se uma das formas mais comuns de obtenção de renda alternativa. A previsão do preço das ações sempre foi muito pesquisada, mas devido à sua natureza dinâmica e volátil, é considerada uma das tarefas mais difíceis no campo da matemática e da ciência da computação. O mercado é afetado por vários fatores macroeconômicos, como políticas governamentais, relações internacionais, cenário econômico, expectativas e psicologia dos investidores, etc. O presente trabalho busca combinar sistemas inteligentes para prever os preços finais das ações do índice Bovespa, que são as mais consolidadas e negociadas no mercado. Conceitos sobre mercado financeiro, aprendizado de máquina e visualização de dados serão estudados para dar suporte ao projeto. Foram propostos modelos para analisar e encontrar padrões nos preços das ações, bem como indicar tendências de longo prazo, e por fim a implementação do projeto para uso geral. Para as análises foram utilizados os dados de janeiro de 2012 a julho de 2022, disponíveis na seção de cotações históricas diretamente do site da B3. O estudo foi desenvolvido utilizando redes neurais e uma máquina de vetores de suporte utilizando dados do preço de fechamento das ações negociadas na bolsa de valores.

Palavras-chave: Redes neurais, redes neurais artificiais, máquina de vetor de suporte, mercado de ações, LSTM, previsão de valores, ciência de dados.

Abstract

The stock market is one of the most popular areas within the financial market, nowadays, through the popularization of the internet, the media and the democratization of information, it has become one of the most common approaches to earning alternative income. Stock price prediction has always been heavily researched, but given its dynamic and volatile nature, it is considered one of the most difficult tasks in the field of mathematics and computer science. The market is affected by several macroeconomic factors such as government policies, international relations, economic scenario, investor expectations and psychology etc. The present work seeks to combine intelligent systems to forecast the final stock prices of the Bovespa index, which are the most established and traded in the market. Concepts about financial market, machine learning and data visualization will be studied to support the project. There were proposed models to analyze and find patterns in stock prices, as well as indicate long-term trends, and finally the implementation of the project for general use. For the analyses, data from January 2012 to July 2022 were used, available in the historical quotations section directly from the B3 website. The study developed using neural networks and a support vector machine using data from the closing price of stocks traded on the stock exchange.

Keywords: Neural networks, artificial neural networks, support vector machine, stock Market, LSTM, price forecasting, data science.

Lista de figuras

Lista de abreviaturas e siglas

AM	Aprendizado de Máquina
API	Application Programming Interface
B3	Brasil, Bolsa, Balcão
LSTM	Long Short-Term Memory
MM	Média móvel
MME	Média Móvel Exponencial
MAE	Mean Absolute Error
MACD	Moving Average Convergence/Divergence
MVS	Máquina de Vetores de Suporte
RL	Regressão Linear
RMSE	Root Mean Squared Error
RN	Redes Neurais
RNA	Redes Neurais Artificiais
RNN	Redes Reurais Recorrentes
PLN	Processamento de Linguagem Natural
SVM	Suport Vector Machine

Sumário

1 Introdução

De acordo com Tan Quek (2007), o mercado de ações é afetado por muitos macro fatores econômicos, como eventos políticos, políticas da empresa, condições econômicas gerais, expectativas dos investidores, decisões de investidores institucionais, movimento de outros mercados de ações, psicologia dos investidores e etc.

Dado o contexto, em conjunto com a velocidade e alcance das informações nas mídias e redes, o número de pessoas físicas que investem na bolsa como alternativa de renda, vem crescendo. Com tamanha demanda, o mercado de investimentos cresceu proporcionalmente com o incentivo de democratizar a educação financeira através de plataformas de investimento, cursos, mídias sociais, entre outros meios de comunicação VIANA (2022). Mesmo com acesso à informação em ascensão, o conhecimento necessário para se analisar investimentos é grande, em razão da complexidade e quantidade dos dados. Isso motiva o uso de aprendizado de máquina como abordagem para sanar tal problema. Uma análise para tomar decisões é uma atividade complexa e requer tempo e estudo para compreender o problema, desenvolver e treinar um modelo que antecipe satisfatoriamente o mercado. Neste trabalho, utilizaremos dois tipos de análises de mercado: a análise fundamentalista e a análise técnica, ambas serão utilizadas para desenvolver o modelo final.

A primeira, respectivamente, baseia-se na análise quantitativa, qualitativa e temporal dos fundamento da empresa, traduzida em diversos índices e indicadores econômico financeiros e de mercado, e visa basicamente, avaliar o desempenho da empresa, como forma de identificar os resultado (consequência) retrospectivos e prospectivos das diversas decisões financeiras tomadas e a partir disso, fazer um prognóstico da empresa Malta (2016).

A segunda, é o estudo da ação do mercado, primariamente por uso de gráficos, com o objetivo de prever as tendências de preços. Outra definição é que a análise técnica é a interpretação da ação do mercado para antecipar ou prever o movimento futuro dos preços LEMOS (2015).

Devido ao crescimento no campo financeiro e na participação da população no Brasil, o intuito deste projeto é contribuir no desenvolvimento e melhoria de modelos de inteligência artificial para prever preços de ações, indicar tendências e expandir o acesso a tais informações complexas através da visualização de dados.

1.1 Problemática

Em meio a crises econômicas, é comum que a população invista em alternativas de renda, como aplicações financeiras. Em consequência, o aquecido mercado se desenvolveu em

razão da demanda por investimentos e crescimento de empresas, o mercado de investimentos cresceu proporcionalmente com o incentivo de democratizar a educação financeira através de plataformas de investimento, cursos, mídias sociais, entre outros. De certa forma, entre tantas alternativas, a credibilidade nos meios de investimento fica baixa, distanciando o público e reduzindo o acesso à informação.

Desse modo, pesquisas e estudos fomentam abordagens para solucionar tais problemas como a incerteza do mercado de ações e o consumo de informações complexas para tomada de decisão, a inteligência artificial e a visualização de dados se tornam indispensáveis.

A previsão do valor de ações é considerada um problema complexo dada sua natureza altamente dinâmica, segundo Abu-Mostafa (1996) é considerada uma tarefa muito difícil a previsão do mercado de ações em previsão de séries temporais financeiras.

1.2 Justificativa

Em fundamento da incerteza do mercado de ações, seu impacto na sociedade e crescimento da participação da população, constata-se a indispensabilidade na contribuição do desenvolvimento de sistemas inteligentes mais precisos, que geram potencial de ganhos a longo prazo para investidores consistentemente e tornam aplicações financeiras acessíveis e viáveis à população. Esse crescimento é muito significante para o país, pois o desenvolvimento do mercado de capitais é uma opção de financiamento e capitalização das empresas, em relação ao investimento em seus projetos, produtos e serviços, além de pessoas físicas terem uma alternativa de renda e criação de patrimônio a longo prazo.

1.3 Objetivos

Contribuir com o mercado de ações desenvolvendo um modelo que descreva a realidade satisfatoriamente, obtendo conclusões e resultados que possam embasar a tomada de decisão e impactar positivamente em uma carteira de ativos. Além de fomentar o mercado e aproximar a população de aplicações financeiras quebrando barreiras de complexidade da informação, proporcionando novas possibilidades de investimentos. Promover boas práticas no desenvolvimento em todas as etapas do processo, com técnicas e linguagens modernas utilizadas no mercado.

1.3.1 Objetivo Geral

Desenvolver um sistema inteligente com precisão satisfatória comparado a modelos recentes, estimular investidores a participarem do mercado de ações, transformar dados complexos em conhecimento para tomada de decisão.

1.3.2 Objetivos Específicos

Analisar, estudar e se aprofundar nos conceitos que compõem o mercado de ações. Coletar, raspar e armazenar dados das ações ao longo de um determinado período. Processar os dados para que além de padronizados, sejam confiáveis e correlacionados. Analisar tecnicamente tendências no mercado e avaliar empresas e seus prospectivos utilizando ferramentas de visualização. Desenvolver modelo para prever o preço final diário das ações consistentemente com precisão satisfatória. Verificar hipóteses e estimativas para impactar positivamente um investidor leigo.

2 Fundamentação Teórica

Nesta seção serão introduzidos conceitos sobre o mercado financeiro, séries temporais, RNAs, MVS. Estes fundamentos serão a base para a compreensão do projeto final.

2.1 Bolsa de valores

Bolsa de valores, é mercado organizado para negociação de ações e títulos. Esses mercados foram originalmente abertos a todos, mas atualmente apenas os membros da associação proprietária podem comprar e vender diretamente. Os associados, ou corretores da bolsa, compram e vendem para si ou para terceiros, cobrando comissões por seus serviços. Uma ação pode ser comprada ou vendida apenas se estiver listada em uma bolsa e não pode ser listada a menos que atenda a certos requisitos estabelecidos pelo conselho de administração da bolsa (TEWELES, 1992).

A B3 é uma das principais empresas de infraestrutura de mercado financeiro no mundo, com atuação em ambiente de bolsa e de balcão. Sociedade de capital aberto – cujas ações (B3SA3) são negociadas no Novo Mercado –, a Companhia integra os índices Ibovespa, IBrX-50, IBrX e Itag, entre outros. Reúne ainda tradição de inovação em produtos e tecnologia e é uma das maiores em valor de mercado, com posição global de destaque no setor de bolsas.

2.1.1 Índice Ibovespa

O Ibovespa é o principal indicador de desempenho das ações negociadas na B3 e reúne as empresas mais importantes do mercado de capitais brasileiro. A partir dele foram selecionadas 4 ações volumosas de setores variados:

PETR4 - Petrobras PN EDJ N2 Indústria de óleo, Gás natural e Energia

ABEV3 - AMBEV S/A ON Consumo e Varejo

ITUB4 - ITAU UNIBANCO PN N1 Varejo e

GOLL4 - GOL PN N2 Transporte e Logística

2.2 Análise técnica

Analisa dados mensuráveis das atividades do mercado de ações, como preços de ações, retornos históricos e volume de negócios históricos; ou seja, informações quantitativas que podem identificar sinais de negociação e capturar os padrões de movimento do mercado de ações.

A análise técnica se concentra em dados históricos e dados atuais, assim como a análise fundamental, mas é usada principalmente para fins de negociação de curto prazo.

Devido à sua natureza de curto prazo, os resultados da análise técnica são facilmente influenciados por notícias.

As metodologias populares de análise técnica incluem média móvel, níveis de suporte e resistência, bem como linhas e canais de tendência.

Este trabalho se baseia na análise técnica onde se entende que os preços dos ativos refletem a expectativa dos investidores no mercado, onde osciladores captam pontos de inflexão na série de preços, e os mistos tentam indicar a psicologia de massa do mercado, isto é, as expectativas dos investidores em relação ao mercado Elder (1993). Neste projeto foram utilizados 4 indicadores técnicos, sendo 3 rastreadores de tendência (média móvel simples, média móvel exponencial, histograma MACD) e 1 oscilador (índice de força relativa).

2.2.1 Média móvel simples

A MM, ou média móvel simples, é uma das médias mais utilizadas por sua facilidade de entendimento e interpretação. O cálculo da mesma é com relação ao preço 21 médio do ativo, na maioria das vezes sendo usado o preço de fechamento, e com base em um número de períodos específicos. A fórmula de calcular a MM é baseada na soma dos preços de fechamento dos últimos "n" períodos e dividir o resultado encontrado pelos mesmos "n" períodos Debastiani (2008). Segue sua equação:

$$MM = \frac{P_1 + P_2 + ... + P_n}{N}$$

2.2.2 Média móvel exponencial

O cálculo da MME é um pouco mais elaborado porque os preços mais próximos dos dias recebem mais peso do que os preços mais antigos. E esta é uma diferença marcante dela para a MM: a base móvel exponencial se baseia em dados históricos recentes em vez de dar um peso igual a todos os preços em um intervalo de tempo. O autor Debastiani (2008) define o cálculo da média móvel exponencial da seguinte forma: Primeiro há a escolha do número de períodos, sendo ele denominado "n". Exemplo = sete dias, com isso, n = 7.

Segundo passo é calcular o coeficiente chamado de K, que será utilizado: K = 2 / (n+1). Seguindo o exemplo de 7 dias têm-se: K = 2 / (7+1), portanto K = 0,25.

Terceiro passo é calcular a média móvel exponencial:

$$MME = P_t * K + MME_{t-1} * (1 - K)$$

Onde o Fech-hoje é representado pelo fechamento do dia. MME-ontem que corresponde ao valor calculado no dia anterior.

Quando calculado a primeira MME, no dia posterior essa média se torna a MME-ontem na fórmula. Assim, continuando o ciclo a cada novo dia de pregão.

2.2.3 MACD

O MACD original é um dos indicadores mais utilizados na análise técnica. Ele é um rastreador de tendência composto por três MMEs. Os períodos das MMEs ficam a critério do investidor, mas os mais utilizados são de 26 dias, de 12 dias, e de 9 dias. O MACD é representado por duas linhas, a linha MACD e a linha de sinal. A linha MACD é composta pela subtração da MME de 12 dias sobre os preços de fechamento pela MME de 26 dias sobre os preços de fechamento. A linha de sinal é a MME de 9 dias sobre a linha MACD. Finalmente o histograma MACD é calculado a partir da subtração da linha MACD pela linha de sinal. O histograma MACD proporciona uma visão não só da tendência mas também da sua força Elder (1993).

2.3 Análise fundamentalista

Avalia as ações de uma empresa examinando seu valor intrínseco, incluindo, entre outros, ativos tangíveis, demonstrações financeiras, eficácia da gestão, iniciativas estratégicas e comportamentos do consumidor; essencialmente todos os fundamentos de uma empresa.

Sendo um indicador relevante para investimentos de longo prazo, a análise fundamentalista se baseia em dados históricos e atuais para medir receitas, ativos, custos, passivos e assim por diante.

2.4 Aprendizado de máquina

O aprendizado de máquina é uma das principais subáreas da inteligência artificial, e é composto por uma coleção de métodos criados a partir de modelos matemáticos baseados na teoria estatística que permitem aos computadores automatizar tarefas com base na descoberta sistemática de padrões nos conjuntos de dados disponíveis ou em experiências passadas (BHAVSAR et al., 2017).

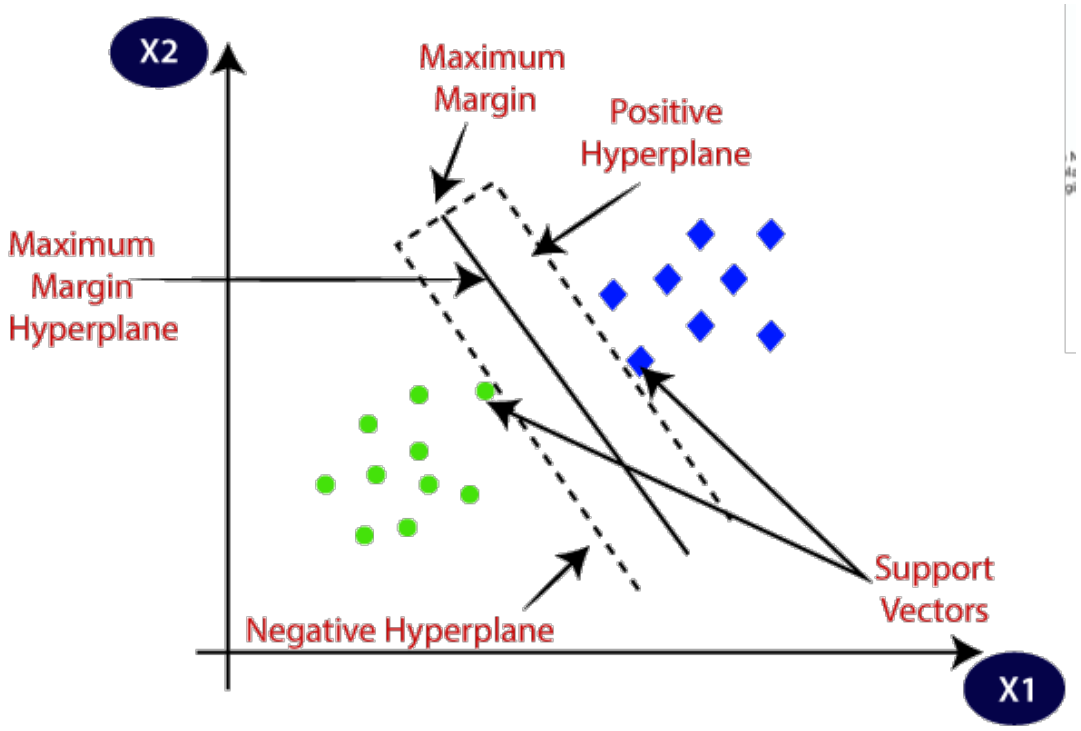
2.4.1 Aprendizagem supervisionada

No aprendizado supervisionado, o modelo recebe um conjunto de entradas com suas respectivas saídas e busca encontrar uma função que estabeleça uma relação aproximada entre elas. Mais formalmente, o modelo baseado no aprendizado supervisionado busca encontrar uma

função h(xi), denominada hipótese, que se aproxime da função f(xi), onde f(xi) é a saída da i-ésima entrada de x (RUSSEL; NORVIG, 2002).

2.5 Máquinas de vetor de suporte

Figura 1 – Representação gráfica da SVM



Fonte: Retirado de LinuxHint, "How to Predict Stock Price Using SVM", através do link: <https://linuxhint.com/predict-stock-price-svm>. Acessado em 20 de dezembro de 2022.

Categoria de redes alimentadas adiante universais, proposta por Boser, Guyon e Vapnik (1996). Assim como Perceptrons multicamadas podem ser utilizadas para classificação de padrões e regressão linear. Basicamente uma máquina de vetor de suporte e uma máquina linear com algumas propriedades especiais: padrões separáveis que podem surgir no contexto de classificação de padrões. A ideia principal é a construção de um Hiperplano como superfície de decisão de tal forma que a margem de separação entre exemplo positivos e negativos seja máxima. A máquina apresenta esta propriedade desejável seguindo uma abordagem fundamentada na teoria da aprendizagem estatística.

Mais precisamente a SVM, é uma implementação do método de minimização estrutural de risco, esse princípio indutivo e baseado no fato de que a taxa de erro de uma máquina de aprendizagem sobre dados de teste (taxa de erro de generalização) é limitada pela soma da taxa de erro de treinamento r por um termo que depende da dimensão de Vapnik-Chervonenkis(V-C). No caso de padrões separáveis, uma SVM produz um valor de zero para o primeiro termo e minimiza o segundo termo, consequentemente a máquina de vetor de suporte pode fornecer

um bom desempenho de generalização em problemas de classificação de padrões, apesar de não incorporar conhecimento do domínio dos problemas. Este atributo é único das SVMs.

Uma noção que é central a construção do algoritmo de aprendizagem por SVM, e o núcleo do produto interno entre um "vetor de suporte" e o vetor retirado do espaço de entrada.

Os vetores de suporte consistem de um pequeno subconjunto dos dados de treinamento extraídos pelo algoritmo. Dependendo de como este núcleo de produto interno é gerado, podemos construir diferentes máquinas de aprendizagem, caracterizadas por superfícies de decisão não-lineares próprias. Desse modo, podemos usar o algoritmo de aprendizagem por vetor de suporte para construir os três seguintes tipos: máquinas de aprendizagem polinomial, redes de função de base radial, perceptrons de duas camadas (com uma única camada oculta).

Para cada uma dessas redes alimentadas adiante podemos usar o algoritmo de aprendizagem por veto de suporte para implementar o processo de aprendizagem, usando um determinado conjunto de dados de treinamento, determinando automaticamente o número necessário de unidades ocultas. Enquanto que o algoritmo de retropropagação é planejado especificamente para treinar um perceptron de múltiplas camadas, o algoritmo de aprendizagem por vetor de suporte e a natureza mais genérica, porque tem uma aplicabilidade mais ampla.

SVM ou máquina de vetor de suporte e uma forma superior de machine learning, usando algoritmos de classificação para problemas de classificação de dois grupos; Quando dadão qualquer conjuntos de dados de treinamento rotulado para categorias separadas, SVM pode categorizar novos textos, podengos performas a sua capacidade ótima quando e provido dados limitados. Dados linearmente separáveis e sua função primária, podendo facilmente classificá-los, ele também tende a ser mais rápido e eficiente em comparação a ANN (quando o conjunto amostral está na casa dos milhares). O algoritmo visa formas a melhor linha para diferenciar espaços de n-dimensões, em grupos para novos dados serem facilmente divididos em suas respectivas categorias nos determinados tempos. O melhor hiperplano e o que maximiza a margem dos dados de treino como na figura. SVM linear e para dados linearmente separáveis (divisíveis estatisticamente em dois grupos por uma única linha reta). Dados linear são classificados com Linear SVM classifier. SVM não-linear opera em dados não-linearmente separáveis (estatisticamente não podem ser divididos em dois grupos usando uma única linha reta). São classificados usando a Non-Linear SVM classifier.

Algumas vantagens do SVM são:

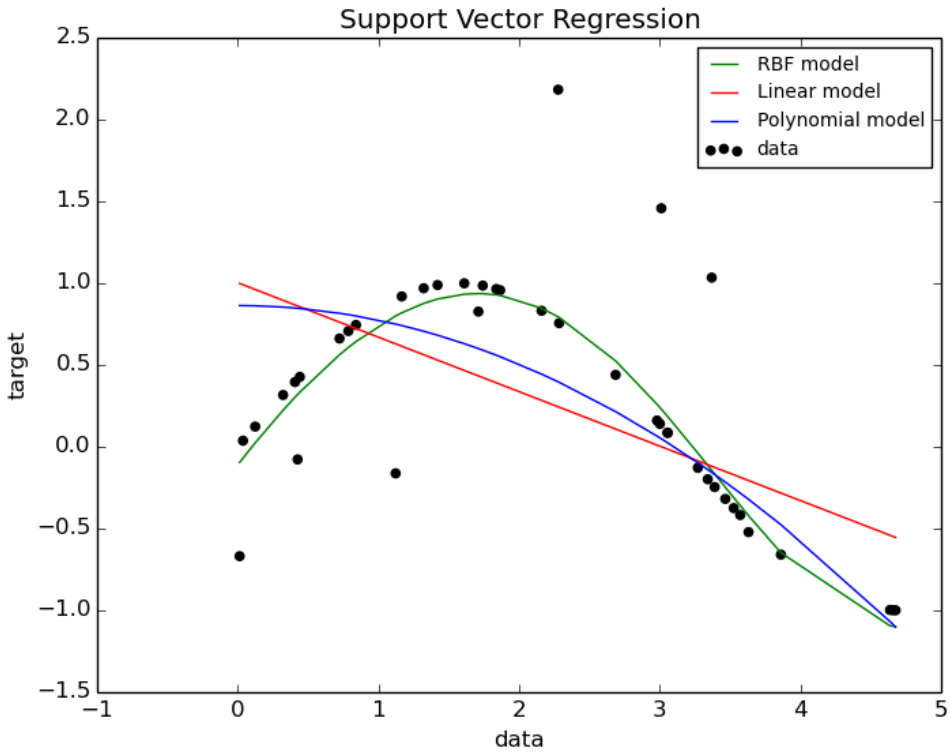
- São mais produtivos com uma clara margem de separação;

- Poderosos em espaços com muitas dimensões;

- Funciona bem quando o número de amostra excede o número de dimensões;

- Também é excelente para lembrar dados, usando um subconjunto de pontos de treinamento chamados vetores de suporte.

Figura 2 – Representação gráfica da SVR



Fonte: Retirado de LinuxHint, "How to Predict Stock Price Using SVM", através do link: <https://linuxhint.com/predict-stock-price-svm>. Acessado em 20 de dezembro de 2022.

Sua maior desvantagem é a performance lenta para grandes quantidades de dados processados devido a mais períodos de treinamento.

2.5.1 Kernel RBF

A função kernel é apenas uma função matemática que converte um espaço de entrada de baixa dimensão em um espaço de dimensão superior. Isso é feito mapeando os dados em um novo espaço de recursos. Neste espaço, os dados serão linearmente separáveis. Isso significa que uma máquina de vetores de suporte pode ser usada para encontrar um hiperplano que separe os dados (PYCODEMATES, 2022).

RBF ou função de base radial, é um kernel muito poderoso usado no SVM. Ao contrário dos kernels lineares ou polinomiais, o RBF é mais complexo e eficiente ao mesmo tempo em que pode combinar vários kernels polinomiais várias vezes de diferentes graus para projetar os dados separáveis não-linearmente em um espaço dimensional superior para que possam ser separáveis usando um hiperplano.

2.5.1.1 C

O parâmetro C compensa a classificação correta dos exemplos de treinamento contra a maximização da margem da função de decisão. Para valores maiores de C, uma margem menor será aceita se a função de decisão for melhor em classificar corretamente todos os pontos de treinamento. Um C menor encorajará uma margem maior, portanto, uma função de decisão mais simples, ao custo da precisão do treinamento. Em outras palavras, C se comporta como um parâmetro de regularização no SVM.

2.5.1.2 Gamma

O parâmetro gama define até onde a influência de um único exemplo de treinamento alcança, com valores baixos significando "longe"e valores altos significando "próximo". Em outras palavras, com gama baixa, pontos distantes da linha de separação plausível são considerados no cálculo da linha de separação. Quando gama alta significa que os pontos próximos à linha plausível são considerados no cálculo. Quando o gama é muito pequeno, o modelo é muito restrito e não consegue capturar a complexidade ou "forma" dos dados. A região de influência de qualquer vetor de suporte selecionado incluiria todo o conjunto de treinamento. O modelo resultante se comportará de forma semelhante a um modelo linear com um conjunto de hiperplanos que separam os centros de alta densidade de qualquer par de duas classes.

Para valores intermediários, podemos ver no segundo gráfico que bons modelos podem ser encontrados em uma diagonal de C e gama. Modelos suaves (valores de gama mais baixos) podem se tornar mais complexos aumentando a importância de classificar cada ponto corretamente (valores de C maiores), portanto, a diagonal de modelos de bom desempenho.

Finalmente, pode-se também observar que, para alguns valores intermediários de gama, obtemos modelos com desempenho igual quando C se torna muito grande. Isso sugere que o conjunto de vetores de suporte não muda mais. O raio do kernel RBF sozinho atua como um bom regularizador estrutural. Aumentar C ainda não ajuda, provavelmente porque não há mais pontos de treinamento em violação (dentro da margem ou classificados incorretamente), ou pelo menos nenhuma solução melhor pode ser encontrada. Sendo as pontuações iguais, pode fazer sentido usar os valores de C menores, uma vez que valores de C muito altos geralmente aumentam o tempo de adaptação.

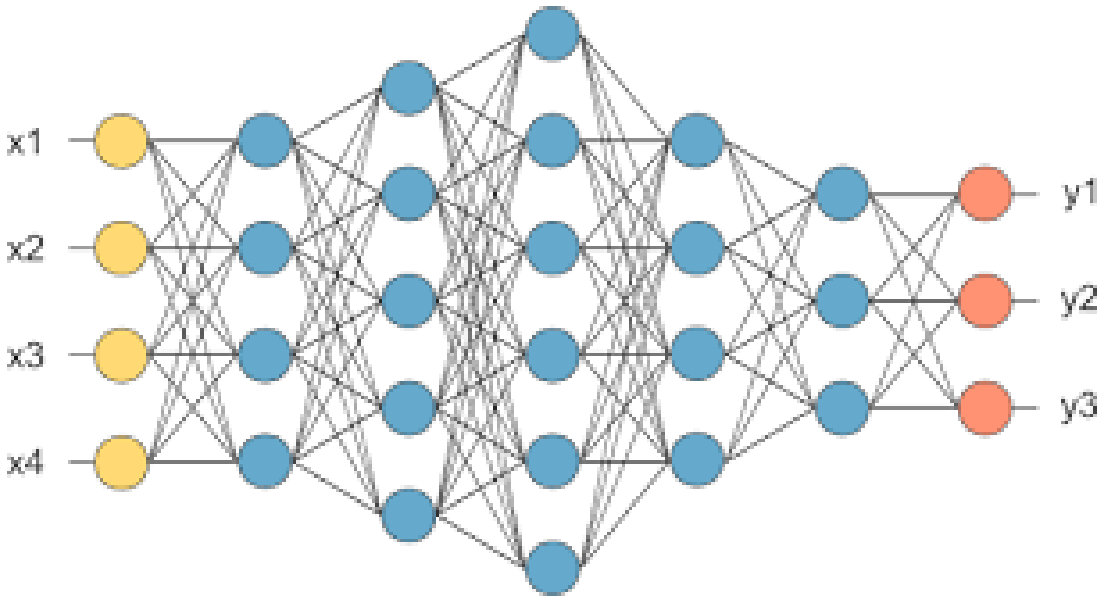
2.6 Redes Neurais

O trabalho em redes neurais artificiais, usualmente denominadas "redes neurais", têm sido motivado desde o começo pelo reconhecimento de que o cérebro humano processa informações de uma forma inteiramente diferente do computador digital convencional. O cérebro é um computador altamente complexo, não-linear e paralelo. Ele tem a capacidade de organizar seus constituintes estruturais, conhecidos por neurônios, de forma a realizar certos

processamentos como reconhecimento de padrões, percepção e controle motor, muitos mais rapidamente que o mais rápido computador digital hoje existente (HAYKIN, 2000).

No momento do nascimento, um cérebro tem uma grande estrutura e habilidade de desenvolver suas próprias regras através do que usualmente denominamos “experiencia”. Na verdade, a experiência vai sendo acumulada com o tempo, sendo que o mais dramático desenvolvimento do cérebro humano acontece durante os dois primeiros anos de vida; mas o desenvolvimento continua para muito além desse estágio.

Figura 3 – Representação de uma rede neural



Fonte: Retirado de Laboratorio Mobilis, "Fundamentos De Redes Neurais", através do link: <http://www2.decom.ufop.br/imobilis/fundamentos-de-redes-neurais/>. Acessado em 3 de dezembro de 2022.

Para Haykin (2000) um neurônio em desenvolvimento é sinônimo de um cérebro plástico: a plasticidade permite que o sistema nervoso em desenvolvimento se adapte ao seu meio ambiente. Assim como a plasticidade parece ser essencial para o funcionamento dos neurônios como unidades de processamento de informação do cérebro humano, também ela o é com relação às redes neurais construídas por neurônios artificiais. Na sua forma mais geral, uma rede neural é uma máquina que é projetada para modelar a maneira como o cérebro realiza uma tarefa particular ou função de interesse; a rede é normalmente implementada utilizando-se componentes eletrônicos ou e simulada por programação em um computador digital. Para alcançarem bom desempenho a s redes neurais empregam uma interligação maciça de celular computacionais simples denominadas "neurônios"e “unidades de processamento”.

Uma rede neural é um processador maciçamente paralelamente distribuído constituído de unidades de processamento simples, que tem a propensão natural para armazenar conhecimento experimental e torná-lo disponível para o uso. Ela se assemelha ao cérebro em dois aspectos:

O conhecimento é adquirido pela rede a partir de seu ambiente através de um processo

de aprendizagem - Forças de conexão entre neurônios, conhecidas como pesos sinápticos, são utilizadas para armazenar o conhecimento adquirido.

O procedimento utilizado para realizar o processo de aprendizagem é chamado de algoritmo de aprendizagem, cuja função é modificar os pesos sinápticos da rede de uma forma ordenada para alcançar um objetivo de projeto desejado. A modificação dos pesos sinápticos e o método tradicional para o projeto de redes neurais, Esta abordagem é bastante próxima da teoria dos filtros adaptativos lineares, que já está bem estabelecida e foi aplicada com sucesso em diversas áreas (HAYKIN, 2000). Entretanto, é possível também para uma rede neural modificar sua própria topologia, o que é motivado pelo fato de os neurônios do cérebro humano poderem morrer e que novas conexões sinápticas possam crescer.

O uso de redes neurais oferece as seguintes propriedades úteis e capacidades: Não-linearidade: Um neurônio artificial pode ser linear ou não-linear. Uma rede neural, constituída por conexões de neurônios não-lineares e ela mesma não-linear. Além disso, a não-linearidade é uma propriedade muito importante, particularmente se o mecanismo físico responsável pela geração do sinal de entrada for inerentemente não-linear.

Mapeamento de entrada-saída: Um paradigma popular de aprendizagem chamado aprendizagem com um professor ou aprendizagem supervisionada envolve a modificação dos pesos sinápticos de uma rede neural pela aplicação de um conjuntos de amostras de treinamento rotuladas ou exemplos da tarefa. Cada exemplo consiste de um sinal de entrada único de uma resposta desejada correspondente. Apresenta-se para a rede um exemplo escolhido ao acaso do conjuntos, e os pesos sinápticos da rede são modificados para minimizar a diferença entre a resposta desejada e a resposta real da rede, produzida pelo sinal de entrada, de acordo com um critério estatístico apropriado, O treinamento da rede e repetido para muitos exemplos de conjuntos até que a rede alcance um estado estável onde não haja mais modificações significativas nos pesos sinápticos.

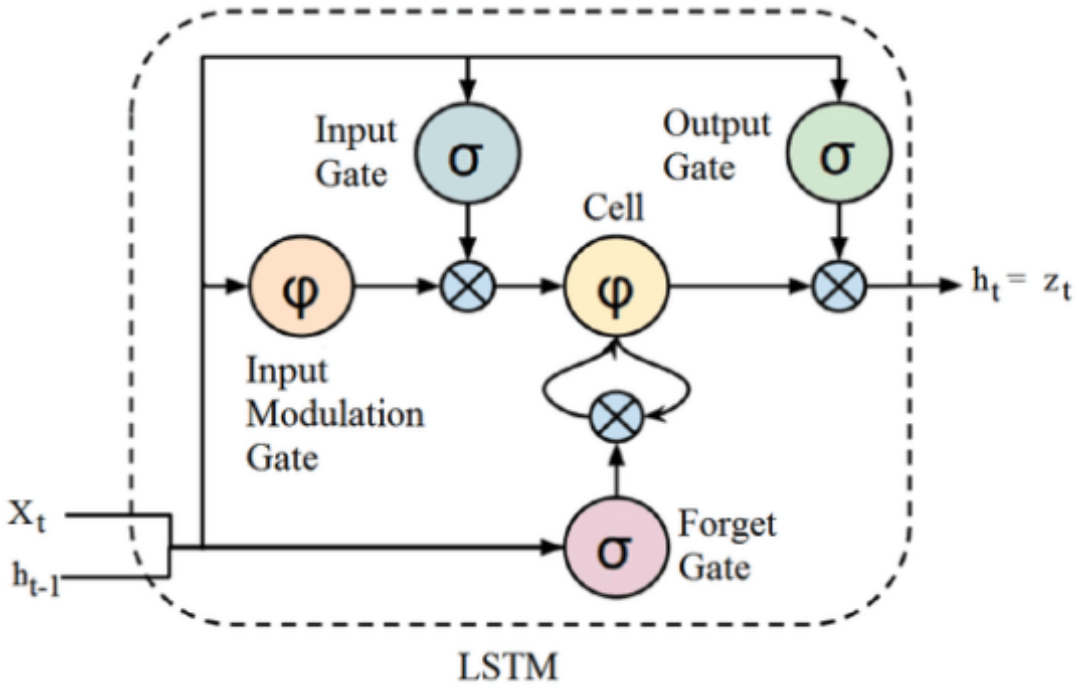
2.6.1 Long short-term memory

LSTM ou redes Long short-term memory são redes neurais recorrentes que podem aprender a dependência de ordem em problemas de previsão de sequência. A LSTM é usada em tradução de máquina, reconhecimento de fala, etc., devido a seus recursos favoráveis para solução de problemas complexos. LSTM pode armazenar informação prévia também. Isso contribui neste projeto, dado que preços anteriores de ações são fundamentais para a previsão de preços futuros.

A estrutura de uma LSTM consiste na célula, portão de input, portão de output, e portão de esquecimento.

A arquitetura da LSTM é um tipo de rede neural recorrente (RNN) utilizada na área de deep learning. Em contraste, para redes neurais feed-forward convencionais, LSTM tem

Figura 4 – Representação da arquitetura da LSTM



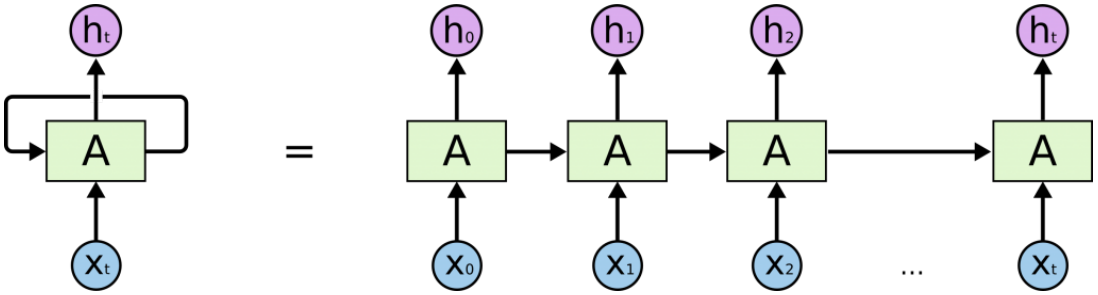
Fonte: Retirado de deeplearningbook, "Capítulo 51 – Arquitetura de Redes Neurais Long Short Term Memory (LSTM)", através do link: <https://www.deeplearningbook.com.br/arquitetura-de-redes-neurais-long-short-term-memory/>. Acessado em 20 de dezembro de 2022.

conexões de feedback. Uma típica unidade de LSTM compreende quatro componentes: a célula, portão de input, portão de output, e portão de esquecimento.

A célula armazena todos valores ao longo do tempo, o input e output dos dados da célula são controlados pelos portões. O de input regula todas novas informações que entraram na rede, o de esquecimento toma conta do conteúdo restante dentro da célula. A porta de saída cuida do limite ao qual a célula usa os dados alimentados e calcula a ativação de saída do algoritmo. O diagrama abaixo descreve a estrutura interna de uma rede LSTM. O portão de modulação de entrada faz parte do portão de entrada e é usado para segregação adicional de dados.

Uma unidade LSTM típica compreende quatro componentes: uma célula, uma porta de entrada, uma porta de saída e uma porta de esquecimento. A célula retém dados em períodos arbitrários e os três portões para controlar o fluxo de entrada e saída de informações. O modelo LSTM possui uma ampla variedade de aplicações; pode modelar linguagens ou gerar textos. O processamento de imagens é outro campo promissor para usar o LSTM, mas o modelo precisa de treinamento extensivo e refinamento para que isso aconteça. O LSTM pode prever notas musicais da mesma forma que a geração de texto, analisando as notas dadas como entrada. O LSTM também é um modelo muito confiável para desenvolver software de tradução de idiomas. O modelo codificador-decodificador LSTM é usado em tais softwares; ele converte as entradas

Figura 5 – Representação do loop da LSTM



Fonte: Retirado de deeplearningbook, "Capítulo 51 – Arquitetura de Redes Neurais Long Short Term Memory (LSTM)", através do link: <https://www.deeplearningbook.com.br/arquitetura-de-redes-neurais-long-short-term-memory/>. Acessado em 20 de dezembro de 2022.

em representações vetoriais e as saídas em sua versão traduzida.

Para Moghar e Hamiche (2020) o objetivo deste artigo é desenvolver um modelo para prever valores futuros do mercado de ações usando Redes Neurais Recorrentes (RNN) e, mais especificamente, o modelo de Memória de Longo-Curto Prazo (LSTM). O objetivo principal é determinar a precisão com que um algoritmo de aprendizado de máquina pode prever e até que ponto as épocas podem aprimorar nosso modelo. Para vários conjuntos de dados, observa-se que o treinamento com muito menos dados e muito mais épocas melhora nosso resultado de teste, ao mesmo tempo em que nos permite obter previsões e valores de previsão superiores. Um estudo futuro tentará identificar os conjuntos ideais em termos de comprimento de dados e épocas de treinamento que melhor se ajustam aos nossos ativos e otimizam nossa precisão de previsão.

2.7 Métodos de Ensemble

O objetivo dos métodos ensemble é combinar as previsões de vários estimadores de base construídos com um determinado algoritmo de aprendizado, a fim de melhorar a generalização/robustez em um único estimador. Scikit-learn (2022) Duas famílias de métodos ensemble são normalmente distinguidas:

Nos métodos de média, o princípio motriz é construir vários estimadores independentemente e, em seguida, calcular a média de suas previsões. Em média, o estimador combinado é geralmente melhor do que qualquer estimador de base única porque sua variância é reduzida.

Por outro lado, nos métodos boosting, os estimadores de base são construídos sequencialmente e tenta-se reduzir o viés do estimador combinado. A motivação é combinar vários modelos fracos para produzir um conjunto poderoso.

2.7.1 Voting Classifier

A ideia por trás do VotingClassifier é combinar conceitualmente diferentes classificadores de aprendizado de máquina e usar um voto majoritário ou as probabilidades médias previstas para prever os rótulos de classe. Tal classificador pode ser útil para um conjunto de modelos com desempenho igualmente bom, a fim de equilibrar suas fraquezas individuais.Scikit-learn (2022)

2.7.1.1 Voto da maioria

Na votação por maioria, o rótulo de classe previsto para uma amostra específica é o rótulo de classe que representa a maioria (moda) dos rótulos de classe previstos por cada classificador individual.

2.8 Métricas de desempenho

Para quantificar o desempenho dos modelos na previsão dos valores foram utilizadas métricas que avaliam os valores previstos em relação aos valores reais. Em seguida temos as métricas usadas no projeto e suas propriedades.

2.8.1 MAPE

O Erro Médio Percentual Absoluto (MAPE – Mean Absolute Percentage Error) obtém as diferenças percentuais entre todos os valores reais e previstos obtidos e faz uma média simples destes valores. Essa métrica é útil para ter um panorama geral do erro médio gerado pelo algoritmo de previsão escolhido. Segue sua equação:

$$MAPE = \frac{1}{N} \quad \sum_{i=1}^{N}((|P_{previsto,i} - P_{real,i})/P_{real,i}) * 100$$

2.8.2 MSE

O Erro quadrático médio (MSE - Mean Squared Error), uma das métricas mais utilizadas para calcular o desempenho de modelos. Ela é calculada a partir da soma da variância e dos quadrados das diferenças obtidas entre os valores reais e previstos. O MSE é uma medida da qualidade de um estimador. Como é derivado do quadrado da distância euclidiana, é sempre um valor positivo que diminui à medida que o erro se aproxima de zero. Segue sua equação:

$$MSE = \frac{1}{N} \quad \sum_{i=1}^{N}(previsto_i - real_i)^2$$

2.8.3 RMSE

Raiz quadrada do erro quadrático médio (RMSE - Root Mean Squared Error) representa a raiz quadrada do segundo momento amostral das diferenças entre os valores preditos e os valores observados ou a média quadrática dessas diferenças. Esses desvios são chamados de resíduos quando os cálculos são realizados sobre a amostra de dados que foi usada para estimativa e são chamados de erros quando calculados fora da amostra. Serve para agregar as magnitudes dos erros nas previsões para vários pontos de dados em uma única medida de poder preditivo. Segue sua equação:

$$RMSE = \sqrt{\frac{1}{N} \sum_{i=1}^{N} (previsto_i - real_i)^2}$$

2.8.4 Acurácia

A acurácia é o método mais simples e mais utilizado de avaliação, sendo chamada de taxa de acerto. Em nosso modelo, se avalia também a direção do movimento em relação ao dia anterior, logo definiremos duas distintas para previsão em relação ao dia anterior e a última previsão feita. Segue sua equação:

$$Acc = \frac{1}{N} \sum_{i=1}^{N} P_i$$

3 Trabalhos Correlatos

Drew Scatterday (2021) conduziu uma análise muito detalhada do algoritmo SVM. Ele previu as ações da Tesla usando um modelo SVM conhecido como Support Vector Regression com sci-kit-learn. Além disso, ele usou um LSTM usando Keras. A regressão linear foi empregada para obter uma linha de melhor ajuste em relação a duas variáveis. A rede neural recebeu os preços das ações dos 36 dias anteriores para, eventualmente, calcular o preço de fechamento do dia seguinte. Os resultados foram quase idênticos aos preços reais. O quadro de dados de teste forneceu as previsões para serem transformadas em valores entre 0 e 1. LSTM e SVM previram o preço para um ano inteiro e os resultados corresponderam perfeitamente.(SCATTERDAY, 2019)

Serafeim (2020) usou os preços de fechamento das ações da Tesla. Inc para testar o modelo LSTM. Os dados foram retirados do ano de 2015 até 2020 (Yahoo Finance). Uma rede neural LSTM de várias camadas deveria ser preparada, o que poderia prever os movimentos do preço das ações. O modelo possui 50 neurônios e quatro camadas ocultas, que atuariam como uma peneira para os dados de saída. Um neurônio foi designado para calcular os preços. O neurônio atribuído foi colocado na camada de saída. Devido ao bloqueio do Covid-19, os preços das ações mostraram movimentos anormais no final que o algoritmo não pôde prever, mas previu sistematicamente os movimentos de preços anteriores. Todos os saltos e quedas foram seguidos com precisão, e pode-se concluir com segurança que o modelo não estava longe de ser preciso.(LOUKAS, 2020)

Henrique et al. (2018) testaram o modelo SVM em três ações blue-chip e três small-caps das bolsas de valores do Brasil, América e China. O número total de ativos somou 18. Dez anos de dados foram extraídos de fontes como Yahoo Finance, Reuters e BMF Bovespa. Os dados foram divididos em 70 para treinamento e 30 para testes e comparações. Para calcular a precisão, RMSE e MAPE foram determinados a partir dos resultados. Os testes provaram que o SVM ou o SVR melhoraram os resultados quando usados com um kernel linear. Os resultados finais foram comparados com um modelo baseado em passeio aleatório. Este teste indicou que o modelo SVM melhorou sem atualizações periódicas e apresentou melhores resultados ao fornecer valores de previsão precisos. Isso prova que o SVM é muito superior a qualquer modelo de decisão aleatória .(HENRIQUE; SOBREIRO; KIMURA, 2018)

Gururaj et al. (2019) pegou o conjunto de dados do site Quandl (preços diários históricos). O experimento foi feito para definir se o SVM era preciso do que a Regressão Linear. O pacote R API de Quandl é altamente ativo e pode coletar preços de ações de empresas para qualquer intervalo de tempo em apenas uma única linha de código. Assim, uma etapa de mineração de dados é eliminada. As ações da Coca-Cola Company foram testadas de 2017 a

2018. Raiz do Erro Quadrado Médio (RMSE), Erro Absoluto Médio (MAE), Erro Percentual Absoluto Médio (MAPE), Erro Quadrado Médio (MSE) e Regressão de Correlação (R) e fator de correlação múltipla não linear foram usados como métricas de avaliação. Depois de realizar o experimento e calcular com precisão os resultados usando os modelos anteriores, concluiu-se que o SVM era mais preciso do que o RL. (V. SHRIYA V.R., 2019)

4 Metodologia

Neste capítulo são apresentados os principais passos para o desenvolvimento do trabalho, e introduzidas as principais ferramentas e bibliotecas utilizadas para a execução do trabalho.

4.1 Etapas

O trabalho foi iniciado com uma pesquisa bibliográfica para compreender a área de aprendizado de máquina e mercado financeiro. Foram estudados também os modelos modernos para previsão no mercado de ações. Durante esta fase foram escolhidas as técnicas, métricas e modelos implementados posteriormente com maior sucesso. Na sequência, foram pesquisadas bibliotecas, módulos e ferramentas que otimizam a implementação dos modelos, tendo sido escolhida, após tais pesquisas, a linguagem Python para programação devido ao grande número de pacotes e serviços com foco nas áreas de Ciência de Dados e AM, os métodos foram então implementados. Foram também colhidos os dados, raspados a fim de serem utilizados na análise exploratória. Finalmente, foram efetuados então o treinamento dos diferentes modelos e obtidas as medidas de acurácia, com suas respectivas visualizações para serem analisadas.

4.2 Ferramentas

A coleta será feita através da sessão de cotações históricas do site da B3. Por fim, para visualização se utilizará de bibliotecas como Plotly para construção de gráficos, e para implementar a Dashboard o framework Dash, e Spyder como IDE capaz de ser acessada e utilizada pela comunidade.

4.2.1 Python

Python é uma linguagem de programação que possui estruturas de dados eficientes de alto nível e uma abordagem simples, mas eficaz, para programação orientada a objeto, juntamente com sua natureza interpretada, o tornam uma linguagem ideal para scripts e desenvolvimento rápido de aplicativos em muitas áreas. Documentação pode ser encontrada em Python (2022).

4.2.2 Scikit-learn

Scikit-learn é uma biblioteca de aprendizado de máquina de código aberto que suporta aprendizado supervisionado e não supervisionado. Ele também fornece várias ferramentas para

ajuste de modelo, pré-processamento de dados, seleção de modelo, avaliação de modelo e muitos outros utilitários. Documentação pode ser encontrada em Scikit-Learn (2022).

4.2.3 Pandas

Pandas é um pacote Python que fornece estruturas de dados rápidas, flexíveis e expressivas projetadas para tornar o trabalho com dados "relacionais" ou "rotulados" fácil e intuitivo. Ele visa ser o bloco de construção fundamental de alto nível para fazer análises práticas de dados do mundo real em Python. Além disso, tem o objetivo mais amplo de se tornar a ferramenta de análise/manipulação de dados de código aberto mais poderosa e flexível disponível em qualquer idioma. Documentação pode ser encontrada em Pandas... (2022).

4.2.4 Dash

Dash é a estrutura original de baixo código para criar rapidamente aplicativos de dados em Python, R, Julia e F (experimental). Escrito sobre Plotly.js e React.js, o Dash é ideal para criar e implantar aplicativos de dados com interfaces de usuário personalizadas. É particularmente adequado para quem trabalha com dados. Documentação pode ser encontrada em Dash (2022).

4.2.4.1 Plotly

A biblioteca de gráficos Python da Plotly cria gráficos interativos com qualidade de publicação. Exemplos de como fazer gráficos de linha, gráficos de dispersão, gráficos de área, gráficos de barras, barras de erro, gráficos de caixa, histogramas, mapas de calor, subplots, eixos múltiplos, gráficos polares e gráficos de bolhas.

4.2.5 Keras

Keras é uma API de aprendizado profundo escrita em Python, executada sobre a plataforma de aprendizado de máquina TensorFlow. Foi desenvolvido com foco em permitir experimentação rápida. Ser capaz de ir da ideia ao resultado o mais rápido possível é a chave para fazer uma boa pesquisa. Documentação pode ser encontrada em Keras.io (2022).

4.2.6 Google Colab

O Colaboratory ou "Colab" é um produto do Google Research, área de pesquisas científicas do Google. O Colab permite que qualquer pessoa escreva e execute código Python arbitrário pelo navegador e é especialmente adequado para aprendizado de máquina, análise de dados e educação. O Colab é um serviço de notebooks hospedados do Jupyter que não requer nenhuma configuração para usar e oferece acesso sem custo financeiro a recursos de computação como GPUs.

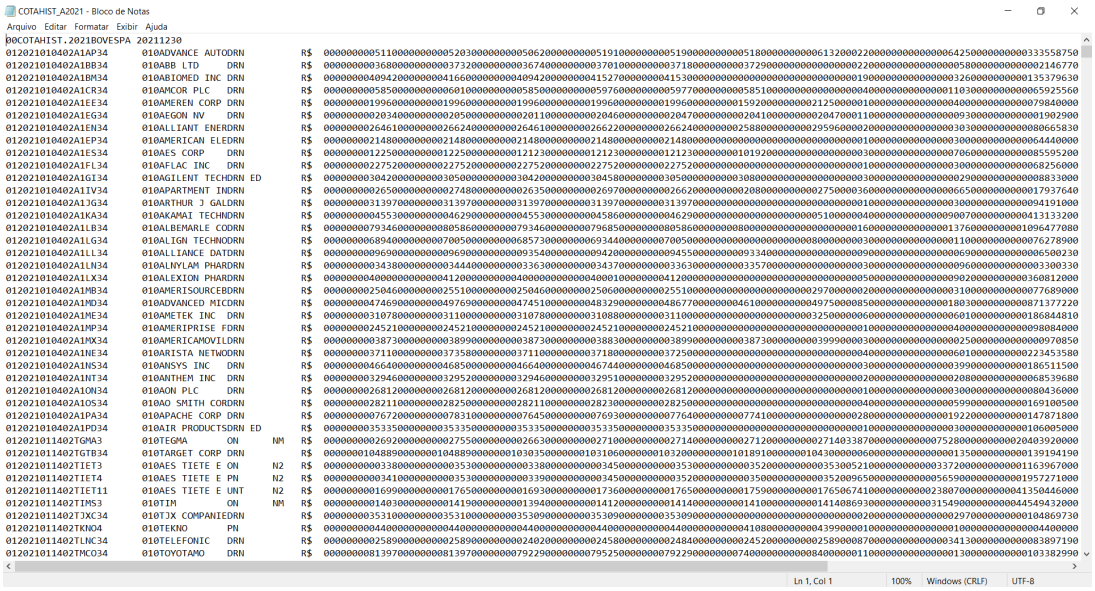
5 Experimentação e Análise dos Resultados

Neste capítulo são detalhados os passos do desenvolvimento do trabalho e seus experimentos, sendo apresentado o processo para a obtenção de dados, discutidas características do corpus e descritos os procedimentos de pré processamento, em seguida são apresentadas as métricas utilizadas para avaliação dos modelos, detalhados os modelos que foram implementados no trabalho, retratados os parâmetros do treinamento dos modelos e finalmente incluídos e discutidos os resultados dos teste.

5.1 Dados

A coleta foi feita através da sessão de cotações históricas do site da B3 no formato FWF (Fixed Width Formatted), cujo arquivo tem largura fixa, com dados de Janeiro de 2012 até Janeiro de 2022.

Figura 6 – Arquivo aberto no formato .TXT



Fonte: Elaborado pelo autor.

5.1.1 Pré-processamento dos Dados

Esta subseção discorre sobre as técnicas de pré-processamento dos dados. Dados após serem raspados.(Figura 7).

Figura 7 – Dataframe das ações da GOLL4

```
df_gol['MACD'] = macd
df_gol
```

Date	Open	High	Low	Close	Adj Close	Volume	MM	ME	MACD
2012-01-03 00:00:00-05:00	13.94	14.18	13.82	14.02	14.017532	857550	NaN	NaN	NaN
2012-01-04 00:00:00-05:00	14.02	14.08	13.74	13.78	13.777574	410050	NaN	NaN	NaN
2012-01-05 00:00:00-05:00	13.74	13.74	13.28	13.38	13.377645	460550	NaN	NaN	NaN
2012-01-06 00:00:00-05:00	13.44	13.50	12.96	13.00	12.997711	629350	NaN	NaN	NaN
2012-01-09 00:00:00-05:00	13.34	13.48	13.18	13.24	13.237669	453150	NaN	NaN	NaN
...	...	...	...	...	...	...	...	...	...
2020-12-23 00:00:00-05:00	9.41	9.60	9.38	9.41	9.410000	1081400	9.938095	9.797018	0.386308
2020-12-24 00:00:00-05:00	9.40	9.52	8.98	9.39	9.390000	621000	9.968571	9.715614	0.322205
2020-12-28 00:00:00-05:00	9.36	9.40	9.11	9.28	9.280000	1032700	9.987619	9.628491	0.259536
2020-12-29 00:00:00-05:00	9.50	9.53	9.34	9.42	9.420000	602400	10.022857	9.586793	0.218646
2020-12-30 00:00:00-05:00	9.49	9.88	9.48	9.65	9.650000	1149500	10.061905	9.599434	0.202466

2264 rows × 9 columns

Fonte: Elaborado pelo autor.

5.1.1.1 Normalização dos dados

Padronizando os dados removendo a média e dimensionando para a variação da unidade. A pontuação padrão de uma amostra x é calculada como:

$$z = \frac{x - u}{s}$$

A centralização e o dimensionamento acontecem independentemente em cada recurso, calculando as estatísticas relevantes nas amostras do conjunto de treinamento. A média e o desvio padrão são armazenados para serem usados em dados posteriores usando a transformação.

A padronização de um conjunto de dados é um requisito comum para muitos estimadores de aprendizado de máquina: eles podem se comportar mal se os recursos individuais não se parecerem mais ou menos com dados padrão normalmente distribuídos.

Muitos elementos usados na função objetivo de um algoritmo de aprendizado (como o kernel RBF) assumem que todos os recursos estão centrados em 0 e têm variância na mesma ordem. Se uma característica tem uma variância que é ordem de grandeza maior do que outras, ela pode dominar a função objetivo e tornar o estimador incapaz de aprender com outras características corretamente.

5.2 Treinamento dos Modelos

Para esse trabalho foram implementados três modelos, o primeiro sendo pelo algoritmo SVM, o segundo LSTM e o terceiro resultado do métodos ensemble, descritos nas seções 2.5, 2.6.1 e 2.7.

5.2.1 Dados de entrada

Para a SVM após normalizar os dados estão prontos para serem consumidos, e para rede LSTM, que recebe uma sequência de dados como entrada, foram ajustados matrizes de 3 dimensões sendo o número de amostras, o tamanho da sequência, e o número de variáveis de entrada.

5.2.2 Treino, validação e teste

A principio para avaliar os modelos, ambos foram treinados com o conjunto de dados raspado. Sendo separados em 3 conjuntos para a LSTM: 70 porcento para treino, 20 porcento para validação e 10 porcento para teste. Já no caso da SVM e 2 conjuntos: 70 porcento para treino e 30 porcento para teste. Calculamos o MSE no conjunto de validação, quanto maior, menor a probabilidade de o modelo generalizar corretamente a partir dos dados de treinamento., evitando superajustamento (overfitting), quando o modelo mostra-se adequado apenas para os dados de treino. Para a SVM com um total de 2729 amostras, foi treinado o modelo com o kernel da função de base radial e função Gamma ajustada em 0,1.

Foi construida a LSTM com 50 neurônios e 4 camadas ocultas. Por fim, atribuiremos 1 neurônio na camada de saída para prever o preço normalizado das ações. Usaremos a função de perda MSE e o otimizador Adam stochastic gradient descent (KINGMA, 2015). O treinamento com 1000 épocas foi o com o melhor desempenho para convergência do modelo. Em relação as funções de ativação, a que desempenhou melhor foi a função de tangente hiperbólica (tanh) em todas as saídas. Não foi encontrado melhora ao aumentar o espaço dimensional da saída da LSTM, por isso esse valor foi mantido como o mesmo número das variáveis de entrada, mas por ser uma rede bidirecional o valor é o dobro do variáveis de entrada. Também não houve melhoras significativas em adicionar mais camadas ao modelo.

5.3 Experimentos

A seguir estão as figuras e quadros com os resultados da performance dos modelos para os conjuntos citados anteriormente.

Gráficos obtidos na análise exploratória sobre os dados para entender como eles se comportam ao longo do tempo.

Figura 8 – Valor de fechamento - ABEV3

Fonte: Elaborado pelo autor.

Figura 9 – Valor de fechamento - ITUB4

Fonte: Elaborado pelo autor.

Com a figura a seguir, podemos observar visualmente as medias móveis ao conjecturar sobre os preços futuros.

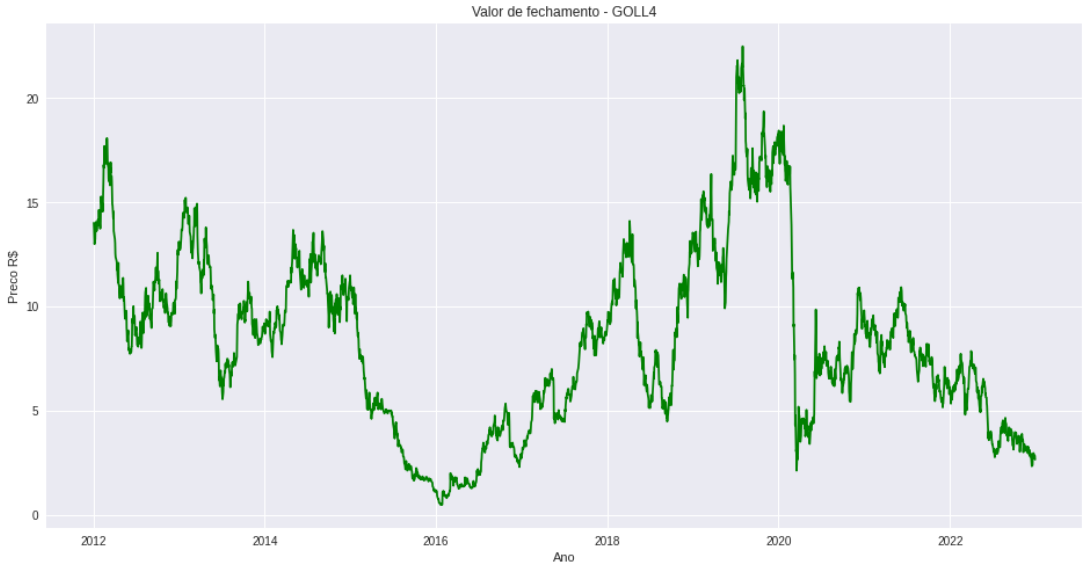
Pode-se observar que as previsões utilizando o modelo LSTM, obtiveram precisão e acurácia maior.

Após alguns testes, ficou claro que a convergência do modelo havia atingido seu mínimo ali, em torno de 800 a 1000 épocas.

Figura 10 – Valor de fechamento - PETR4

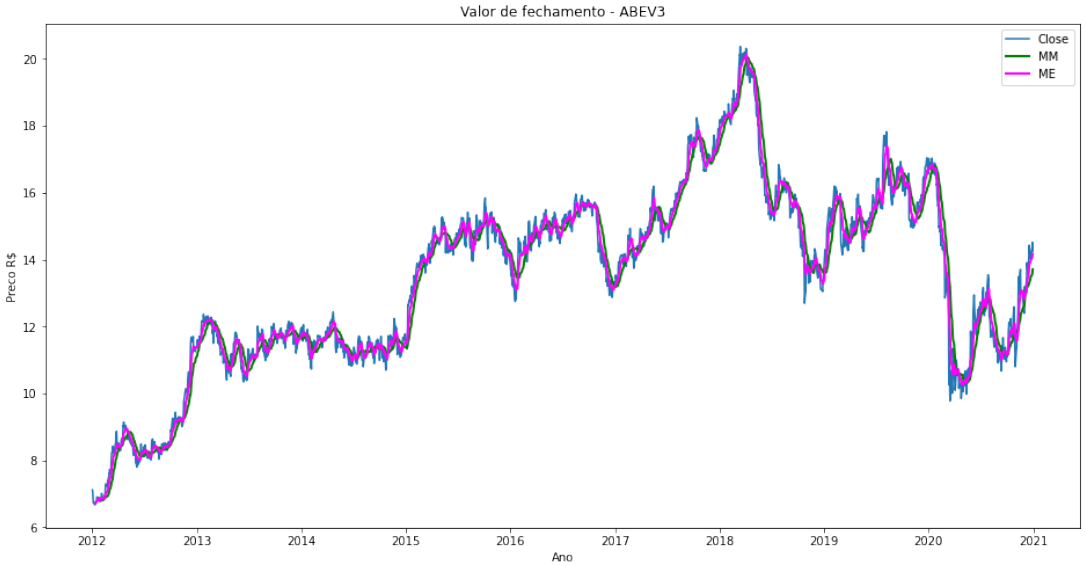
Fonte: Elaborado pelo autor.

Figura 11 – Valor de fechamento - GOLL4



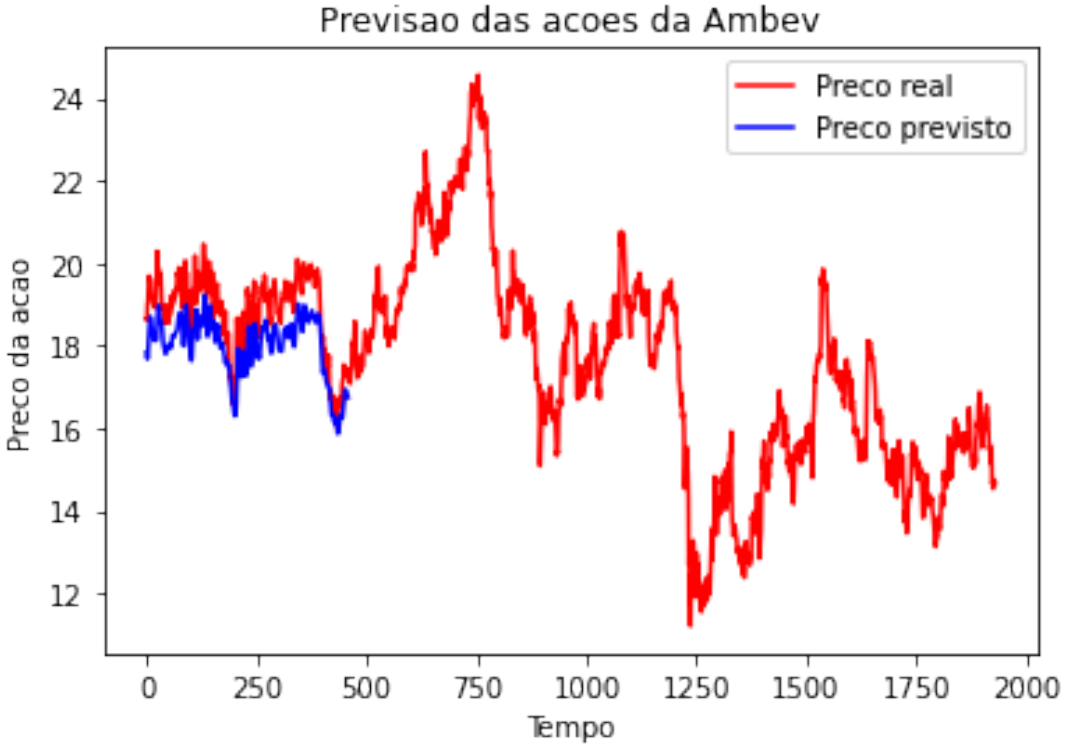
Fonte: Elaborado pelo autor.

Figura 12 – Média móvel simples e exponencial comparados com o valor real fechado



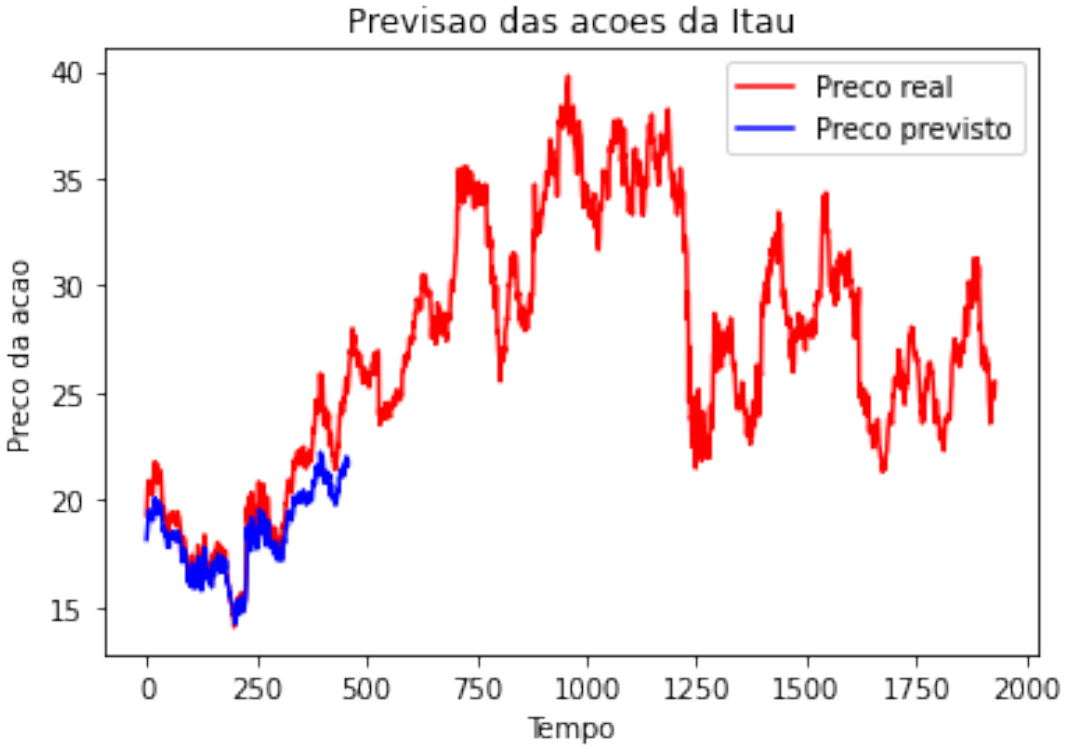
Fonte: Elaborado pelo autor.

Figura 13 – Previsão de ações da empresa Ambev pelo modelo LSTM



Fonte: Elaborado pelo autor.

Figura 14 – Previsão de ações da empresa Itaú pelo modelo LSTM

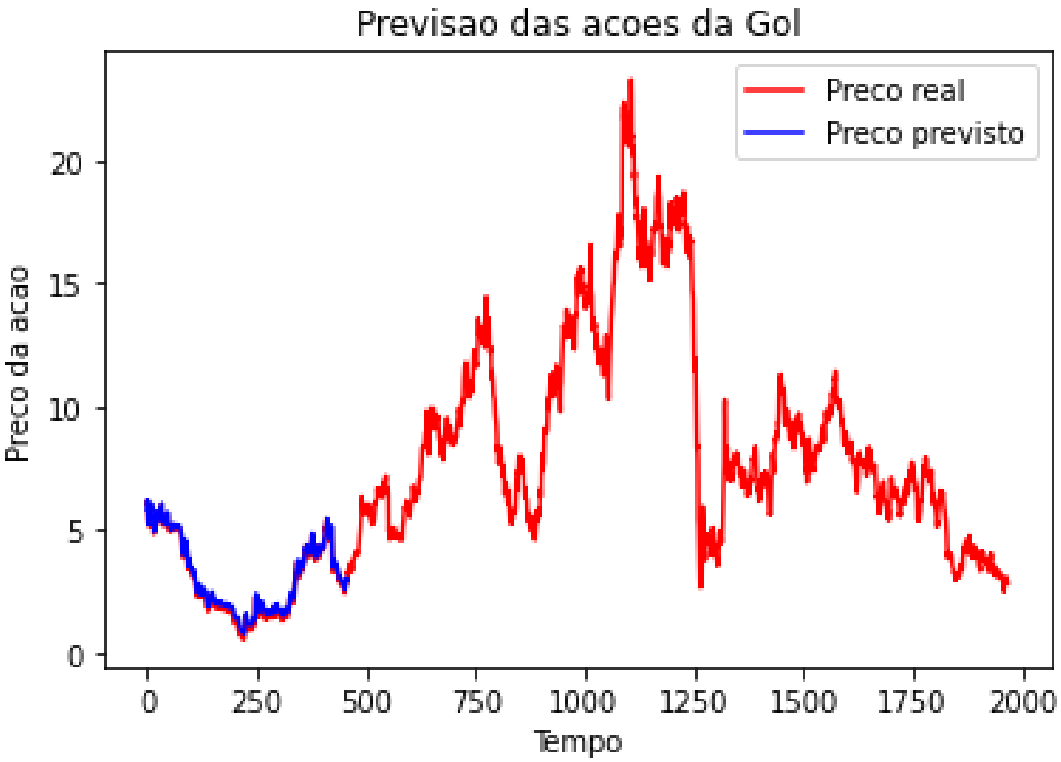


Fonte: Elaborado pelo autor.

Figura 15 – Previsão de ações da empresa Petrobras pelo modelo LSTM

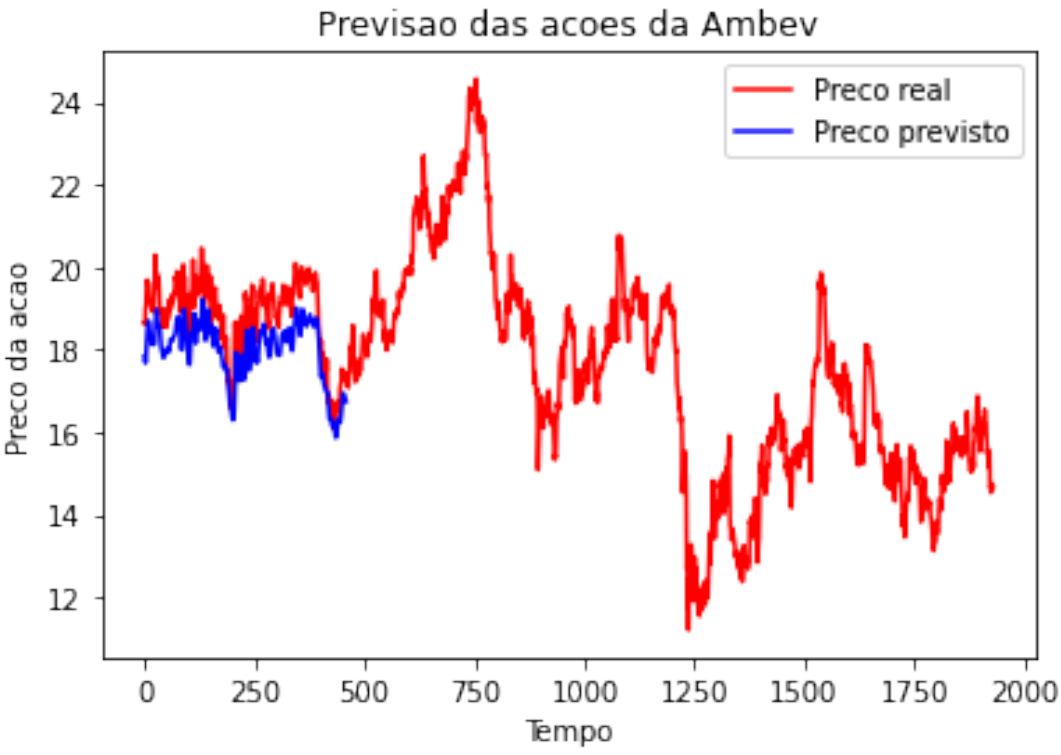
Fonte: Elaborado pelo autor.

Figura 16 – Previsão de ações da empresa Gol pelo modelo LSTM



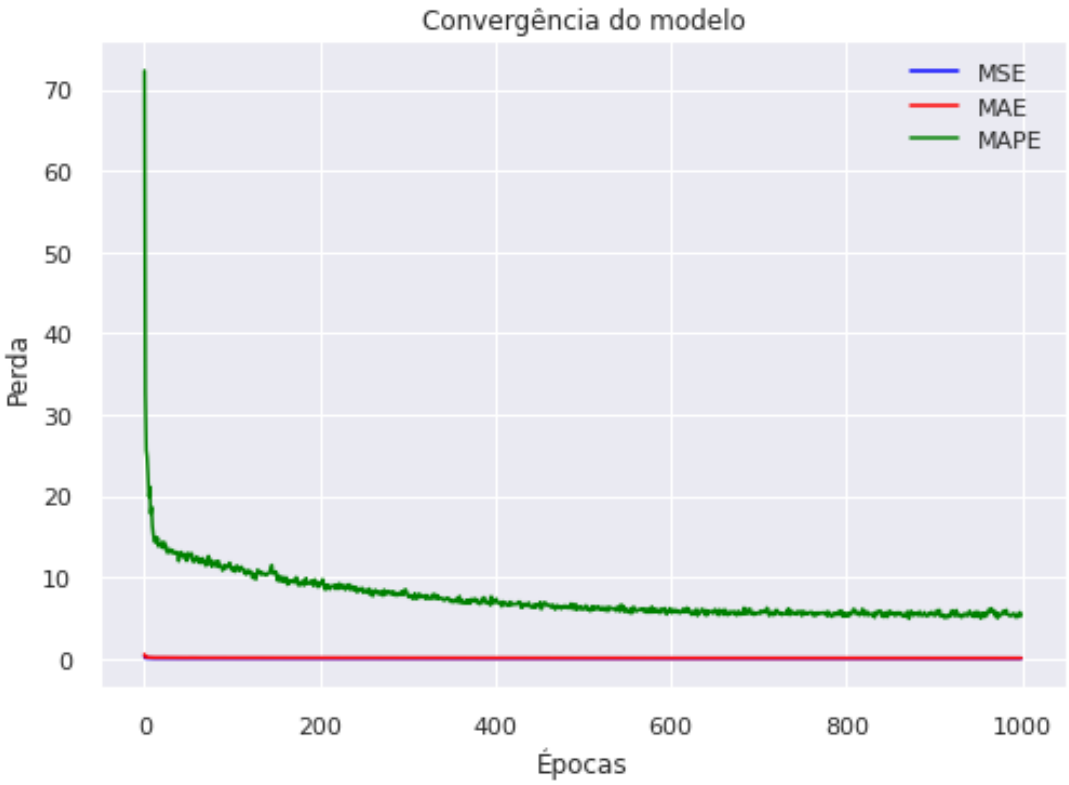
Fonte: Elaborado pelo autor.

Figura 17 – Previsão de ações da empresa Ambev pelo modelo LSTM



Fonte: Elaborado pelo autor.

Figura 18 – Convergência do modelo em épocas



Fonte: Elaborado pelo autor.

Figura 19 – Métricas do modelo SVR

GOLL4	MAE	MSE	MAPE
SVR	1.0924883515869446	2.364033384188477	0.1482539037763215

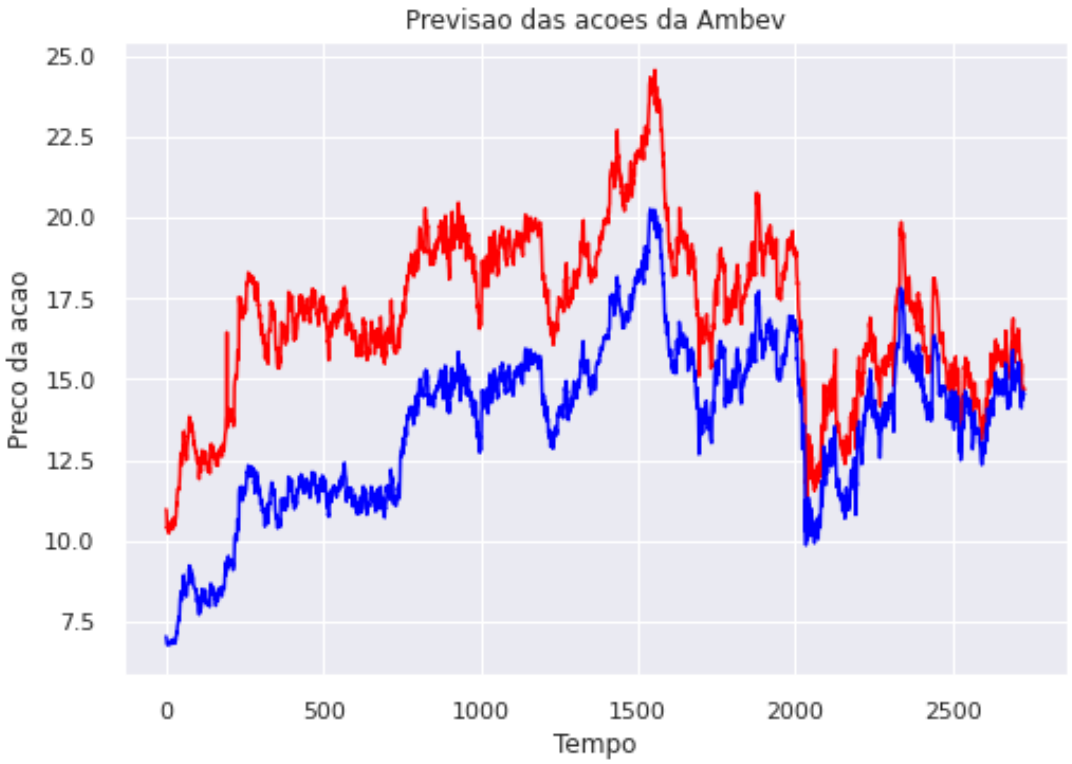
ITUB4	MAE	MSE	MAPE
SVR	1.1136217017846606	2.3958081931958732	0.055749299084890416

PETR4	MAE	MSE	MAPE
SVR	0.821091256404947	1.5298671257972725	0.08963196213248963

ABEV3	MAE	MSE	MAPE
SVR	0.6110377030869935	0.6350332684291058	0.04554646175688439

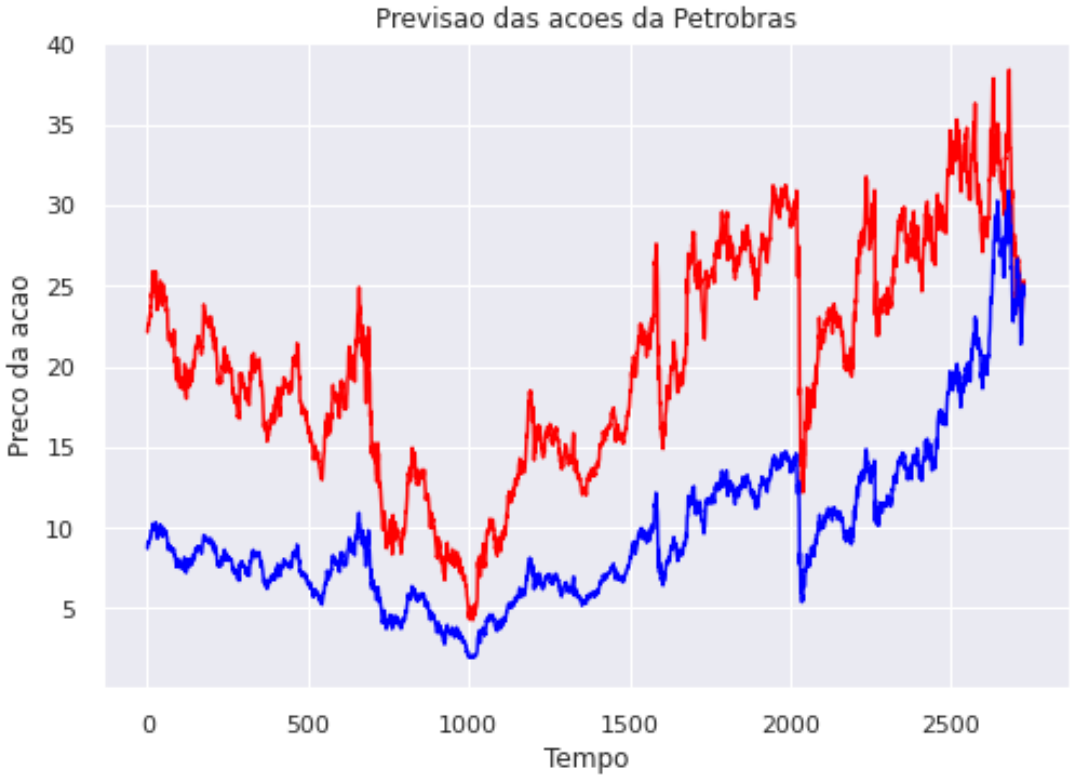
Fonte: Elaborado pelo autor.

Figura 20 – Previsão de ações da empresa Ambev pelo modelo SVR



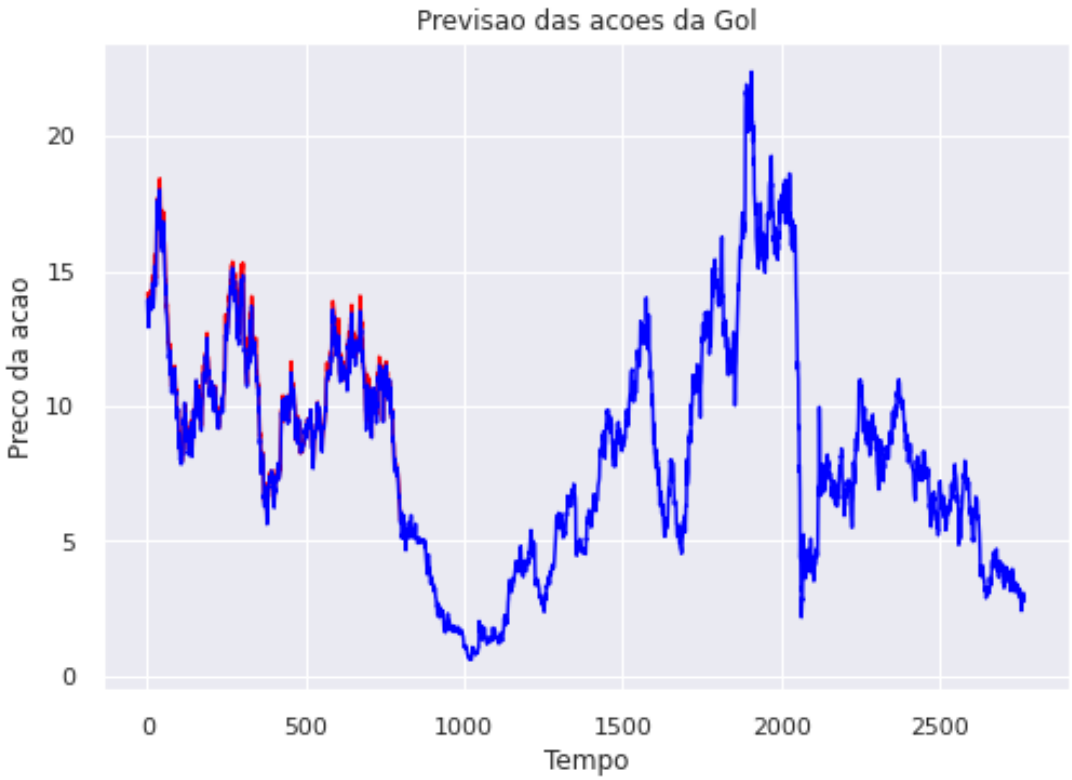
Fonte: Elaborado pelo autor.

Figura 21 – Previsão de ações da empresa Petrobras pelo modelo SVR



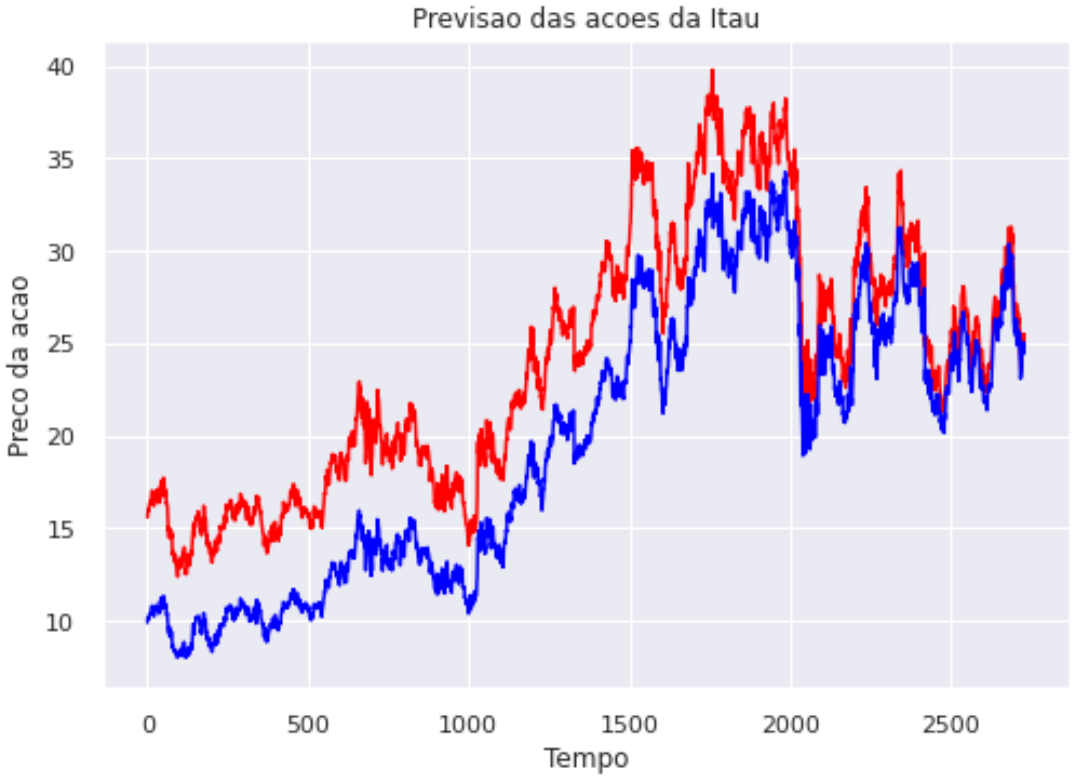
Fonte: Elaborado pelo autor.

Figura 22 – Previsão de ações da empresa Gol pelo modelo SVR



Fonte: Elaborado pelo autor.

Figura 23 – Previsão de ações da empresa Itaú pelo modelo SVR



Fonte: Elaborado pelo autor.

6 Considerações Finais

O presente trabalho propõe análises com o uso de redes neurais recorrentes em combinação com máquinas de vetor de suporte para regressão não-linear, a fim de prever valores futuros e indicar tendências. A priori foi necessário realizar um estudo sobre o mercado de ações e seus ativos, além de pesquisas relacionadas à previsão de sistemas dinâmicos e temporais. A seguir, um aprofundamento na teoria dos modelos que seriam utilizados para a análise e também nas bibliotecas e frameworks usados para desenvolvê-la.

Foram propostos dois algoritmos, de naturezas diferentes, para encontrar padrões nos dados de negociações coletadas, realizar previsões e obter insights para melhorar a tomada de decisão de potenciais investidores.

Os resultados apresentados foram satisfatórios, dada a dificuldade da tarefa, e pode-se, entender a análise técnica como uma abordagem para investimentos. Os modelos nos ajudam a entender padrões em grandes volumes de dados e auxiliam na tomada de decisão. Neste projeto foram analisados dados de fechamento diário, aumentando a granularidade dos dados poderíamos obter preços e negociações na ordem de minutos ou segundos, observando tendências ao longo do dia.

Infelizmente, para ações com menor volume e instáveis, os valores futuros se mostram menos determinísticos através de valores históricos. Portanto, se torna menos confiável e preciso modelos que se baseiam puramente em dados numéricos.

Tanto para estes casos, quanto para o projeto, tem se mostrado mais promissor redes neurais mais complexas que combinam análises técnica e fundamentalista. Como incorporar dados históricos a análise de sentimento em notícias e mídias sociais em relação ao mercado de ações em geral, bem como a uma determinada ação de interesse.

Referências

ABU-MOSTAFA, A. Introduction to financial forecasting. applied intelligence. An Introduction to Cluster Computing, n. 1, p. 205–213, 1996.

BHAVSAR, P.; SAFRO, I.; BOUAYNAYA, N.; POLIKAR, R.; DERA, D. Chapter 12 - machine learning in transportation data analytics. Elsevier, p. 283–307, 2017. Disponível em: https://www.sciencedirect.com/science/article/pii/B9780128097151000122.

BOSER, B.; GUYON, I.; VAPNIK, V. A training algorithm for optimal margin classifier. *Proceedings of the Fifth Annual ACM Workshop on Computational Learning Theory*, v. 5, 08 1996.

DASH. *Dash Documentation*. 2022. Disponível em: https://dash.plotly.com/. Acesso em: 12 dez. 2022.

DEBASTIANI, C. A. *Análise Técnica de Ações: Identificando Oportunidades de Compra e Venda*. [S.l.]: Novatec, 2008. v. 1.

ELDER, A. *Aprenda a operar no mercado de ações: um guia completo para trading*. [S.l.]: Alta Books, 1993. v. 1.

HAYKIN, S. *Redes Neurais: Princípios e Prática*. [S.l.]: Bookman, 2000. v. 2.

HENRIQUE, B. M.; SOBREIRO, V. A.; KIMURA, H. Stock price prediction using support vector regression on daily and up to the minute prices. *The Journal of Finance and Data Science*, v. 4, n. 3, p. 183–201, 2018. Disponível em: https://www.sciencedirect.com/science/article/pii/S2405918818300060.

KERAS.IO. *Keras*. 2022. Disponível em: https://keras.io/about/. Acesso em: 12 dez. 2022.

KINGMA, J. L. B. D. P. Adam: A method for stochastic optimization. 2015.

LEMOS, F. Análise técnica dos mercados financeiros: um guia completo e definitivo dos métodos de negociação de ativos. *São Paulo: Saraiva Educação*, 2015.

LOUKAS, S. *Time-series forecasting: Predicting stock prices using an LSTM model*. 2020. Disponível em: https://towardsdatascience.com/lstm-time-series-forecasting-predicting-stock-prices-using-an-lstm-model-6223e9644a2f. Acesso em: 26 nov. 2022.

MALTA, C. Variáveis da análise fundamentalista e dinâmica e o retorno acionário de empresas brasileiras entre 2007 e 2014. *Rege - Revista de Gestão*, 2016.

MOGHAR, A.; HAMICHE, M. Stock market prediction using lstm recurrent neural network. *Procedia Computer Science*, v. 170, p. 1168–1173, 2020. Disponível em: https://www.sciencedirect.com/science/article/pii/S1877050920304865.

PANDAS Documentation. 2022. Disponível em: https://pandas.pydata.org/docs/index.html. Acesso em: 12 dez. 2022.

PYCODEMATES. *The RBF kernel in SVM: A Complete Guide*. 2022. Disponível em: https://www.pycodemates.com/2022/10/the-rbf-kernel-in-svm-complete-guide.html. Acesso em: 17 dez. 2022.

PYTHON. *Python Documentation*. 2022. Disponível em: https://docs.python.org/3. Acesso em: 12 dez. 2022.

RUSSEL, S.; NORVIG, P. *Artificial Intelligence: A Modern Approach*. [S.l.]: Pearson, 2002. v. 2.

SCATTERDAY, D. *Walking through support vector regression and LSTMs with stock price prediction*. 2019. Disponível em: https://towardsdatascience.com/walking-through-support-vector-regression-and-lstms-with-stock-price-prediction-45e11b620650. Acesso em: 22 nov. 2022.

SCIKIT-LEARN. *Ensemble Methods*. 2022. Disponível em: https://scikit-learn.org/stable/modules/ensemble.html#. Acesso em: 9 dez. 2022.

SCIKIT-LEARN. 2022. Disponível em: https://scikit-learn.org/stable/index.html. Acesso em: 12 dez. 2022.

SCIKIT-LEARN. *Voting Classifier*. 2022. Disponível em: https://scikit-learn.org/stable/modules/ensemble.html#voting-regressor. Acesso em: 9 dez. 2022.

TAN QUEK, S. Biological brain-inspired genetic complementary learning for stock market and bank failure prediction. *Computational Intelligence*, p. 236–261, 2007.

TEWELES, E. S. B. R. J. *The Stock Market*. [S.l.]: John Wiley Sons, 1992. v. 7.

V. SHRIYA V.R., A. K. G. *Stock market prediction using linear regression and support vector machines*. 2019. Disponível em: http://www.ripublication.com. Acesso em: 22 nov. 2022.

VIANA, J. S. Mercado financeiro brasileiro: uma análise da percepção dos egressos do curso de secretariado executivo da universidade federal do ceará acerca de uma atuação nesse campo profissional. *Trabalho de Conclusão de Curso (Graduação em Secretariado Executivo) – Faculdade de Economia, Administração, Atuária e Contabilidade*, Universidade Federal do Ceará, Fortaleza, 2022.