

AVALIAÇÃO DE DESEMPENHO DE REDES NEURAS RECORRENTES DO TIPO LSTM PARA DECISÃO DE INVESTIMENTO DE CURTO PRAZO NO MERCADO DE AÇÕES

PERFORMANCE EVALUATION OF RECURRENT NEURAL NETWORKS LSTM FOR SHORT-TERM INVESTMENT DECISION IN STOCK MARKET

Alexandre Pereira da Silva*

Silas Santiago Lopes Pereira[†]

Mário Wedney de Lima Moreira[‡]

RESUMO

Um desafio significativo para investir no mercado de ações é tomar decisões complexas e avaliar diferentes cenários e tendências nos relatórios financeiros. Uma alternativa capaz de atender a essas necessidades é a utilização de algoritmos de negociação. Essa estratégia pode envolver o uso de inteligência artificial para tomar decisões eficazes. Assim, este trabalho propõe o desenvolvimento de uma solução de apoio à tomada de decisão na compra e venda de ativos financeiros a partir da aplicação de técnicas de aprendizado de máquina. Para tanto, avaliou-se o desempenho de uma arquitetura baseada em rede neural recorrente, do tipo *long short-term memory* (LSTM), quanto à sua eficiência na previsão de curto prazo. Os modelos gerados foram avaliados com base nas métricas do erro quadrático médio e do erro percentual absoluto médio. Os resultados mostram que o algoritmo LSTM foi capaz de prever os preços das ações com um baixo erro, aproximando-se do valor real.

Palavras-chave: Mercado de ações. Inteligência artificial. Redes neurais artificiais. Long short-term memory.

ABSTRACT

A significant challenge to invest in the stock market is to make complex decisions and evaluate different scenarios and trends in financial reports. An alternative capable of attending to these needs is the use of algorithms for trading. This strategy may involve the use of artificial intelligence

*Alexandre Pereira é graduando em ciência da computação no Instituto Federal do Ceará (IFCE) e pesquisador bolsista pelo Programa Institucional de Bolsas de Iniciação Científica (PIBIC). Endereço eletrônico: alexandre.bcc1@gmail.com.

[†]Silas Santiago é professor do IFCE, mestre em ciências da computação na universidade estadual do ceará (UECE) e bacharel em ciência da computação pela UECE. Endereço eletrônico: silas@lar.ifce.edu.br.

[‡]Mário Wedney é professor do IFCE, doutor em engenharia informática pela Universidade da Beira Interior (UBI) em Portugal, mestre em ensino de ciências e matemática na Universidade Federal do Ceará (UFC) e licenciado em matemática pela UFC. Endereço eletrônico: mario.wedney@ifce.edu.br.

approaches to make effective decisions. Hence, this paper proposes the development of a solution to support decision-making in the purchase and sale of stocks from the application of machine learning techniques. For this purpose, the performance of a recurrent neural network architecture, also known as long short-term memory (LSTM), was evaluated concerning its efficiency in the short-term prediction. The generated models were evaluated based on the root-mean-square error and mean absolute percentage error metrics. Results show the LSTM algorithm was capable of predicting stock prices with a low error approximately.

Keywords: Stock Market. Artificial intelligence. Artificial neural networks. Long short-term memory.

1 INTRODUÇÃO

Mercados financeiros se caracterizam por sua alta volatilidade dado o alto nível de incertezas e riscos associados (DEORUKHKAR et al., 2019). Cotidianamente, bolsas de valores iniciam suas negociações eufóricas com as constantes variações de preços, o que permite ao investidor grandes oportunidades. Estima-se que diariamente sejam transacionados mais de US\$ 3 trilhões no mercado de *Forex*. Com um mercado tão inconstante e perigoso, é uma difícil tarefa para um investidor profissional e mais ainda para um leigo, acompanhar todas as variações de preço em tempo integral e a partir de uma análise determinar em tempo hábil a melhor estratégia de investimento para cada ocasião. Desta forma, o investimento no mercado de ações traz consigo a necessidade de tomar decisões complexas, avaliar diferentes cenários e tendências nas informações contidas em cotações nas bolsas de valores, o que requer além da *expertise* do especialista, um rigoroso planejamento e gerenciamento de riscos sobre o capital investido.

Dessa maneira, é muito recorrente no mundo dos investimentos a distinção de quatro perfis de estratégias a se considerar no mercado de ações, a saber, *scalper*, *daytrade*, *swing trading* e longo prazo. O *scalper* é caracterizado por negociações de curtíssimo prazo, que duram em média entre um e quinze minutos. Nesta forma de investimento é de crucial importância a precisão do investidor e velocidade no processamento de dados relacionados ao contexto financeiro que a empresa em negociação se encontra. *Daytrade* por sua vez é reconhecido por negociações que duram até um dia. Assim como o *scalper*, o *daytrade* é um perfil agressivo, devido a grande instabilidade suscetível que a cotação de uma empresa é sujeita. Dessa forma, tanto o *scalper* quanto o *daytrade* são utilizados por investidores de perfil agressivo que se expõem a mais riscos em seus investimentos. Em contrapartida a estes perfis de estratégias de investimentos, o *swing trade* e o longo prazo são característicos de um investimento conservador. Isto se dá pelo fato de que tanto o *swing trading*, que lida com investimentos entre três dias a um mês, quanto o longo prazo, que trata de investimentos superiores a um mês, possuem em essência menor risco ao capital.

Dada a caracterização individual de um perfil para o investidor, surge outra questão a

se definir como parâmetro para análise de investimento, *i.e.*, a métrica de interpolação usada para previsão. Atualmente, existem três formas predominantes de se fazer previsões sobre cotações de preços futuros dos ativos financeiros, que são a análise fundamentalista, a análise gráfica e a análise técnica. Conforme Eapen, Bein e Verma (2019, p. 264) e Ahmad Bagheri, Hamed Mohammadi e Mohsen Akbari (2014, p.6235), a análise fundamentalista se baseia em dados econômicos como fator determinante para sua previsão, logo o conjunto de dados relacionado a esta análise é intimamente ligado aos balanços patrimoniais e receitas, dentre outras informações de cunho contábil (EAPEN; BEIN; VERMA, 2019; BAGHERI; PEYHANI; AKBARI, 2014). Nessa forma de análise o preço do ativo, que é o produto negociado, pouco é considerado e não se utiliza de recursos gráficos para uma tomada de decisão. Em contrapartida à análise fundamentalista, a análise gráfica busca interpretar o ativo a partir de padrões e formas geométricas que, de forma empírica, são identificados nos gráficos de cotação de preços dos ativos financeiros. Segundo Cervelló-Royo, Guijarro e Michniuk (2015, p. 5963) esse método se baseia nos padrões comportamentais que os investidores externam em seus investimentos, o que acaba por serem representados no gráfico de cotação por meios de símbolos (CERVELLÓ-ROYO; GUIJARRO; MICHNIUK, 2015). Seguindo os conceitos da análise fundamentalista e gráfica, a análise técnica visa mesclar os fundamentos de ambas as técnicas. Esta admite que o comportamento dos preços tem origem estocástica. Logo, usa-se como ferramenta o aspecto probabilístico na forma de interpretar as informações. Portanto, o uso de séries temporais da estatística, bem como cálculo financeiro, têm grande importância na previsão do comportamento dos ativos. Liu, Liao e Ding (2018, p. 2787) enfatizam que o uso da probabilidade como tomada de decisão na análise técnica o coloca em outro patamar, posto que deixa de se sujeitar ao empirismo dos investidores e seus fatores externos como emoções, e baseia-se exclusivamente em dados estatísticos (LIU; LIAO; DING, 2018).

Dado o exposto, grande parte dos novos investidores sentem dificuldades em compreender a dinâmica do mercado financeiro de ações. Isto gera como consequência um distanciamento desta modalidade de investimento, privando-os de rentabilizarem o capital próprio. Diante dessa perspectiva, existe a necessidade de buscar alternativas que viabilizem a aplicação de recursos capitais que, além de fornecerem consistência, também disponham de um baixo risco e um grande potencial de lucro no contexto de finanças. Dessa forma, técnicas de *machine learning* surgem como candidatas a suprir essa necessidade, principalmente quando se trata de Redes Neurais Artificiais (RNAs) e aprendizagem profunda, como redes neurais recorrentes.

Baseado nessa premissa, Hatcher e Yu (2018, p. 24412) sugerem que a atratividade da aplicação de RNAs frente as técnicas convencionais é justificada a partir da capacidade de modelagem de sistemas mais complexos envolvendo um grande número de variáveis (HATCHER; YU, 2018). RNAs são aptas a resolver problemas de características gerais, como aproximação, classificação, categorização e previsão. Em previsão de séries temporais, a qual é uma tarefa que envolve a construção de modelos a partir de observações em um dado período, a aplicação de modelos clássicos requer a compreensão profunda do problema estudado. Na modelagem com RNAs, uma vez que estas se caracterizam por uma modelagem não linear e não paramétrica,

reduz-se a necessidade do entendimento do processo.

Em paralelo à aplicação de técnicas de ML no ambiente financeiro, também existem diversos programas paramétricos, denominados *expert advisors (EAs)*, que automatizam estratégias baseada em experiências empíricas e tomam posições de compra e venda a partir de regras previamente estabelecidas pelo desenvolvedor. Uma grande deficiência em EAs é a incapacidade de detectarem as mudanças que ocorrem no panorama comercial, o que compromete sua eficácia a longo prazo. Dada a capacidade de adaptação e previsão existente nas RNAs, espera-se um desempenho superior com relação aos EAs e uma consistência nos mais diversificados panoramas comerciais.

Dessa forma, esta pesquisa tem o objetivo de avaliar, aplicar e analisar técnicas de ML com foco em redes neurais long short-term memory (LSTM) para se criar um modelo de predição que sirva de apoio à tomada de decisões no que concerne a compra e venda de ativos financeiros. A principal matéria prima para construção dos modelos preditivos será o contexto histórico comercial, que é composto de informações sobre aberturas, fechamentos, máximas e mínimas dos ativos financeiros investigados. O objetivo deste estudo é avaliar um modelo de rede neural recorrente do tipo LSTM para antecipar decisões estratégicas relacionadas à compra e venda de ativos financeiros a partir do contexto histórico. Assim, pretende-se contribuir para a geração de oportunidades de negociações buscando o baixo risco e alto potencial de lucro (FISCHER; KRAUSS, 2018).

Ao decorrer do trabalho de desenvolvimento e pesquisa desta obra, obteve-se o seguinte artigo publicado em conferência internacional:

- “*Performance Evaluation of Recurrent Neural Networks for Short-Term Investment Decision in Stock Market*”. Artigo publicado na *International Conference on Advanced Communication and Networking (ACN 2019)*. O artigo apresenta uma arquitetura de previsão utilizando ML para prever o preço dos ativos financeiros no mercado de ações.

O presente artigo está organizado da seguinte forma. Os trabalhos relacionados são revisados na Seção 2. As Seções 3 e 4 detalham os conceitos centrais do mercado financeiro e das redes neurais LSTM. Em seguida, a Seção 5 descreve a metodologia utilizada para desenvolver a solução proposta. Na seção 6, os resultados obtidos são discutidos. Finalmente, na Seção 7, são apresentadas as conclusões e sugestões para trabalhos futuros.

2 TRABALHOS RELACIONADOS

Nesta seção é apresentada uma breve descrição dos estudos relacionados encontrados no atual estado da arte. O uso de ML é apresentado como ferramenta de previsão dos preços das ações e sua aplicação durante os últimos anos.

Em Deorukhkar *et al.* (2019), os autores apresentam uma abordagem do tipo *ensemble learning* para predição do preço de fechamento do próximo dia (DEORUKHKAR *et al.*, 2019). Este modelo é baseado em uma combinação de três diferentes modelos, a saber, Uma rede neural

recorrente do tipo LSTM, um modelo ARIMA (*Autoregressive Integrated Moving Average*) e um modelo baseado em análise de sentimentos. O intuito da combinação das técnicas descritas está em utilizar as intuições históricas dos modelos ARIMA e LSTM, além de agregar o estado corrente do mercado a partir da aplicação de análise de sentimentos. A saída de cada um destes modelos preditivos é combinada em uma rede neural *feed forward* para a predição do preço de fechamento do papel da ação. Os autores argumentam que a abordagem proposta traz dinamicidade ao considerar aspectos técnicos e fundamentais da ação.

Thiele e Adami (2016, p. 115) propõem um modelo preditivo para previsão de tendência das séries temporais financeiras que é baseado na combinação de três RNAs (THIELE; ADAMI, 2016). Para um dado ativo financeiro, uma rede receberá como entrada as variáveis dependentes deste ativo, que no estudo de caso consiste de abertura, fechamento, máxima e mínima do ativo. Outra rede usará as variáveis independentes que consistem em indicadores de análise técnica que são oriundos do tratamento das variáveis dependentes. A última rede, denominada RNA de fusão, é utilizada para combinar a saída dos dois modelos anteriores. Os autores argumentam que esta estratégia tem o objetivo de melhorar a previsão das duas primeiras redes treinadas individualmente, agregando conhecimento. Os experimentos consideraram três ações do mercado brasileiro (ITUB4, PETR4 e VALE5) para a previsão de tendência de alta ou de baixa. Neste processo, a arquitetura construída propôs somente a previsão da tendência do próximo dia útil, não explorando a interpolação desse intervalo. Os autores utilizaram os resultados obtidos pelo treinamento da rede e simularam com uma base de teste possíveis resultados para aplicação do modelo. Foi obtida uma média de acerto em torno de 57% considerando os três ativos analisados. Com relação a simulação de retorno financeiro, um saldo de 13% foi obtido, mesmo no cenário de desvalorização corrente no período de teste. Os autores concluem que o sistema preditivo tem a capacidade de gerar maiores lucros em momentos favoráveis no mercado de ações.

Em Arvidsson e Ånhed (2017), os autores fazem uma comparação da capacidade de previsão de séries temporais financeiras a partir de redes neurais recorrentes LSTM e redes neurais recorrentes tradicionais (ARVIDSSON; ÅNHED, 2017). Este estudo investiga a capacidade de redes neurais recorrentes LSTM na modelagem do comportamento do mercado financeiro e sua acurácia quando comparada para modelos tradicionais. Para execução dos experimentos foram utilizados dados históricos do ativo Euro - Dólar (EUR/USD) em um tempo gráfico de 10 minutos. Os dados de testes são compostos por preço de abertura, máxima, mínima e fechamento a partir do dia 9 de fevereiro de 2017 à 23 de fevereiro do mesmo ano. A arquitetura de modelagem para previsão usando a rede LSTM consistia de uma camada de entrada com 2.048 células com um *dropout* de 20%. Seguido da camada de entrada, uma camada escondida com 512 células com um *dropout* de mesmo valor da anterior. Finalizando o modelo, uma camada densa com funções de ativação do tipo linear foi adicionada a arquitetura. Os modelos foram configurados de forma a predizer 10 minutos adiante a partir de uma arquitetura LSTM do tipo *many-to-many*. Os resultados mostram que, para um conjunto de dados de séries temporais financeiras, o modelo LSTM tem melhor desempenho na generalização desta categoria de problema.

Complementando essa ideia, em Pawar, Jalem e Tiwari (2018), os autores apresentam

um comparativo entre algoritmos tradicionais de ML (análise de regressão, *support vector machine* (SVM), *random forest* e RNAs) e um misto de redes LSTM com redes recorrentes tradicionais na tarefa de previsão de séries temporais financeiras (PAWAR; JALEM; TIWARI, 2018). Neste estudo, uma diversificação de arquiteturas híbridas entre LSTM e um conjunto de arquiteturas de redes recorrentes LSTM e redes neurais recorrentes são testadas e avaliadas para a tarefa de previsão de movimentação no mercado de ações. Os resultados mostraram que o modelo baseado em redes neurais recorrentes obteve melhor desempenho quando comparado aos métodos convencionais de aprendizado, ratificando os trabalhos anteriores apresentados. Além disso, os autores argumentam que o modelo baseado em redes neurais-LSTM gerado pode ser útil tanto para *traders* individuais como para investidores corporativos.

Em Eapen, Bein e Verma (2019), os autores propõem uma abordagem de aprendizado profundo para prever preços de ações no mercado financeiro (EAPEN; BEIN; VERMA, 2019). A abordagem proposta envolve a previsão do preço de um ativo com uma semana de antecedência a partir de uma sequência de dados baseada em uma série temporal. A abordagem teve um desempenho melhor quando comparada com os resultados do estado da arte atual. O modelo proposto combina o uso de uma rede neural recorrente bidirecional LSTM e redes neurais convolucionais (CNNs). A CNN é usada para obter outra sequência de tamanho pequeno. Os autores realizaram os experimentos utilizando um conjunto de dados disponível no *Yahoo Finances*. Este estudo verifica que a combinação de camadas convolucionais e redes LSTM permitem um melhor desempenho quando comparadas com um regressor baseado em SVM.

A relevância das pesquisas discutidas é notória, dado os estudos do estado da arte atual relacionados ao uso de ML na previsão de séries temporais financeiras (HIRANSHA et al., 2018; REDDY, 2018; LIU; LIAO; DING, 2018; ZHANG et al., 2019). No entanto, estas pesquisas são limitadas quando extrapolada a inferência de períodos mais extensos. Poucos estudos investigaram o comportamento das projeções financeiras de longo prazo e a literatura relacionada ao domínio da previsão de preços no mercado de ações é significativamente baixa. Assim, a presente pesquisa busca diferenciar-se das demais explorando o comportamento da previsão de longo prazo e investigando o desempenho de técnicas de ML com foco nas RNAs como agente de apoio à compra e venda de ativos financeiros em bolsas de valores nos mais diversos panoramas comerciais.

3 FUNDAMENTAÇÃO TEÓRICA

Nesta seção é apresentada uma contextualização teórica sobre as ideias e conceitos utilizados para desenvolver esta obra.

3.1 Mercado Financeiro

Os mercados financeiros permitem a compra e venda de ações emitidas por empresas públicas e privadas. É papel do mercado financeiro fornecer um meio seguro e eficiente para que os investidores possam negociar, tanto com estratégias de curto quanto de longo prazo, o que

garante alta liquidez diariamente. As bolsas contêm uma infinidade de produtos financeiros, *e.g.*, títulos do governo, tesouro direto, índices, *commodities*, entre outros. Dentre esses produtos, o mercado de ações e o câmbio de moedas são os mais explorados pelos investidores. A previsão do preço das ações é um campo de pesquisa de grande interesse por parte dos investidores, bancos e empresários. Para esta tarefa são utilizados mecanismos com princípios distintos. Destacam-se as seguintes filosofias de previsão, a saber, análise fundamentalista, gráfica e técnica (CERVELLÓ-ROYO; GUIJARRO; MICHNIUK, 2015).

A análise fundamentalista tem fundamentos econômicos e fiscais para inferir previsões. A relação entre as ações e a saúde financeira das empresas, juntamente com seu panorama financeiro, são a melhor fonte de dados para se prever o futuro dos preços e tendências das ações. Esses aspectos abrangem o campo da microeconomia e macroeconomia. Enquanto a microeconomia se preocupa com os aspectos do equilíbrio regional, a macroeconomia aponta para questões nacionais e internacionais. Os principais recursos para este tipo de análise são os balanços, relatórios de produção, perspectivas de expansão e dados administrativos como salários, folha de pagamento e outros (EAPEN; BEIN; VERMA, 2019). A análise técnica e a análise gráfica são baseadas nas ideias formuladas por Charles Henry Dow, que foram alavancadas pelos estudos de Ralph Nelson Elliott de 1934, publicados em sua obra *The Wave Principle*. A análise gráfica compartilha o pensamento de que as informações necessárias para a previsão financeira são denotadas nos gráficos que cotam seus preços. Esta mentalidade implica a expressão do comportamento dos agentes econômicos por meio do registro dos movimentos das cotações de preços (CERVELLÓ-ROYO; GUIJARRO; MICHNIUK, 2015). A análise técnica diz respeito a uma perspectiva de natureza estatística. A sua gestão na previsão é orientada por teorias de probabilidade e pela utilização de correlação de variáveis financeiras. Cotações de preços de valores máximos e mínimos, bem como seus indicadores, que são produtos de funções matemáticas que tentam medir o grau de volatilidade e o volume financeiro, são usados para modelar as previsões de ações (LIU; LIAO; DING, 2018).

3.2 Redes Neurais recorrentes do tipo long short-term memory

O ML é uma subárea da inteligência artificial (IA) que visa criar técnicas de aprendizagem computacional, bem como o desenvolvimento de soluções que podem aprender automaticamente. Mitchell (1997, p. 2) conceitua o ML como um sistema de aprendizagem capaz de tomar decisões com base na experiência adquirida (MITCHELL, 1997). Com base nisso, os sistemas baseados em ML trazem características particulares que permitem o tratamento de problemas de classificação, predição, categorização entre outros. Uma especificação geral de modelos baseados em ML está relacionada à ideia de seu processo de aprendizagem, que pode ser supervisionado ou não supervisionado.

Em termos de problemas de previsão, o uso de RNAs tem ganhado muita notoriedade, principalmente com a evolução do desempenho computacional ocorrida nas últimas décadas. Assim, com base na ideia de simular o funcionamento do sistema nervoso de seres vivos no

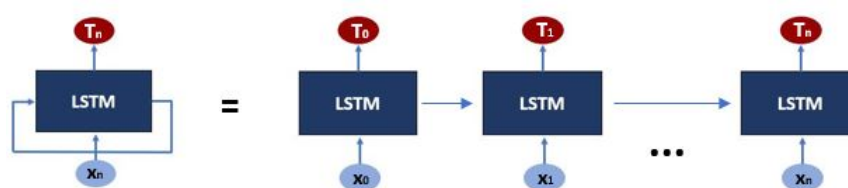
processo de resolução de problemas, as RNAs se destacam por se adaptarem a problemas não lineares. Existem várias arquiteturas de RNAs. *Perceptron* multicamadas, rede Adaline e recorrente são exemplos de arranjos estruturais usados na resolução de problemas. Por causa das RNNs, estas técnicas são usadas ativamente em tarefas de regressão e problemas em que a sequência é um dos principais fatores para sua resolução (HAYKIN, 2001).

Aprendizado profundo é uma subárea do ML que se concentra no uso de RNAs para mapear intuições entre conjuntos de dados e categorias (DEORUKHKAR et al., 2019). Hoje em dia, as redes neurais profundas são caracterizadas por múltiplas camadas e neurônios, maneiras distintas e sofisticadas de conectar camadas e neurônios, utilização do poder computacional disponível nas placas de processamento gráfico atuais para treinamento e extração automática de atributos. Assim, uma rede neural profunda pode ser definida como uma RNA com um grande número de parâmetros e camadas (PATTERSON; GIBSON, 2017). Redes neurais recorrentes são RNAs que incorporam uma arquitetura de *feedback*, em que seus resultados são usados para alimentação. Isso significa que as soluções obtidas pelo seu processamento são atuais, pois levam em consideração valores anteriores e não apenas respostas independentes como nas redes tradicionais (HAYKIN, 2001).

Baseadas na arquitetura de rede neural recorrente de Hopfield, as redes LSTM têm características semelhantes em sua operação. A sua criação deve-se à motivação para resolver problemas relacionados com a sequência, atributo das redes recorrentes, mas em uma perspectiva de médio prazo. Ao contrário das redes recorrentes tradicionais, que armazenam informações de curto prazo, fornecendo apenas o contexto mais recente dos dados, as redes LSTM procuram expandir a base do problema definindo a situação de longo prazo (BOCCATO, 2013). Essa característica implica que, além das informações mais recentes armazenadas em redes recorrentes, as informações passadas também são perpetuadas no ciclo de processamento, conferindo-lhes o recurso de memória de longo prazo.

A especificidade da arquitetura LSTM é conferida pelo suporte à capacidade sequencial de sua estrutura, conforme mostrado na Figura 1. Desta forma, garante a perpetuação da informação por um período dinâmico, que pode ser de curto prazo, quando poucas células são alocadas sequencialmente, ou longo prazo, quando muitas células são dispostas sequencialmente (PAWAR; JALEM; TIWARI, 2019). Esta operação é promovida por técnicas de aprendizado profundo, que teorizam a distribuição paralela nas camadas de células neurais.

Figura 1 – Rede neural *long short-term memory*.



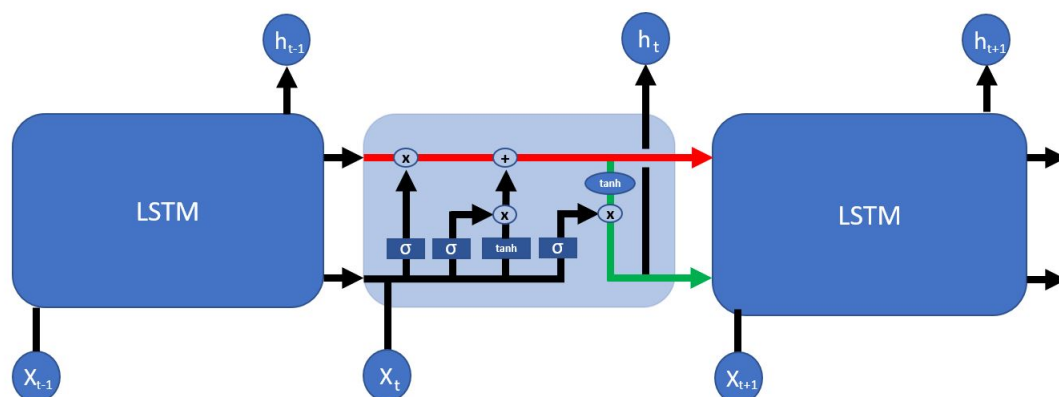
A Figura 1 apresenta a recorrência da informação em que, após o processamento da rede, sua saída é redirecionada para alimentá-la novamente. Esse processo se perpetua em cascata

e, além de ser realimentado com o resultado do processamento, é repassado para as próximas camadas.

Dessa forma, o modelo de arquitetura LSTM cria um *loop* de informações, armazenado em células de memória, que permite a expansão da conformidade contextual para resolver problemas naturais de série temporal. Essa capacidade de armazenamento permite o *feedback* de informações em diferentes níveis. Isso acarreta o acúmulo de conhecimento, levando ao ganho de informações na sua operação, principalmente nas tarefas de previsão.

A Figura 2 apresenta o detalhamento da arquitetura da célula LSTM. Neste modelo observa-se que quatro estruturas fazem parte da constituição da célula LSTM. Elas estão demarcadas pela cor azul marinho e seguem a formação de duas camadas de rede neural com função de ativação sigmoide, uma camada de rede neural com função de ativação tangente hiperbólica e, finalmente, uma nova camada de rede neural com função de ativação sigmoide. As camadas com função de ativação sigmoide, tem como função atuar como represa de informações inseridas em X_t , sendo denominadas “portas”. Como o resultado da função sigmoide variam de valores entre 0 e 1, ficam responsáveis por determinar as quantidades de informações dos dados que serão repassadas para as camadas seguintes. A camada formada por tangente hiperbólica, por sua vez, cria um vetor de novos valores que são adicionados no processamento da RNA. Dessa forma, com o devido treinamento destas células, a LSTM elimina dados que não agregam a predição e armazena informações em suas portas que são repassadas através de atuadores, indicados pelos vetores orientados. Dessa forma, no vetor orientado superior, indicado pela cor vermelha, temos o armazenamento de informações ponderadas pelas camadas e que caracterizam o armazenamento de informações de curto prazo. Já no vetor orientado inferior de cor verde, fica-se armazenado o valor resultante da previsão, indicando a predição que também é repassada para célula LSTM seguinte, orientando-a sobre a perspectiva de longo prazo que está inserida.

Figura 2 – Detalhamento de uma rede neural *long short-term memory*.



A necessidade de contextualização da memória de longo prazo justifica a adoção de redes neurais LSTM no problema de predição de séries financeiras (KIM; WON, 2018). Esta solução é reiterada pelos postulados da análise gráfica, em que a precificação de ações e títulos financeiros são justificados por seu histórico. Assim, dada a capacidade de conhecimento de curto e longo

prazo da rede LSTM, ela se apresenta como um candidato potencial para a tarefa de prever o comportamento financeiro.

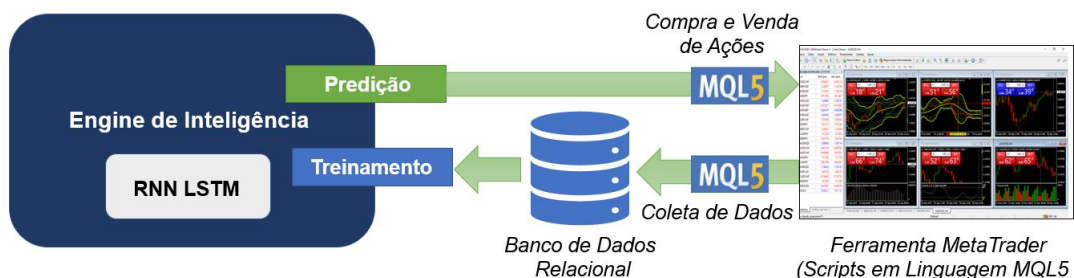
4 CAMINHOS METODOLÓGICOS

Esta secção detalha o processo de implementação e execução, bem como o planejamento utilizado para se realizar os experimentos estudados.

4.1 Arquitetura da Solução Proposta

A Figura 3 descreve a arquitetura do sistema de predição proposto. Neste esquema, a ferramenta *MetaTrader*¹ é inicialmente usada para automatizar a captura de ativos e executar ordens de compra e venda, pois sua plataforma permite a criação de *scripts* automatizados usando a linguagem MQL5. A partir da coleta de dados, estes serão armazenados em um banco de dados relacional, que servirá para construir o modelo preditivo sugerido, integrando um motor de inteligência na forma de uma interface de programação de aplicativos (API). Após o treinamento e validação do modelo preditivo. As predições feitas pelo modelo serão integradas com o *MetaTrader*, em que serão utilizadas para orientar a compra ou venda dos ativos financeiros.

Figura 3 – Arquitetura da solução proposta.



4.2 Preparação de Dados

A fase de pré-processamento dos dados busca a construção do conjunto de dados final, que *a priori* consiste na cotação histórica das aberturas do ativo Euro-Dólar, a partir desses dados brutos inicialmente coletados e necessários para a realização das previsões. Nessa etapa, também foram verificadas a qualidade dos dados, *outliers* e tratamento dos dados, como discretização e normalização. Seguindo essa heurística, cada atributo do conjunto de dados foi escalonado usando a normalização Min-Max disponível na biblioteca *Scikit-Learn* (PEDREGOSA et al., 2011), conforme expresso em (1).

$$x_{scaled} = (x - x_{min}) / (x_{max} - x_{min}) \quad (1)$$

¹<https://www.metatrader5.com/>

Em Liu, Liao e Ding (2018) e Pang *et al.* (2018), os autores apresentam modelos de predição que usam arquiteturas de rede LSTM (LIU; LIAO; DING, 2018; PANG et al., 2018). Nestes estudos, a previsão é feita para inferir um elemento futuro de cada vez. No entanto, dada a necessidade de estimar mais de um elemento, esta pesquisa avalia uma arquitetura baseada em LSTM que prevê mais de um elemento. Portanto, considerando a integridade temporal da amostra, os dados contidos no conjunto de dados *A* foram manipulados. Como uma série temporal tem índices que variam de 1 a t , que representam respectivamente o primeiro e o último elemento coletados, a meta da previsão é o conjunto de elementos cujos índices variam de $t + 1$ a $t + n$. Portanto, com base no conjunto de dados *A*, este estudo buscou como objetivo prever os elementos $t + 1$, $t + 2$ e $t + 3$. Dada a razão, a alteração da periodicidade da série temporal do conjunto de dados *A* para dois e três períodos, respectivamente, foi considerada, resultando no conjunto de dados *B* e *C*, nessa ordem, representados nas figuras 5 e 6.

Figura 4 – Conjunto de dados de treinamento para a rede de memória de longo prazo *A*.

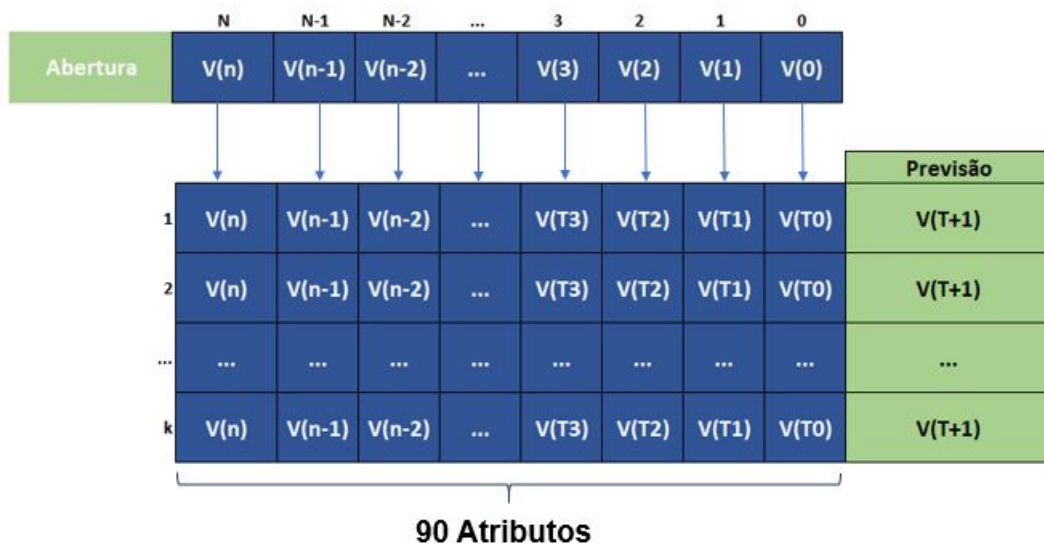


Figura 5 – Conjunto de dados de treinamento para a rede de memória de longo prazo *B*.

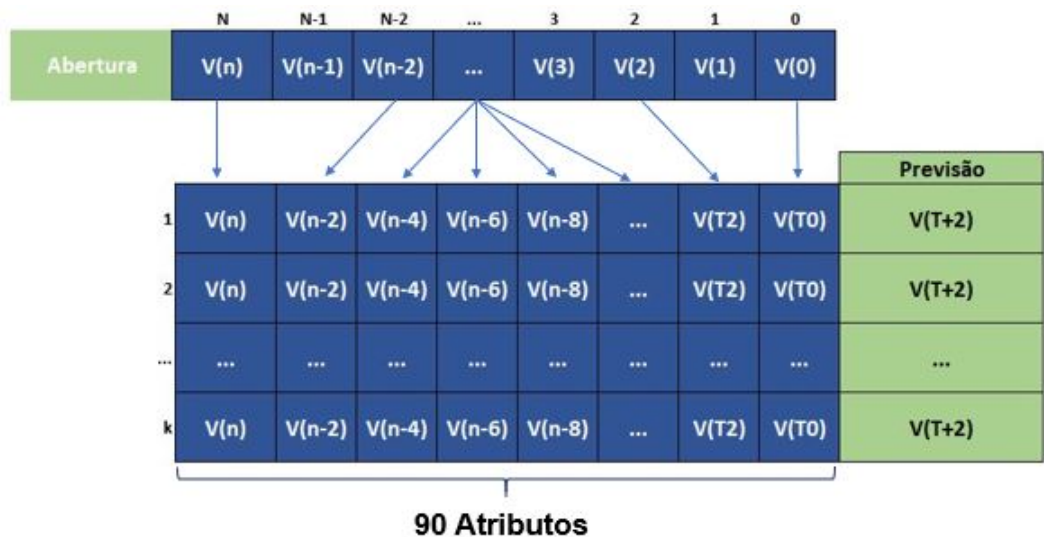
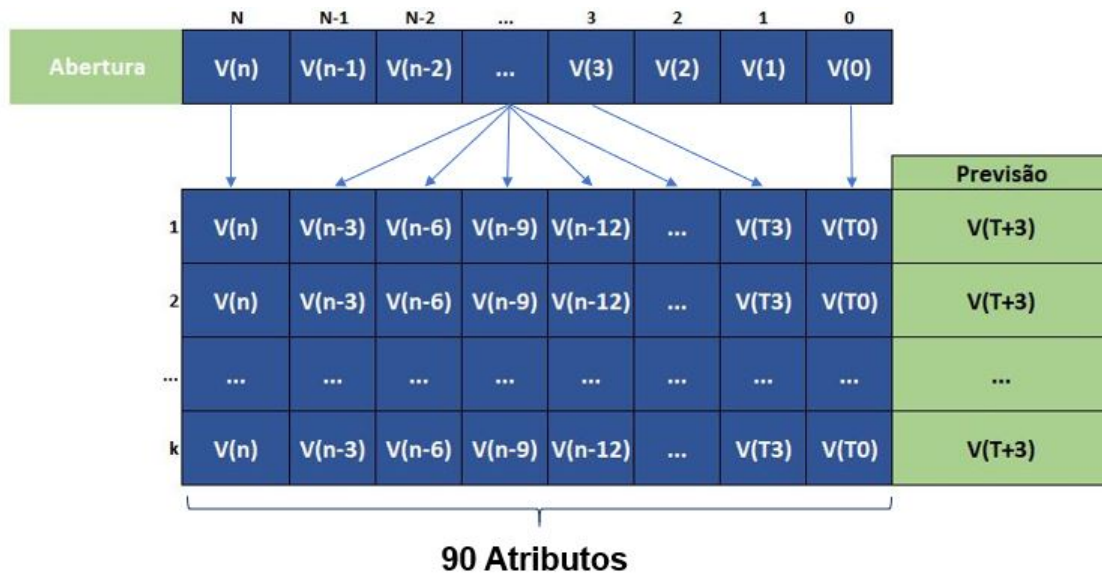


Figura 6 – Conjunto de dados de treinamento para a rede de memória de longo prazo C.



A partir dos três conjuntos de dados, três estruturas de rede neural LSTM foram modeladas, as quais permitem a previsão de três elementos futuros, a saber, os elementos $t + 1$, $t + 2$ e $t + 3$.

4.3 Experimentos

Para a avaliação das técnicas de ML foram utilizadas bibliotecas externas como *Scikit-learn* (PEDREGOSA et al., 2011) e a linguagem de programação *Python* para tarefas de ciência de dados, permitindo a rápida prototipagem em tarefas de mineração. A biblioteca *Keras*² foi usada para permitir a construção de RNAs em alto nível, possibilitando o processamento via *TensorFlow*³. A raiz do erro quadrático médio (RMSE) e o erro médio percentual absoluto (MAPE), descritos em (2) e (3), respectivamente, foram utilizados para analisar o desempenho dos modelos.

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n \left(\frac{d_i - f_i}{\sigma_i} \right)^2} \quad (2)$$

$$MAPE = \frac{1}{n} \sum_{t=1}^n \left| \frac{A_t - F_t}{A_t} \right| \times 100 \quad (3)$$

A arquitetura da RNA utilizada neste estudo consiste em uma camada de entrada que recebe um vetor de entrada com 90 atributos, referente a cada instância nos conjuntos de dados. Esses atributos descrevem os últimos 90 valores de abertura do preço do estoque analisado. A arquitetura da rede é composta por três camadas ocultas, cada uma composta por 50 nós e uma camada de saída com apenas um nó, que fornecerá o valor estimado para um estoque. A função de ativação utilizada foi linear, descrita pela expressão $f(x) = ax$. Este estudo também

²<https://keras.io>

³<http://tensorflow.org/>

Tabela 1 – Arquitetura de rede neural artificial.

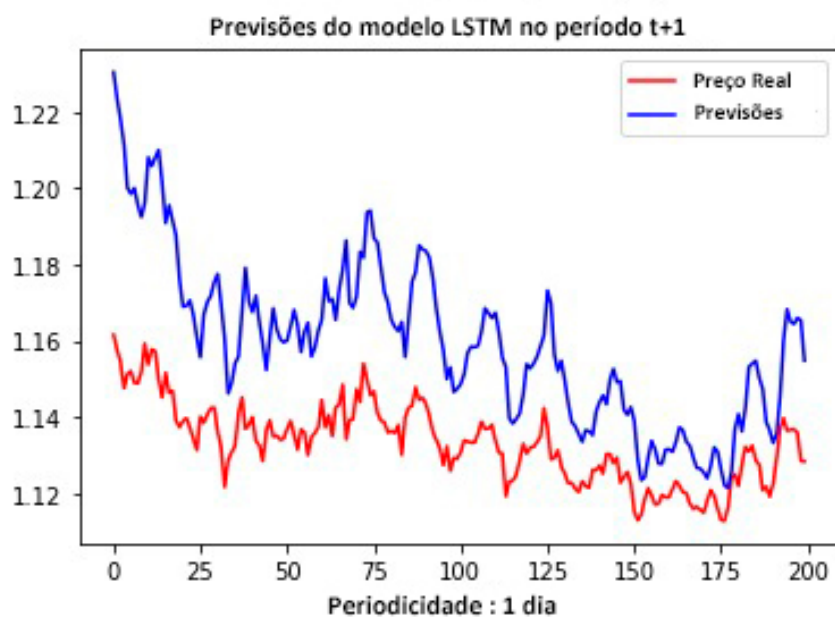
Camada	Tipo	Forma de saída	Parâmetros
<i>Camada de entrada</i>	(LSTM)	(90, 100)	40800
<i>Camada Oculta 1</i>	(LSTM)	(90, 50)	30200
<i>dropout 1 (Dropout)</i>	(90, 100)	0	
<i>Camada Oculta 2</i>	(LSTM)	(90, 50)	20200
<i>dropout 2 (Dropout)</i>	(90, 50)	0	
<i>Camada Oculta 3</i>	(LSTM)	(50)	20200
<i>dropout 3 (Dropout)</i>	(50)	0	
<i>Camada de Saída</i>	(Dense)	(1)	51

aplicou um *dropout* de 30% em cada camada oculta para evitar *overfitting*. A Tabela 1 apresenta a arquitetura da RNA. Esta estrutura de rede foi escolhida arbitrariamente e contém um total de 111.451 parâmetros treináveis.

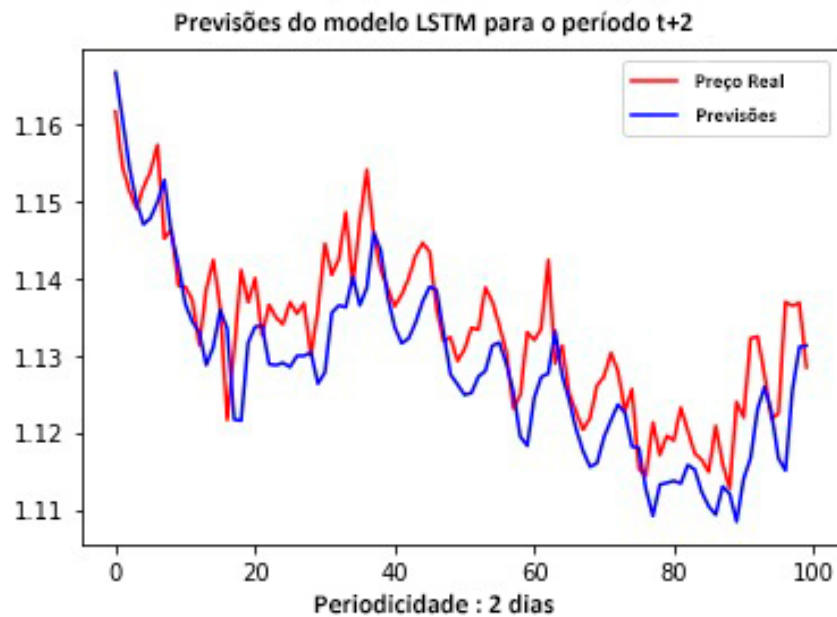
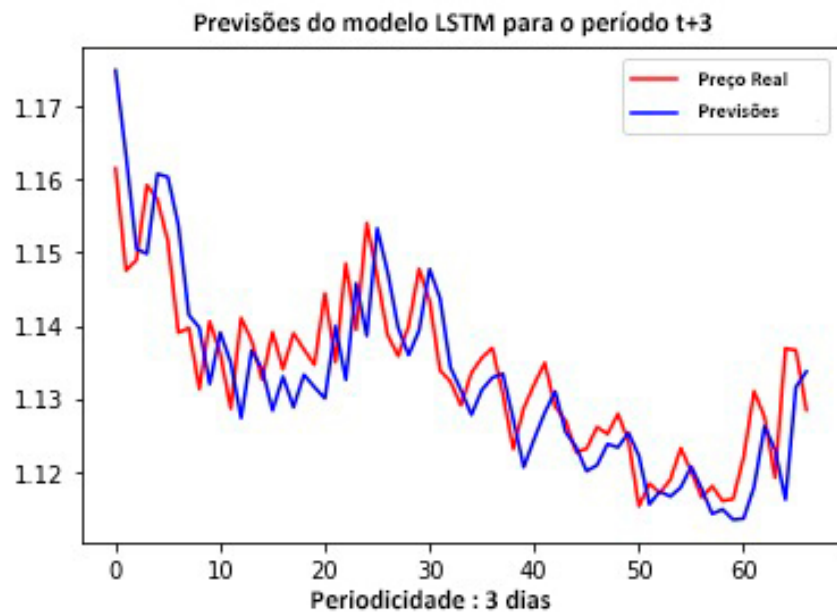
5 RESULTADOS E DISCUSSÃO

A partir da realização das etapas de treinamento e avaliação dos modelos preditivos, esta seção apresenta os resultados individuais das previsões obtidas após a normalização dos dados e, finalmente, o resultado combinado das previsões. Os testes realizados para avaliar a precisão do modelo baseiam-se nos dados dos ativos Euro-Dólar, que vão de setembro de 2018 a julho de 2019.

Figura 7 – Conjunto de previsões da rede A.



Individualmente, é possível perceber que os resultados obtidos pelas projeções generalizaram-se bem ao cenário atual. A rede A obteve o maior erro médio absoluto de 0,0298, enquanto as redes B e C obtiveram erros de 0,0198 e 0,0195, respectivamente.

Figura 8 – Conjunto de previsões da rede *B*.Figura 9 – Conjunto de previsões da rede *C*.

A Tabela 2 apresenta as métricas de desempenho RMSE e MAPE. A RNA *A* apresentou muitos ruídos. Embora o modelo tenha se adaptado bem, quando comparado com as redes *B* e *C*, seus resultados foram abaixo do esperado. Devido a isso, as previsões do modelo foram substancialmente menores do que o esperado, embora o RMSE e o MAPE estivessem dentro da faixa aceitável, conforme mostrado na tabela.

Figura 10 – Previsões resultantes do modelo proposto.

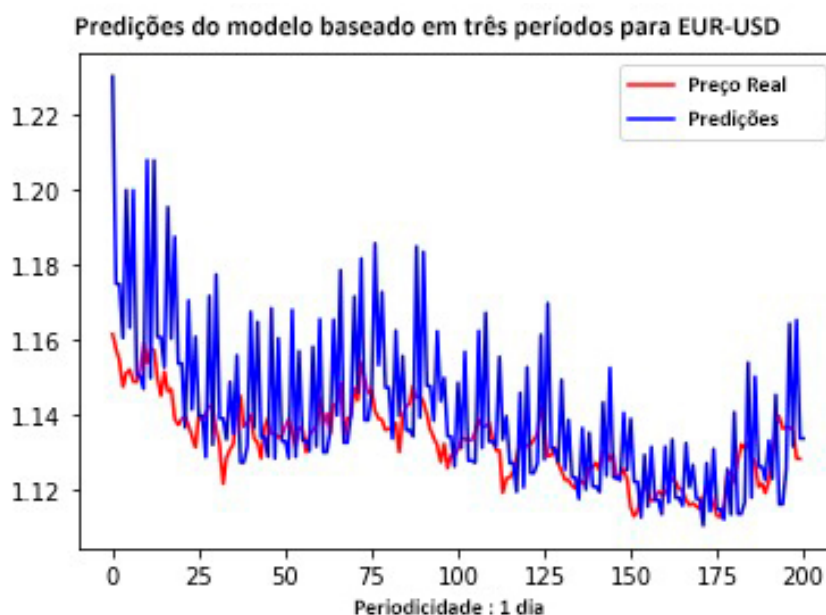


Tabela 2 – Avaliação de redes de memória de longo prazo.

Métrica	Modelo A	Modelo B	Modelo C
RMSE	0.000861	0.000006	0.0000055
MAPE	2.319	0.54187	0.5179

6 CONCLUSÕES E TRABALHOS FUTUROS

O uso de redes neurais LSTM no processo de previsão de séries temporais financeiras tem demonstrado grande potencial para inferência no curto prazo. A aplicação da arquitetura proposta para inferir mais de um elemento é, portanto, ainda objeto de análise uma vez que os resultados obtidos ainda não são satisfatórios para auxiliar efetivamente em uma negociação, embora defina claramente a tendência do ativo analisado.

Direcionamentos futuros desta pesquisa envolvem a busca pela combinação dos modelos descritos com outras técnicas de ML, a fim de obter melhores previsões. Além disso, pretende-se investigar outras estratégias sobre o comportamento da projeção das séries financeiras no curto prazo com múltiplos períodos à frente.

REFERÊNCIAS

ARVIDSSON, P.; ÅNHED, T. **Sequence-to-sequence learning of financial time series in algorithmic trading**. 2017.

BAGHERI, A.; PEYHANI, H. M.; AKBARI, M. Financial forecasting using anfis networks with quantum-behaved particle swarm optimization. **Expert Systems with Applications**, Elsevier, v. 41, n. 14, p. 6235–6250, 2014.

BOCCATO, L. **Novas propostas e aplicações de redes neurais com estados de eco**. 210 p. Tese (Doutorado) — Universidade Estadual de Campinas, Unicamp, SP, 2013.

CERVELLÓ-ROYO, R.; GUIJARRO, F.; MICHNIUK, K. Stock market trading rule based on pattern recognition and technical analysis: Forecasting the djia index with intraday data. **Expert systems with Applications**, Elsevier, v. 42, n. 14, p. 5963–5975, 2015.

DEORUKHKAR, O. S. et al. Stock price prediction using combination of lstm neural networks, arima and sentiment analysis. **International Research Journal of Engineering and Technology**, v. 6, n. 3, p. 3497–3503, 2019.

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)**. [S.l.], 2019. p. 0264–0270.

FISCHER, T.; KRAUSS, C. Deep learning with long short-term memory networks for financial market predictions. **European Journal of Operational Research**, Elsevier, Amsterdam, The Netherlands, v. 270, n. 2, p. 654–669, 2018.

HATCHER, W. G.; YU, W. A survey of deep learning: platforms, applications and emerging research trends. **IEEE Access**, IEEE, v. 6, p. 24411–24432, 2018.

HAYKIN, S. Neural networks: principles and practice. **Bookman**, 2001.

HIRANSHA, M. et al. Nse stock market prediction using deep-learning models. **Procedia computer science**, Elsevier, v. 132, p. 1351–1362, 2018.

KIM, H. Y.; WON, C. H. Forecasting the volatility of stock price index: A hybrid model integrating lstm with multiple garch-type models. **Expert Systems with Applications**, Elsevier, v. 103, p. 25–37, 2018.

LIU, S.; LIAO, G.; DING, Y. Stock transaction prediction modeling and analysis based on lstm. In: **13th IEEE Conference on Industrial Electronics and Applications (ICIEA)**, May 31-June 2. Wuhan, China: IEEE, 2018. p. 2787–2790.

MITCHELL, T. Machine learning. **Burr Ridge, IL: McGraw Hill**, 1997.

PANG, X. et al. An innovative neural network approach for stock market prediction. **The Journal of Supercomputing**, Springer, (*in press*), p. 1–21, 2018.

PATTERSON, J.; GIBSON, A. **Deep Learning: A Practitioner's Approach**. [S.l.]: "O'Reilly Media, Inc.", 2017.

PAWAR, K.; JALEM, R. S.; TIWARI, V. Stock market price prediction using lstm rnn. In: _____. **Emerging Trends in Expert Applications and Security**. Singapore: Springer, 2018. p. 493–503.

PAWAR, K.; JALEM, R. S.; TIWARI, V. Stock market price prediction using lstm rnn. In: **Emerging Trends in Expert Applications and Security**. [S.l.]: Springer, 2019. p. 493–503.

PEDREGOSA, F. et al. Scikit-learn: Machine learning in Python. **Journal of Machine Learning Research**, v. 12, p. 2825–2830, Oct. 2011.

REDDY, V. K. S. Stock market prediction using machine learning. **International Research Journal of Engineering and Technology (IRJET)**, v. 5, n. 10, p. 1033–1035, 2018.

THIELE, C. C.; ADAMI, A. G. Previsão de séries temporais financeiras: modelo baseado em redes neurais artificiais. **Revista Brasileira de Computação Aplicada**, v. 8, n. 2, p. 113–128, 2016.

ZHANG, K. et al. Stock market prediction based on generative adversarial network. **Procedia computer science**, Elsevier, v. 147, p. 400–406, 2019.