

Aplicação de Técnicas de Inteligência Artificial para Previsão e Recomendação de Investimentos no Mercado Financeiro

Gabriel S. Rosa
Centro Federal de Educação
Tecnológica de Minas Gerais
DECOM - CEFET-MG
Divinópolis, MG, Brasil
ms.gabriel.rosa@gmail.com

Rodrigo M. A. Silva
Centro Federal de Educação
Tecnológica de Minas Gerais
PPGMMC - CEFET-MG
Belo Horizonte, MG, Brasil
rodrigopiece27@gmail.com

Alisson Marques Silva
Centro Federal de Educação
Tecnológica de Minas Gerais
DECOM - PPGMMC - CEFET-MG
Divinópolis, MG, Brasil
alisson@cefetmg.br

Resumo—Este estudo propõe a aplicação de técnicas de inteligência artificial (IA) no mercado financeiro para previsão e recomendação estratégica de investimentos. O trabalho aborda a teoria da eficiência de mercado, destacando a influência de emoções e comportamentos na formação dos preços das ações, e investiga a eficiência de mercados emergentes, revelando oportunidades decorrentes de desvios sistemáticos nos preços. A metodologia proposta abrange a extração, geração e seleção de variáveis, culminando na implementação de uma máquina de previsão que combina modelos estatísticos, de classificação e de regressão por meio de técnicas de *ensemble*. Os experimentos foram conduzidos com três ativos financeiros (PETR3, WINFUT e WDOFUT), em diferentes granularidades temporais. Os resultados indicam a viabilidade de se alcançar retornos significativos, com a máquina de previsão superando a estratégia *buy and hold* em 83,33% dos casos. Essa abordagem integrativa mostra-se promissora para orientar decisões financeiras mais informadas e estratégicas.

Index Terms—Previsão de ações, mercado financeiro, inteligência artificial, séries temporais, análise de dados financeiros, *ensemble*.

I. INTRODUÇÃO

Os primeiros registros de mercados financeiros similares aos contemporâneos surgiram no século XVII na Holanda, com a criação da Bolsa de Valores de Amsterdam, onde comerciantes negociavam ações da Companhia das Índias Orientais [1]. Posteriormente, no final do século XIX, nos Estados Unidos, Charles H. Dow criou o *Dow Jones Industrial Average* (DJIA), aplicando uma média ponderada dos preços das ações de 12 empresas industriais americanas [2], marcando o início do mercado de índices.

A teoria da eficiência de mercado afirma que os preços dos ativos refletem todas as informações disponíveis publicamente [3], tornando improvável obter lucros anormais consistentes baseando-se apenas em informações públicas [9]. No entanto, existem evidências de que os preços podem ser influenciados por fatores emocionais e comportamentais [10]. Um estudo de [11] investigou os coeficientes de Hurst das séries de preços de ações, indicando que mercados emergentes são menos eficientes que os desenvolvidos, resultando em oportunidades para investidores obterem retornos consideráveis.

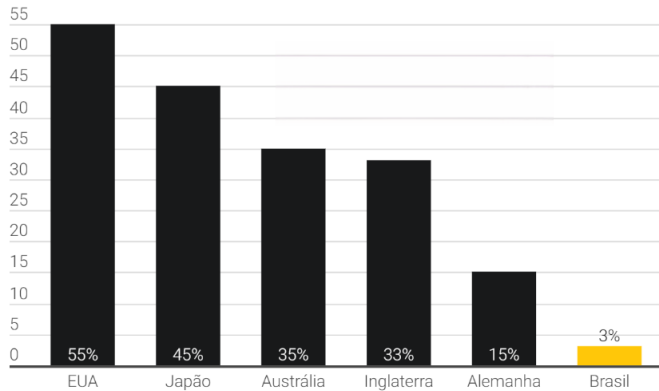
Neste cenário, a busca por estratégias menos arriscadas originou diversos modelos aplicáveis ao mercado financeiro, baseados em Análise Técnica, Análise Fundamentalista, Análise Quantitativa e Análise de Sentimentos. Como consequência, surgiram modelos baseados em Inteligência Artificial (IA) que combinam essas abordagens, permitindo a análise de grandes volumes de dados e o reconhecimento de padrões complexos [12], [13].

Gonzalo e Pitarakis [18] mostraram que a previsão de séries temporais pode ser uma ferramenta útil na especulação de retornos de ações. Técnicas de IA podem ser aplicadas a tarefas de previsão de sentido e/ou valor de ativos financeiros, devido à sua capacidade de identificar padrões dificilmente observados por humanos [19]. Além disso, essas técnicas se justificam diante do recente cenário de períodos de crise, como a pandemia de COVID-19, a volatilidade pós-pandemia [6] e as recentes tensões geopolíticas. Nesses contextos, a eficiência dos mercados financeiros tende a se deteriorar temporariamente, abrindo espaço para padrões mais previsíveis no curto prazo [7]. Estudos mostram que, nesses momentos, os preços apresentam maior autocorrelação, persistência e rupturas na aleatoriedade, o que indica desvios da Hipótese dos Mercados Eficientes [8]. Com isso, os modelos que empregam IA se tornam ferramentas eficazes para explorar ineficiências transitórias e desenvolver estratégias de negociação mais precisas em ambientes marcados por alta instabilidade.

O crescente interesse e participação de pessoas físicas no mercado financeiro têm sido observados globalmente. A globalização dos mercados e o aumento da renda das famílias em países em desenvolvimento são fatores que impulsionam essa tendência [14]. Além disso, a facilidade de acesso à informação e o uso de tecnologias financeiras têm contribuído para o aumento da participação de pessoas físicas na bolsa [15]. Um estudo da XP Investimentos [16] revela que o número de pessoas físicas investindo em ações está diretamente relacionado à estabilidade do mercado e à qualidade de vida nos países analisados. Na bolsa brasileira (B3), houve um aumento de mais de 2,8 milhões de investidores individuais em 2020,

representando um crescimento de 93,7% em relação a 2019 [17]. A Figura 1 ilustra a porcentagem de pessoas que investem em ações em diversos países. Destaca-se, nessa figura, que nos Estados Unidos 55% das pessoas investem em ações, enquanto no Brasil esse número é de apenas 3%.

Figura 1. Porcentagem de pessoas por país que investem em ações.



Diante desse contexto, torna-se relevante o desenvolvimento de modelos baseados em IA que possam apoiar decisões de investimento. O objetivo geral deste trabalho é a aplicação de técnicas de IA para previsão e recomendação de investimentos no mercado financeiro. Os objetivos específicos incluem: (i) revisar a literatura sobre mercado financeiro e técnicas de IA; (ii) propor e implementar uma abordagem para previsão de valores e sentido de variação; (iii) avaliar o desempenho da abordagem em diferentes ativos e granularidades; (iv) propor e implementar uma abordagem para recomendação de compra e venda; e (v) avaliar os resultados em comparação com estratégias tradicionais.

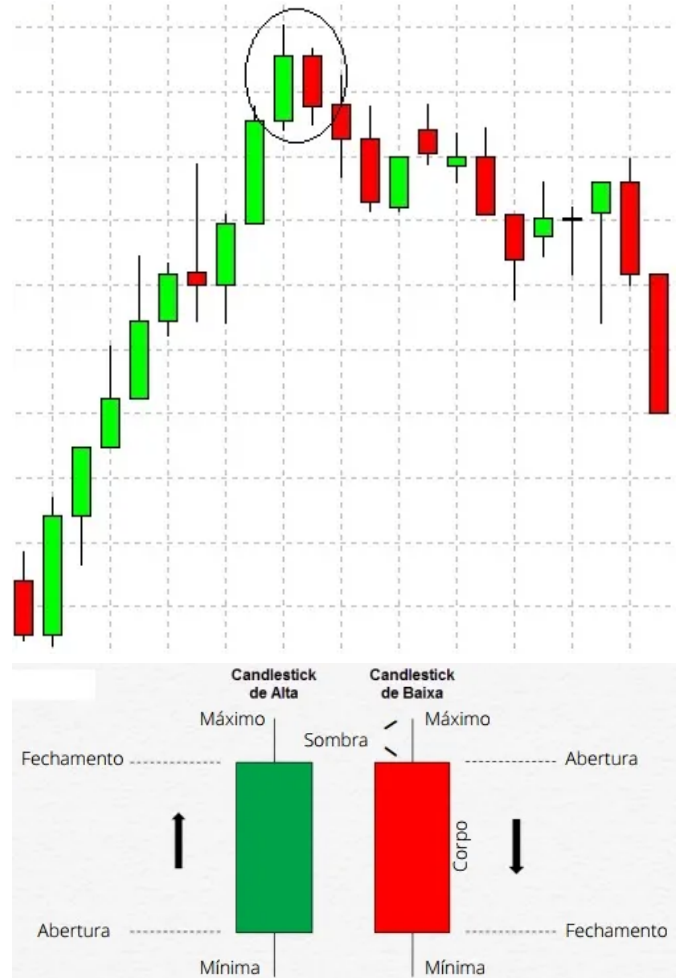
O restante deste artigo está organizado da seguinte forma. A próxima seção apresenta uma breve revisão da literatura sobre técnicas aplicadas na previsão e na recomendação de investimento no mercado financeiro. Em seguida, a Seção III detalha a abordagem proposta neste trabalho para previsão e recomendação de investimentos. A abordagem é avaliada por meio de experimentos computacionais na Seção IV. Por fim, a Seção V apresenta as considerações finais do trabalho, destacando suas contribuições e perspectivas para trabalhos futuros.

II. REVISÃO DA LITERATURA

Uma série temporal é uma sequência de observações coletadas em um intervalo de tempo, que reflete a dependência serial dos dados. Para analisá-la, comumente decompõe-se a série em componentes de tendência, ciclo e sazonalidade [20]. A tendência descreve o comportamento da série ao longo do tempo, os ciclos são padrões que se repetem em curto intervalo, e a sazonalidade refere-se a variações regulares em intervalos fixos.

O mercado financeiro gera dados que podem ser analisados como séries temporais, permitindo identificar tendências, ciclos e sazonalidades nos preços dos ativos. O gráfico de velas (vide Figura 2) é amplamente adotado para discretizar dados sem perda de informações relevantes [21], fornecendo valores de abertura, máxima, mínima e fechamento (OHLC, do inglês *Open, High, Low, Close*) para cada período.

Figura 2. Gráfico de velas e suas características observáveis.



Na literatura recente, diversas abordagens têm sido propostas para previsão no mercado financeiro. Sezer, Gudelek e Ozbayouglu [29] apresentaram uma revisão abrangente sobre o uso de deep learning para previsão financeira, destacando que *Convolutional Neural Networks* (CNN) e *Recurrent Neural Networks* (RNN) têm demonstrado resultados promissores. Além disso, Jiang, Xu e Liang [30] exploraram aplicações de aprendizado por reforço no mercado financeiro, mostrando como agentes podem aprender estratégias de negociação otimizadas por meio de interações com o ambiente de mercado.

Modelos híbridos também têm ganhado destaque na literatura recente. Zhang, Liu e Tan [32] propuseram um modelo que combina análise *wavelet* com redes *Long Short-Term Memory* (LSTM) para capturar tanto componentes de frequência

quanto dependências temporais em séries financeiras. Já Mutinda e Langat [4] combinaram o modelo *Generalized Autoregressive Conditional Heteroskedasticity* (GARCH), utilizado para estimar a volatilidade condicional dos retornos da ação da Airtel (empresa indiana), com três arquiteturas de redes neurais para séries temporais: LSTM, *Gated Recurrent Unit* (GRU) e *Transformer*. Nesse processo, as variâncias previstas pelo GARCH são inseridas como entradas adicionais nas redes, aproveitando simultaneamente os parâmetros estatísticos extraídos pelo GARCH e as capacidades individuais de modelagem temporal de cada arquitetura.

Safari e Ghaemi [5] propõem o modelo chamado NeuroFuzzyMan, uma abordagem híbrida que associa sistemas *neuro-fuzzy* a redes *Bidirectional Long Short-Term Memory* (BiLSTM) para realizar previsão de séries temporais financeiras. A lógica *fuzzy* é aplicada para lidar com as incertezas presentes nos dados de mercado, enquanto a BiLSTM é responsável por captar padrões temporais complexos e dependências de longo prazo. Segundo os autores, ao integrar a capacidade de tratar incertezas da lógica *fuzzy* com o poder de memória das BiLSTM, o NeuroFuzzyMan fornece análises mais robustas para embasar decisões de investimento.

A consideração de fatores comportamentais também tem sido incorporada aos modelos de previsão. Chen, Qiao, Wang, Du e Stanley [31] investigaram a integração de sentimentos de investidores extraídos de mídias sociais, demonstrando que essa informação pode melhorar significativamente a precisão das previsões.

Além da modelagem, a seleção de características é uma etapa crítica. Cavalcante, Brasileiro, Souza, Nobrega e Oliveira [33] apresentaram uma revisão sobre técnicas computacionais para essa tarefa, destacando a importância de métodos como *Principal Component Analysis* (PCA) e algoritmos genéticos. Patel, Shah, Thakkar e Kotecha [34] compararam quatro modelos de previsão — *Artificial Neural Network* (ANN), *Support Vector Machine* (SVM), *Random Forest* (RF) e *Naive Bayes* — e concluíram que a transformação dos dados de entrada em valores discretos de tendência melhora significativamente o desempenho dos modelos. Na previsão de tendência ou preço de ativos, destacam-se abordagens baseadas em análise gráfica [22], análise de sentimento [23] e análise fundamentada em dados numéricos [24].

Técnicas de *ensemble* são frequentemente incorporadas nessa etapa para aumentar a capacidade preditiva do modelo resultante [25]. As técnicas de *ensemble*, em particular, têm ganhado destaque na previsão financeira devido à sua capacidade de combinar múltiplos modelos para obter resultados mais robustos. Krauss, Do e Huck [35] aplicaram um *ensemble* de *deep learning*, *Gradient-Boosted Trees* (GBT) e *Random Forests* para previsão de retornos de ações do S&P 500, alcançando retornos anualizados de 0,45% antes dos custos de transação. De forma semelhante, Booth, Gerding e McGroarty [36] propuseram um sistema automatizado de negociação baseado em *ensemble* de modelos, demonstrando que a combinação de diferentes estratégias pode reduzir o risco e aumentar a estabilidade dos retornos.

Com uma abordagem integrativa, [37] desenvolveram um modelo de *ensemble* hierárquico que integra análise técnica, fundamentalista e de sentimento, evidenciando melhorias significativas na precisão da previsão. Por fim, [38] compararam as técnicas de *boosting* *AdaBoost*, *Gradient Boosting* (GB) e *Extreme Gradient Boosting* (XGBoost), concluindo que o XGBoost apresenta melhor desempenho em termos de precisão e velocidade de treinamento. Complementarmente, [39] propuseram um *ensemble* baseado em *Empirical Mode Decomposition* (EMD) e redes neurais, que superou modelos individuais em diversos indicadores.

Com uma abordagem integrativa, Kin, Shamsuddin e Lim [37] desenvolveram um modelo de *ensemble* hierárquico que integra análise técnica, fundamentalista e de sentimento, evidenciando melhorias significativas na precisão da previsão. Por fim, Wang, Sun, Liu e Cao [38] compararam as técnicas de *boosting* *AdaBoost*, *Gradient Boosting* (GB) e *Extreme Gradient Boosting* (XGBoost), concluindo que o XGBoost apresenta melhor desempenho em termos de precisão e velocidade de treinamento. Complementarmente, Li, Wang, Cao, Min e Deng [39] propuseram um *ensemble* baseado em *Empirical Mode Decomposition* (EMD) e redes neurais, que superou modelos individuais em diversos indicadores.

Os estudos avaliados, em geral, apresentam uma estrutura metodológica semelhante, que inclui as etapas de extração de dados, geração de variáveis, seleção de variáveis, treinamento de modelos e análise de resultados. A Figura 3 ilustra, de forma genérica, todas as etapas da estrutura metodológica adotada nos trabalhos analisados.

Figura 3. Estrutura metodológica típica dos estudos de previsão e recomendação de investimentos.



III. ABORDAGEM PROPOSTA

A abordagem proposta é estruturada em quatro etapas principais: construção dos conjuntos de dados, seleção de variáveis, previsão e recomendação de investimento.

A. Construção dos Conjuntos de Dados

O processo de construção dos conjuntos de dados é dividido em duas fases: extração de dados e geração de novas variáveis. A Figura 4 apresenta uma visão geral desse processo.

Na etapa de extração, foram utilizadas informações de plataformas como o *Investing*¹ e o *Profit*², que disponibilizam dados

¹<https://br.investing.com/>

²<https://www.nelogica.com.br/produtos/profit-pro>

históricos e em tempo real em diferentes intervalos. Após a extração, aplicaram-se diversas técnicas para gerar variáveis relevantes à tarefa de previsão, calculando novas variáveis a partir de cada valor de OHLC, utilizando diferentes fórmulas econômicas ajustadas a parâmetros específicos abordados na literatura.

Entre as técnicas utilizadas para a geração de variáveis, destacam-se:

- Valores Passados (VP) - obtido por meio do deslocamento de 7 valores passados;
- *Simple Moving Average* (SMA) - calculada a partir da média móvel simples aplicada a diferentes intervalos;
- *Exponential Moving Average* (EMA) - média móvel exponencial com mais peso atribuído aos preços recentes;
- *Moving Average Convergence Divergence* (MACD) - derivado da diferença entre médias móveis exponenciais;
- *Commodity Channel Index* (CCI) - indicador que mede a variação do preço em relação à média;
- *Average Directional Index* (ADX) - indicador que mede a força da tendência;
- *Rate of Change* (ROC) - mede a variação percentual do preço em relação a um período anterior;
- *True Strength Index* (TSI) - indicador de *momentum* que mostra a força por trás dos movimentos de preço;
- *Stochastic Oscillator* %K e %D - indicadores que comparam o preço de fechamento com a faixa de preço;
- *Williams %R* - indicador de *momentum* que mostra níveis de sobrecompra e sobrevenda.

Como resultado desse processo, foram geradas um total de 117 variáveis, proporcionando uma variedade de padrões de entrada para os modelos de previsão.

B. Seleção de Variáveis

Embora a base de dados contenha um grande número de variáveis, nem todas possuem a mesma relevância para a tarefa de previsão. Por isso, a seleção de variáveis foi realizada em duas etapas, conforme ilustrado na Figura 5.

Nas duas etapas, foram utilizados dois grupos de algoritmos: um voltado para modelos de classificação (Fisher, Gini, R2 e *Random Forest*) e outro para modelos de regressão (*Random Forest*, *Lasso*, *ElasticNet* e *Recursive Feature Elimination*). Após essa etapa, realizou-se um *merge* em cada grupo, unificando as variáveis selecionadas por cada método, sem repetição, e aplicando o teste de Kruskal-Wallis para ranquear

Figura 4. Visão macro do processo de construção do conjunto de dados.

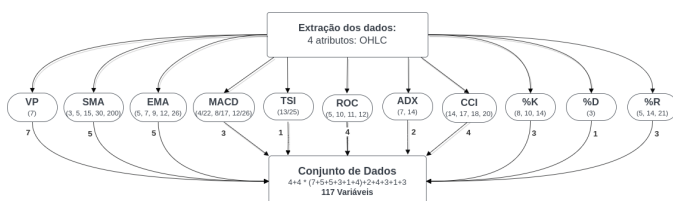
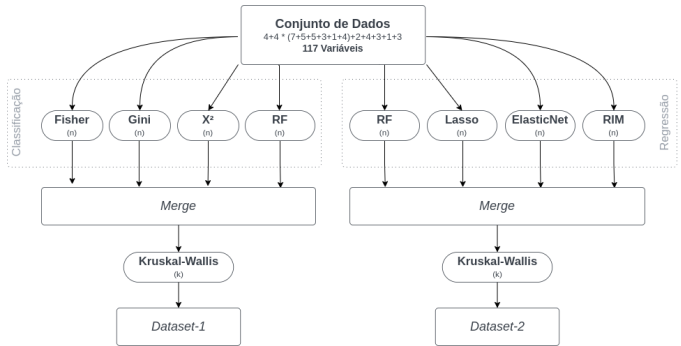


Figura 5. Visão macro do processo de seleção de variáveis.

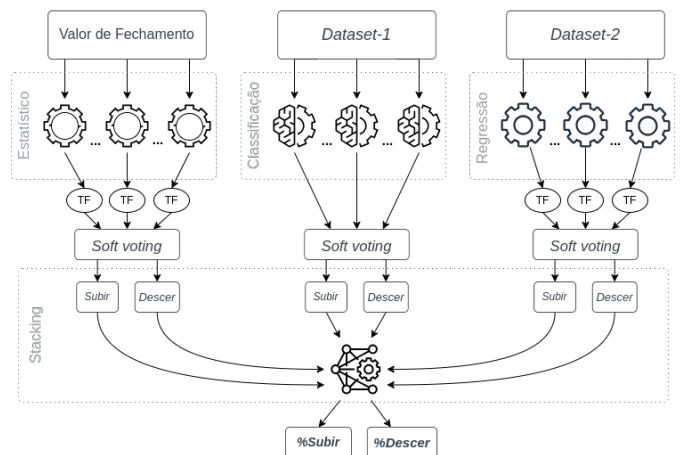


as variáveis mais relevantes. Isso resultou em dois conjuntos de dados: o *Dataset-1*, referente ao grupo de classificação, e o *Dataset-2*, referente ao grupo de regressão.

C. Máquina de Previsão

A etapa seguinte da abordagem é a construção da máquina de previsão, que utiliza duas técnicas de *ensemble*: *Soft Voting* e *Stacking*. A Figura 6 apresenta uma visão macro dessa etapa.

Figura 6. Visão macro do processo de previsão.



Para isso, foram construídos três grupos de algoritmos distintos:

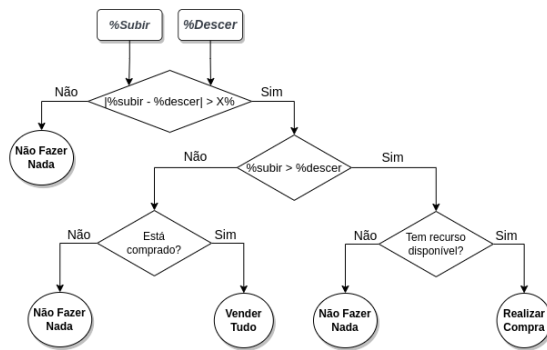
- **Modelos estatísticos:** *Autoregressive Integrated Moving Average* (ARIMA), *Seasonal ARIMA* (SARIMA) e *Generalized Autoregressive Conditional Heteroskedasticity* (GARCH), que utilizam apenas o valor de fechamento como entrada;
- **Técnicas de Inteligência Artificial para classificação:** *Support Vector Machine* (SVM), *Logistic Regression* (LR) e *k-Nearest Neighbors* (KNN), que utilizam o *Dataset-1* como entrada;
- **Algoritmos de Inteligência Artificial para regressão:** *Multilayer Perceptron* (MLP), *Support Vector Regression* (SVR) e *Random Forest* (RF), que utilizam o *Dataset-2* como entrada.

As técnicas de *Soft Voting* e *Stacking* são aplicadas de forma sequencial. O *Soft Voting* combina as previsões de cada grupo de modelos, ponderando a contribuição de cada um deles, resultando em dois valores: os graus de pertinência das classes 'subir' e 'descer'. Em seguida, o *Stacking* utiliza um algoritmo de IA para combinar as saídas do *Soft Voting* de cada grupo, gerando duas saídas finais referentes aos percentuais de correlação com cada classe.

D. Recomendação de Investimento

Com base nos percentuais de subida ou descida obtidos pela máquina de previsão, inicia-se o processo de recomendação de compra e venda, com o objetivo de reduzir operações de baixa confiabilidade e maximizar os retornos financeiros. A Figura 7 apresenta uma visão macro desta etapa.

Figura 7. Visão macro do processo de recomendação de investimentos.



A estratégia de recomendação começa com a aplicação de um filtro para verificar se a diferença percentual entre as classes de subida e descida é significativa. Caso não seja, recomenda-se não realizar nenhuma operação e manter a posição atual. Se a diferença for relevante, compara-se os sinais de 'subir' e 'descer' para identificar a classe com maior percentual de pertinência.

Se o sinal de subida for mais relevante, entende-se que há alta probabilidade de o próximo valor ser superior ao atual. Nesse caso, verifica-se a disponibilidade de recursos na carteira de investimentos. Se houver, recomenda-se a compra do ativo. Por outro lado, se o sinal de descida for predominante, verifica-se a existência de ativos já comprados. Se existirem, recomenda-se a venda total, adotando uma posição defensiva para evitar perdas potenciais.

IV. EXPERIMENTOS E RESULTADOS

A. Configuração Experimental

Os experimentos foram conduzidos utilizando três ativos: PETR3, WINFUT e WDOFUT, em duas granularidades distintas — 30 e 60 minutos. As bases de dados foram divididas em 10% para otimização de hiperparâmetros, 70% para treinamento e 20% para teste.

Foram avaliados dez modelos de previsão, agrupados em três categorias: modelos estatísticos (ARIMA, SARIMA,

GARCH), técnicas de IA para classificação (SVM, KNN, LR), e algoritmos de IA para regressão (MLP, SVR, RF), além do modelo final MLP OUT, que representa a máquina de previsão completa.

As métricas de avaliação incluíram Acurácia, F1-score e Matriz de Confusão para as tarefas de classificação, além da variação percentual acumulada e do número de operações para a etapa de recomendação de investimento.

B. Resultados Experimentais

A Tabela I ilustra os resultados obtidos pela abordagem proposta na classificação e na recomendação de investimentos com granularidades de 30 e 60 minutos para o ativo PETR3. Na previsão de direção (subida ou descida) com granularidade de 30 minutos, o melhor resultado foi obtido pelo modelo LR, seguido por SVM e MLP OUT empatados em segundo lugar. Na sequência, vieram KNN e MLP em quarto lugar e, por fim, o ARIMA na sexta posição. Com granularidade de 60 minutos, o modelo MLP apresentou o melhor desempenho, enquanto KNN e MLP OUT empataram na segunda colocação. Em seguida, vieram LR, SVM e ARIMA. Na tarefa de recomendação de investimentos, todos os modelos propostos superaram a estratégia *buy and hold* nos cenários avaliados. O destaque foi para o ARIMA na granularidade de 30 minutos, com desempenho 19,09% superior ao *buy and hold*, e para o KNN na granularidade de 60 minutos, com uma variação percentual de 40,51% acima da mesma referência.

Tabela I. Resultados de previsão de direção (subida ou descida) e recomendação de investimentos para PETR3.

Modelo	30 minutos		60 minutos	
	Acurácia (%)	Var. (%)	Acurácia (%)	Var. (%)
ARIMA	52,35	19,09	50,21	12,34
SVM	54,20	7,25	51,50	25,67
KNN	53,65	5,32	52,36	40,51
LR	54,75	6,98	51,93	32,45
MLP	53,65	4,13	52,79	22,34
MLP OUT	54,20	8,09	52,36	35,67

A Tabela II apresenta os resultados da abordagem proposta aplicada ao ativo WINFUT, considerando intervalos temporais de 30 e 60 minutos. No cenário com granularidade de 30 minutos, o SVM obteve a maior acurácia entre os modelos avaliados (51,50%), seguido de perto por LR, KNN e MLP OUT, que apresentaram desempenhos bastante próximos. Já o SVR e o ARIMA ficaram nas últimas posições em termos de acurácia. Para a granularidade de 60 minutos, o modelo KNN se sobressaiu na tarefa de classificação, com acurácia de 51,07%, superando os demais algoritmos. Na recomendação de investimentos, todos os modelos da abordagem proposta conseguiram superar a estratégia *buy and hold*. Nos experimentos com intervalo temporal de 30 minutos, o melhor resultado foi obtido pelo SVR, seguido por LR, SVM, KNN e ARIMA. O MLP OUT apresentou um desempenho semelhante, porém ligeiramente superior à estratégia *buy and hold*. Já com granularidade de 60 minutos, o KNN foi o destaque, apresentando o melhor resultado. Todos os modelos alcançaram desempenho superior a 2,34% em relação à estratégia *buy and hold*.

Tabela II. Resultados de previsão de direção (subida ou descida) e recomendação de investimentos para WINFUT.

Modelo	30 minutos		60 minutos	
	Acurácia (%)	Var. (%)	Acurácia (%)	Var. (%)
ARIMA	48,35	3,45	50,21	2,34
SVM	51,50	7,65	50,64	4,56
KNN	50,64	6,54	51,07	5,07
LR	51,07	8,76	50,64	4,32
SVR	49,21	9,21	50,21	4,12
MLP OUT	50,64	0,16	50,64	4,56

A Tabela III apresenta os resultados da abordagem proposta aplicada ao ativo WDOFUT, com intervalos de 30 e 60 minutos. No cenário com granularidade de 30 minutos, o modelo KNN obteve o melhor desempenho na tarefa de previsão de direção (50,64%), seguido por SVM e MLP OUT, ambos com acurácias próximas. O modelo SVR ficou logo atrás, enquanto ARIMA e GARCH apresentaram os menores índices de acurácia, embora com desempenhos similares. Com granularidade de 60 minutos, os resultados de classificação foram mais equilibrados, com o KNN novamente alcançando a maior acurácia (50,64%), ligeiramente à frente dos demais modelos. Contudo, na tarefa de recomendação de investimentos, todos os algoritmos apresentaram resultados inferiores à estratégia *buy and hold*, indicando maior dificuldade da abordagem em identificar oportunidades lucrativas nesse cenário. O modelo GARCH foi o que apresentou o menor prejuízo, com variação de apenas -0,88%, destacando-se entre os demais.

Tabela III. Resultados de previsão de direção (subida ou descida) e recomendação de investimentos para WDOFUT.

Modelo	30 minutos		60 minutos	
	Acurácia (%)	Var. (%)	Acurácia (%)	Var. (%)
ARIMA	49,21	3,45	48,78	-1,23
GARCH	49,21	3,78	48,78	-0,88
SVM	50,21	5,67	50,21	-2,345
KNN	50,64	6,54	50,64	-3,45
SVR	49,64	7,11	49,21	-4,56
MLP OUT	50,21	4,233	50,21	-3,06

C. Análise Comparativa

Em 5 dos 6 cenários analisados (83,33%), a máquina de previsão superou a estratégia *buy and hold*, com destaque para o ativo PETR3 na granularidade de 60 minutos, que apresentou um ganho de +40,51% em relação à estratégia de referência. Mesmo com acurácias próximas de 50%, diversos modelos foram capazes de superar significativamente o *buy and hold*, o que indica que conseguiram identificar corretamente momentos cruciais para a tomada de decisão.

Observou-se ainda que os modelos com melhor desempenho em termos de acurácia nem sempre foram os mais eficazes na recomendação de investimentos. Esse resultado reforça a limitação das métricas tradicionais de avaliação — como a acurácia — quando aplicadas a contextos em que o retorno financeiro é o principal objetivo.

Entre os ativos analisados, o PETR3 apresentou os melhores resultados, possivelmente por ser mais previsível do que os contratos futuros de índices. Além disso, a granularidade de 60 minutos mostrou-se mais vantajosa na maioria dos casos,

provavelmente por reduzir os efeitos de ruídos de curtíssimo prazo que dificultam a modelagem de séries temporais no mercado financeiro.

V. CONCLUSÃO

Este trabalho aplicou técnicas de Inteligência Artificial no contexto do mercado financeiro, desenvolvendo uma máquina de previsão que integra modelos estatísticos, de classificação e de regressão por meio de técnicas de ensemble. Os experimentos realizados com três ativos financeiros, em diferentes granularidades temporais, demonstraram a viabilidade da abordagem, com a máquina de previsão superando a estratégia *buy and hold* em 83,33% dos casos avaliados.

Verificou-se, no entanto, que modelos com maior acurácia na previsão de direção nem sempre resultaram nos melhores retornos financeiros. Esse comportamento indica a importância de adotar métricas mais alinhadas aos objetivos de maximização de ganhos. Em relação à granularidade, o intervalo de 60 minutos geralmente produziu resultados superiores, com exceção do ativo WDOFUT, possivelmente em função da menor influência de ruídos de curto prazo nesse horizonte.

Como direções futuras, propõe-se refinar o cálculo do erro do modelo, priorizando classificações corretas com maior impacto financeiro. Sugere-se substituir modelos com baixo desempenho e adotar abordagens evolutivas com aprendizado contínuo. A inclusão de variáveis oriundas de análise de sentimento e dados fundamentalistas pode ampliar a capacidade preditiva. Também se recomenda testar a robustez do modelo em diferentes granularidades temporais. Outras métricas financeiras e os custos operacionais devem ser considerados na avaliação comparativa dos modelos.

AGRADECIMENTOS

Alisson Marques da Silva agradece à Fundação de Amparo à Pesquisa do Estado de Minas Gerais (FAPEMIG) pelo apoio concedido por meio do projeto APQ-03224-24 e ao Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq), processo nº 305452/2025-8.

REFERÊNCIAS

- [1] French, D. (2006). The Dutch monetary environment during tulipmania. *The Quarterly Journal of Austrian Economics*, 9(1), 3-14.
- [2] Dow, J. (1986). The Dow Theory. *Barron's*, 20(28), 1-2.
- [3] Fama, E. F. (1970). Efficient capital markets: A review of theory and empirical work. *The Journal of Finance*, 25(2), 383-417.
- [4] Mutinda, J. K. e Langat, A. K. (2024). Stock price prediction using combined GARCH-AI models. *Scientific African*, 26, e02374.
- [5] Safari, A. e Ghaemi, S. (2025). NeuroFuzzyMan: A hybrid neuro-fuzzy BiLSTM stacked ensemble model for financial forecasting and analysis: Dataset case studies on JPMorgan, AMZN and TSLA. *Expert Systems with Applications*, 266, 126037.
- [6] Wang, J., Wang X. (2021). COVID-19 and financial market efficiency: Evidence from an entropy-based analysis. *Finance Research Letters*, 42, 101888.
- [7] Izzeldin, M., Muradoglu, Y., Pappas, V., Petropoulou, A., Sheeja S. (2023). The impact of the Russian-Ukrainian war on global financial markets. *International Review of Financial Analysis*, 87, 102598.
- [8] Shahid, M. (2022). COVID-19 and adaptive behavior of returns: evidence from commodity markets. *Humanities and Social Sciences Communications*, 9(1).
- [9] Malkiel, B. G. (2003). The efficient market hypothesis and its critics. *Journal of Economic Perspectives*, 17(1), 59-82.

- [10] Shiller, R. J. (2000). Irrational exuberance. Princeton University Press.
- [11] Cajueiro, D. O., Tabak, B. M. (2004). The Hurst exponent over time: testing the assertion that emerging markets are becoming more efficient. *Physica A: Statistical Mechanics and its Applications*, 336(3-4), 521-537.
- [12] Agrawal, J. G., Chourasia, V. S., Mittra, A. K. (2022). Stock market prediction using machine learning algorithms. *International Journal of Advanced Computer Science and Applications*, 13(1), 1-8.
- [13] Vadlamudi, S. (2017). Stock market prediction using machine learning: A systematic literature review. *American Journal of Trade and Policy*, 4(3), 123-128.
- [14] Kaniel, R., Saar, G., Titman, S. (2008). Individual investor trading and stock returns. *The Journal of Finance*, 63(1), 273-310.
- [15] UNCTAD. (2020). World Investment Report 2020: International Production Beyond the Pandemic. United Nations Conference on Trade and Development.
- [16] XP Investimentos. (2023). Relatório de Investidores Pessoa Física na Bolsa de Valores.
- [17] B3. (2023). Histórico de pessoas físicas na Bolsa de Valores. B3 - Brasil, Bolsa, Balcão.
- [18] Gonzalo, J., Pitarakis, J. Y. (2002). Estimation and model selection based inference in single and multiple threshold models. *Journal of Econometrics*, 110(2), 319-352.
- [19] Korczak, J., Hernes, M. (2017). Deep learning for financial time series forecasting in A-Trader system. In *Proceedings of the 2017 Federated Conference on Computer Science and Information Systems* (pp. 905-912).
- [20] Morettin, P. A., Toloi, C. M. (2006). *Análise de séries temporais*. Blucher.
- [21] Bulkowski, T. N. (2012). *Encyclopedia of candlestick charts*. John Wiley & Sons.
- [22] Matsura, E. (2017). *Comprar ou vender? Como investir na bolsa utilizando análise gráfica*. Saraiva Educação SA.
- [23] Igarashi, W., Prado, E. V., Souza, J. P. (2021). Análise de sentimento no mercado de ações: uma revisão sistemática da literatura. *Revista de Gestão da Tecnologia e Sistemas de Informação*, 18, e202118007.
- [24] Halil, F., Mesut, A., Hasan, A. (2019). Predicting the direction of gold prices using artificial neural network approach. *Social Sciences Research Journal*, 8(1), 27-36.
- [25] Sagi, O., Rokach, L. (2018). Ensemble learning: A survey. *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*, 8(4), e1249.
- [26] B3. (2023). Horário de negociação. B3 - Brasil, Bolsa, Balcão.
- [27] Wang, S., Gao, Y., Jing, F., Wang, J. (2013). Soft voting clustering ensemble. In *International Conference on Intelligent Science and Big Data Engineering* (pp. 307-314). Springer, Berlin, Heidelberg.
- [28] Dietterich, T. G. (2002). Ensemble learning. *The handbook of brain theory and neural networks*, 2, 110-125.
- [29] Sezer, O. B., Gudelek, M. U., Ozbayoglu, A. M. (2020). Financial time series forecasting with deep learning: A systematic literature review: 2005–2019. *Applied Soft Computing*, 90, 106181.
- [30] Jiang, Z., Xu, D., Liang, J. (2021). A deep reinforcement learning framework for the financial portfolio management problem. *Expert Systems with Applications*, 167, 113678.
- [31] Chen, L., Qiao, Z., Wang, M., Wang, C., Du, R., Stanley, H. E. (2021). Which artificial intelligence algorithm better predicts the Chinese stock market?. *IEEE Access*, 9, 52177-52191.
- [32] Zhang, G., Liu, Y., Tan, X. (2020). Deep learning approach for financial time series prediction. In *2020 IEEE International Conference on Systems, Man, and Cybernetics (SMC)* (pp. 4082-4089). IEEE.
- [33] Cavalcante, R. C., Brasileiro, R. C., Souza, V. L., Nobrega, J. P., Oliveira, A. L. (2016). Computational intelligence and financial markets: A survey and future directions. *Expert Systems with Applications*, 55, 194-211.
- [34] Patel, J., Shah, S., Thakkar, P., Kotecha, K. (2015). Predicting stock and stock price index movement using trend deterministic data preparation and machine learning techniques. *Expert Systems with Applications*, 42(1), 259-268.
- [35] Krauss, C., Do, X. A., Huck, N. (2017). Deep neural networks, gradient-boosted trees, random forests: Statistical arbitrage on the S&P 500. *European Journal of Operational Research*, 259(2), 689-702.
- [36] Booth, A., Gerding, E., McGroarty, F. (2014). Automated trading with performance weighted random forests and seasonality. *Expert Systems with Applications*, 41(8), 3651-3661.
- [37] Kim, J. H., Shamsuddin, A., Lim, K. P. (2019). Stock return predictability and the adaptive markets hypothesis: Evidence from century-long US data. *Journal of Empirical Finance*, 18(5), 868-879.
- [38] Wang, J., Sun, T., Liu, B., Cao, Y., Wang, D. (2018). Financial markets prediction with deep learning. In *2018 17th IEEE International Conference on Machine Learning and Applications (ICMLA)* (pp. 97-104). IEEE.
- [39] Li, X., Xie, H., Wang, R., Cai, Y., Cao, J., Wang, F., Min, H., Deng, X. (2019). Empirical analysis: Stock market prediction via extreme learning machine. *Neural Computing and Applications*, 27(1), 67-78.