

DESENVOLVIMENTO E COMPARAÇÃO DE MODELOS DE APRENDIZADO DE MÁQUINA APLICADOS À PREDIÇÃO DA TENDÊNCIA DE DESEMPENHO DO IBOVESPA

JOÃO VITOR DE PAIVA MARCOTTI ¹

LEANDRO MATHEUS FOGATE ²

RESUMO

Este trabalho aborda o desenvolvimento de modelos de aprendizado de máquina aplicados à predição da tendência de desempenho do Ibovespa, sendo elaborados por meio da linguagem de programação Python, com uso das bibliotecas Pandas, Numpy, Matplotlib, Seaborn, Scikit-learn, Keras e TensorFlow, e comparados em períodos de 2010, 2015 e 2020 a 2023. Tendo como objetivo lucrar apenas com posições compradas do índice, jamais apostando em sua queda, estabeleceu-se que para um modelo ser considerado satisfatório, sua precisão deve superar o percentual de dias em que o Ibovespa se valorizou no período de teste, vencendo assim o modelo base, que prevê a valorização do índice todos os dias. Quatro experimentos iniciais foram realizados, alimentando os modelos com diferentes variáveis de entrada, como os indicadores de análise técnica K e D estocásticos, média móvel simples e ponderada do valor de fechamento, momento e a razão de Larry William's. Os modelos de Redes Neurais Recorrentes (RNNs), LSTM e GRU, apresentaram as melhores performances médias, tanto de acurácia quanto de precisão. Um quinto experimento foi realizado com base nos resultados dos anteriores, e consistiu na implementação e teste de modelos híbridos LSTM-GRU e GRU-LSTM, alimentados pelas mesmas variáveis de entrada do terceiro experimento. No entanto, a hibridização dos modelos de melhor performance nos experimentos iniciais não superou as suas performances individuais. Para um trabalho futuro, pode-se trabalhar com a previsão da tendência do Ibovespa para um prazo mais longo, como meses ou anos, ao invés de apenas 1 dia.

Palavras-chave: Aprendizado profundo. Estatística. Inteligência artificial. Python. Séries temporais.

¹Engenheiro eletricista pela Universidade Estadual de Maringá – UEM e graduando em ciências econômicas pelo Centro Universitário Cesumar – UniCesumar.

²Mestre em Economia pela Universidade Estadual de Maringá.

1 INTRODUÇÃO

Um modelo preditivo possui como finalidade a antecipação de eventos com base em acontecimentos passados, auxiliando na tomada de decisão e sendo fundamental na solução de problemas complexos (ALVES, 2019).

Segundo Alves (2019), o desenvolvimento de um modelo preditivo se caracteriza pela análise de uma coleção de dados históricos, a fim de definir uma maneira que permita prever um evento futuro de interesse. Desse modo, o grau de exatidão de um modelo preditivo envolve analisar a equivalência entre a previsão realizada por ele e o que de fato aconteceu.

A parte mais importante de um modelo preditivo é a identificação de quais são as melhores variáveis para a definição de um acontecimento ou característica (ALVES, 2019).

Modelos de aprendizado de máquina que empregam Redes Neurais Artificiais (ANN: *Artificial Neural Networks*) são a técnica mais utilizada em algoritmos de previsão, sendo confirmado o fato de que fornece bons desempenhos na previsão do preço de ações. Além disso, algoritmos baseados em Árvores de Decisão (DT: *Decision Trees*) também são muito empregados neste tipo de aplicação, descrevendo muito bem a relação de causa e efeito de informações (TSAI; WANG, 2009).

Redes Neurais Artificiais, são sistemas baseados em uma coleção de unidades ou nós conectados chamados neurônios artificiais, que são inspirados nos neurônios de um cérebro biológico animal. Assim como as sinapses em um cérebro biológico, cada conexão pode transmitir um sinal para outros neurônios. Um neurônio artificial recebe sinais, os processa e então pode sinalizar os neurônios conectados a ele. Normalmente, os neurônios são agregados em camadas, sendo que diferentes camadas podem realizar diferentes transformações em suas entradas. Os sinais viajam da primeira camada (de entrada) para a última camada (de saída), depois de atravessar as camadas intermediárias várias vezes (HARDESTY, 2017).

O modelo mais básico de ANN se trata do Perceptron de Múltiplas Camadas (MLP: *Multilayer Perceptron*). O objetivo deste modelo é aproximar uma função F . Para um classificador, $Y = F(X)$ mapeia uma entrada X para uma categoria de Y , definindo um mapeamento $Y = F(X; \theta)$ e aprendendo os valores do parâmetro θ que

resultam na melhor aproximação da função Y . Este modelo se trata de uma rede neural do tipo *feedforward*, pois a informação flui pelos neurônios das camadas iniciais para as finais, até chegar à saída, sem conexões de *feedback*. Caso houvesse tais conexões, tratar-se-ia de uma Rede Neural Recorrente (RNN: *Recurrent Neural Networks*) (GOODFELLOW; BENGIO; COURVILLE, 2016).

Como mencionado anteriormente, RNN se trata de um tipo de ANN no qual há conexões de *feedback* entre as camadas de neurônios. Esta propriedade permite que redes neurais deste tipo armazenem informações ao processar novas entradas, tornando-as ideais para tarefas nas quais se deve considerar entradas anteriores à presente, sendo, portanto, propícias para aplicações que envolvem uma grande sequência de valores, como uma série temporal. *Long Short Term Memory* (LSTM) e *Gated Recurrent Unit* (GRU) são exemplos de RNNs (GOODFELLOW; BENGIO; COURVILLE, 2016).

Árvores de Decisão se trata de uma técnica que pode ser dividida em dois tipos, variando de acordo com o tipo de valor assumido pela variável de destino: Árvores de Classificação, assumindo valores discretos, e Árvores de Regressão, assumindo valores contínuos (WU et al., 2017). No entanto, segundo Silva e Sátiro (2022), o uso de algoritmos deste tipo têm se tornado cada vez menos comum em razão do surgimento de modelos que combinam múltiplas Árvores de Decisão, sendo conhecidos como modelos de bagging e boosting. De acordo com Nabipour (2020), estes modelos apresentam melhores resultados do que aplicações de Árvores de Decisão individuais.

Modelos de *bagging*, como o *Random Forest*, trabalham com a associação de Árvores de Decisão em paralelo, sendo o resultado final dado pela média dos resultados individuais. Já modelos de boosting, como *Gradient Boosting*, *Light Gradient Boosting* e *X Gradient Boosting*, a associação ocorre em série, sendo o objetivo de cada Árvore de Decisão melhorar o resultado da árvore anterior (SILVA; SÁTIRO, 2022).

Ainda, o método Máquina de Vetores Sequenciais (SVM: *Support Vector Machines*) também pode ser empregado em modelos preditivos. Segundo Géron (2019), tem-se em SVM um modelo muito versátil, capaz de performar em classificações lineares e não lineares, regressões e até mesmo detecção de *outliers*. Este modelo busca traçar hiperplanos de forma a separar os dados de entrada, de

forma que o hiperplano ideal maximiza a distância entre os dois pontos da amostra mais próximos.

Modelos Naive Bayes possuem uma base matemática relativamente simples, sendo fundamentados em princípios de probabilidade estatística por meio da aplicação do teorema de Bayes e da suposição de independência condicional entre suas variáveis de entrada. Entre os tipos de algoritmos Naive Bayes, 3 se destacam: Gaussiano, Bernoulli e Multinomial, sendo nomeados de acordo com a distribuição de dados em que são baseados (BROWNLEE, 2020).

Ainda, segundo Tsai e Wang (2009), temos que a combinação de diferentes técnicas de aprendizado de máquina tem apresentado melhores performances em comparação ao emprego de uma única técnica.

Desta forma, empresas e gestores relacionados à mercados de ações de todo o mundo têm buscado técnicas e ferramentas que possibilitem a predição do desempenho de ativos em um determinado período de tempo e auxiliem na tomada de decisão.

2 MODELOS DE PREDIÇÃO NO MERCADO FINANCEIRO

A predição do valor de ações negociadas em bolsa se trata de um processo desafiador e complexo, em razão da natureza imprevisível de seus movimentos no longo prazo (NABIPOUR et al., 2020).

Segundo Mohri et al. (2018), há duas metodologias de predição do valor de ações: análise fundamentalista e análise técnica. A análise fundamentalista foca mais nas informações contábeis da empresa do que em sua ação negociada em bolsa, sendo a análise baseada no histórico de seus demonstrativos financeiros. Já a análise técnica se baseia nos movimentos passados da ação da empresa, bem como os padrões formados pelo mesmo (MOHRI et al., 2018). Assim, é possível se aplicar técnicas e modelos de *machine learning* à análise técnica de uma ação.

Segundo Obthong et al. (2020), nos últimos anos, muitos pesquisadores têm utilizado algoritmos de aprendizado de máquina para criar ferramentas e analisar dados financeiros históricos para auxiliar na tomada de decisões de investimentos.

Notícias relacionadas ao mercado financeiro e dados coletados de redes sociais foram utilizados por Jeong, Kim e Yoon (2018) em seu modelo para auxiliar na tomada de decisões relacionadas a ações.

Resultados precisos dependem, diretamente, da escolha correta do modelo de aprendizado de máquina empregado e do tipo de informação coletada (OBTHONG et al., 2020).

O Quadro 1, obtido a partir de Obthong et al. (2020), resume o uso de diferentes modelos quantitativos em aplicações de predição de desempenho de ações:

Quadro 1 - Aplicações de diferentes modelos quantitativos na predição do desempenho de ações.

Modelo	Utilizado
Clusterização	
<i>K-Means</i>	X
<i>SOM - Self Organizing Maps</i>	X
Clusterização hierárquica	X
Clusterização	
Predição	
<i>RF - Random Forest</i>	X
<i>SVM - Support Vector Machine</i>	X
<i>MLP - Multilayer Perceptron</i>	X
<i>LSTM - Long Short-Term Memory</i>	X
<i>RNN - Recurrent neural networks</i>	X
<i>GAs - Genetic Algorithms</i>	X
<i>KNN - K Nearest Neighbour</i>	X
<i>SVR - Support Vector Regression</i>	X
<i>MCS - Monte Carlo Simulation</i>	X
<i>ANNs - Artificial Neural Networks</i>	X
<i>CART - Classification and Regression Trees</i>	X
<i>GP - Gaussian Processes</i>	X
<i>BSM - Black Scholes Model</i>	X

<i>GRNN - Generalized Regression Neural Network</i>	X
<i>RBF - Radial Basis Function Neural Networks</i>	
<i>BPNN - Back propagation neural network</i>	X
<i>LR - Logistic Regression</i>	X
<i>HMM - Hidden Markov Model</i>	X
Classificação	
<i>SVM - Support Vector Machine</i>	X
<i>KNN - K Nearest Neighbour</i>	X
<i>LR - Logistic Regression</i>	
<i>ANNs - Artificial Neural Networks</i>	

Fonte: Obthong et al. (2020).

Por sua vez, Kumar et al. (2021) apresenta o percentual de cada modelo de aprendizado de máquina utilizado para a predição do valor de ações presentes em 30 trabalhos selecionados e estudados pelo autor (Quadro 2).

Quadro 2 - Percentual de cada modelo utilizado nos 30 trabalhos estudados pelo autor.

Técnica	Percentual de uso (%)
<i>SVM - Support Vector Machine</i>	21
<i>NN - Neural Network</i>	33
<i>ANN - Artificial neural network</i>	24
<i>CNN - Convolutional neural network</i>	6
<i>RNN - Recurrent neural network</i>	6
<i>SVR - Support vector regression</i>	3
<i>GAN - Generative adversarial network</i>	3
NB - Naïve Bayes	3
Abordagem híbrida	9

Fonte: Kumar et al. (2021).

Dados de séries temporais são grupos de informações contínuas coletadas ao longo do tempo (anualmente, mensalmente, semanalmente, diariamente ou a cada hora, minuto ou segundo). O preço de fechamento de uma ação é um exemplo de dado de série temporal (OBTHONG et al., 2020).

Segundo Obthong et al. (2020), uma série temporal é formada por 4 componentes: tendência, ciclo, sazonalidade e irregularidade.

A tendência mostra a direção de movimento dos dados em um período longo, podendo ser estável, crescente ou decrescente (OBTHONG et al., 2020).

Já o ciclo se trata dos padrões de movimento que a ação apresenta em períodos superiores a 1 ano, sendo afetada, normalmente, por condições relacionadas a ciclos econômicos ou empresariais (OBTHONG et al., 2020).

Por sua vez, a sazonalidade é um movimento dos dados influenciado por um período específico e acontece no mesmo período do ano repetidamente. Pode ocorrer em razão de causas naturais, de negócios, sociais ou culturais (OBTHONG et al., 2020).

Por fim, a irregularidade se trata de movimentos irregulares de curto período presentes na série temporal e podem ser causados por desastres, guerras ou doenças. Costuma afetar as empresas por um período curto de tempo (OBTHONG et al., 2020).

Desta forma, não se tem garantia de que um modelo de predição baseado em dados históricos será preciso, mesmo que apresente boa performance nas etapas de treinamento e teste. Isso ocorre em razão de incertezas que podem acontecer no futuro, eventos esses que dependem de condições políticas e econômicas (OBTHONG et al., 2020). Exemplo: pandemia de Covid-19 e guerra entre Ucrânia e Rússia.

De acordo com Obthong et al. (2020), certos modelos de aprendizado de máquina tem apresentado boa performance ao serem aplicados para predição do valor de ativos, com baixo percentual de erros. São eles: ANN, RNN, LSTM, SLSTM e BLSTM. Já no trabalho de Bluvol (2022), modelos ARIMA performaram melhor do que modelos de redes neurais artificiais.

3 PROCEDIMENTOS METODOLÓGICOS

Este trabalho desenvolveu e comparou a performance de múltiplos modelos de aprendizado de máquina utilizados na previsão da tendência de variação do Ibovespa no dia seguinte, classificando-a como valorização (1) ou desvalorização (0), em 5 experimentos.

No primeiro experimento realizado, utilizou-se como variáveis explicativas a razão de fechamento e a tendência nos últimos 3, 5 e 10 dias úteis dos índices Ibovespa, S&P500, Gold, Euronext 100 e SSE Composite - Shanghai, bonds de 5 anos do tesouro americano e a cotação das moedas dólar e euro em relação ao real.

No segundo experimento realizado, calculou-se a relação linear entre cada uma das variáveis explicativas do primeiro experimento e o valor do Ibovespa no dia seguinte, alimentando os modelos apenas com as 5 maiores correlações positivas e 5 menores correlações negativas.

No terceiro experimento, utilizou-se como variáveis explicativas os seguintes indicadores de análise técnica: média móvel simples do valor de fechamento, média móvel ponderada do valor de fechamento, momento, K estocástico, D estocástico, RSI, MACD, linha de sinal, razão de Larry Williams e CCI, calculados em relação aos valores de fechamento do Ibovespa de 3, 5 e 10 dias úteis anteriores e por meio das equações propostas por Kara, Boyacioglu e Baykan (2011), exibidas na Figura 1. Os valores obtidos estão dispostos de forma contínua ao longo do tempo.

Figura 1 - Indicadores de análise técnica utilizados

Name of indicators	Formulas
Simple 10-day moving average	$\frac{C_t + C_{t-1} + \dots + C_{t-10}}{10}$
Weighted 10-day moving average	$\frac{((n) \times C_t + (n-1) \times C_{t-1} + \dots + C_{t-10})}{(n + (n-1) + \dots + 1)}$
Momentum	$C_t - C_{t-n}$
Stochastic K%	$\frac{C_t - LL_{t-n}}{HH_{t-n} - LL_{t-n}} \times 100$
Stochastic D%	$\frac{\sum_{i=0}^{n-1} K_{t-i} \%}{n}$
RSI (Relative Strength Index)	$100 - \frac{100}{1 + (\sum_{i=0}^{n-1} Up_{t-i}/n) / (\sum_{i=0}^{n-1} Dw_{t-i}/n)}$
MACD (moving average convergence divergence)	$MACD(n)_{t-1} + 2/n + 1 \times (DIFF_t - MACD(n)_{t-1})$
Larry William's R%	$\frac{H_n - C_t}{H_n - L_n} \times 100$
A/D (Accumulation/Distribution) Oscillator	$\frac{H_t - C_{t-1}}{H_t - L_t}$
CCI (Commodity Channel Index)	$\frac{M_t - SM_t}{0.015D_t}$

C_t is the closing price, L_t the low price, H_t the high price at time t , $DIFF_t$: $EMA(12)_t - EMA(26)_t$, EMA exponential moving average, $EMA(k)_t$: $EMA(k)_{t-1} + \alpha \times (C_t - EMA(k)_{t-1})$, α smoothing factor: $2/1 + k$, k is time period of k day exponential moving average, LL_t and HH_t mean lowest low and highest high in the last t days, respectively, M_t : $H_t + L_t + C_t/3$; SM_t : $(\sum_{i=1}^n M_{t-i+1})/n$, D_t : $(\sum_{i=1}^n |M_{t-i+1} - SM_t|)/n$, Up_t means the upward price change, Dw_t means the downward price change at time t .

Fonte: Kara, Boyacioglu e Baykan (2011).

No quarto experimento, utilizou-se os mesmos indicadores de análise técnica do terceiro experimento, porém transformados para a forma determinística, indicando a tendência de aumento ou diminuição do Ibovespa em 1 e 0 respectivamente.

A transformação dos valores da forma contínua para a determinística foi feita seguindo as instruções de Patel, Shah, Thakkar e Kotecha (2015), de forma que 1 indica uma tendência de alta e 0 de queda:

- Média móvel simples = 1 se o valor de fechamento for superior à média móvel simples dos N dias passados. Caso seja menor ou igual, iguale-se a 0.
- Média móvel ponderada = 1 se o valor de fechamento for superior à média móvel ponderada dos N dias passados. Caso seja menor ou igual, iguale-se a 0.
- K estocástico = 1 se o valor do K estocástico do dia N for maior do que do dia N-1. Caso seja menor ou igual, iguale-se a 0.
- D estocástico = 1 se o valor do D estocástico do dia N for maior do que do dia N-1. Caso seja menor ou igual, iguale-se a 0.
- Razão de Larry William's = 1 se o valor da razão de Larry Williams do dia N for maior do que do dia N-1. Caso seja menor ou igual, iguale-se a 0.
- Convergência e divergência da média móvel (MACD) = 1 se o valor da MACD do dia N for maior do que do dia N-1. Caso seja menor ou igual, iguale-se a 0.
- Índice de força relativa (RSI) = 1 se a RSI do dia N for superior a 70 ou menor do que a RSI do dia N-1. Caso contrário, iguale-se a 0.
- Índice do canal de commodities (CCI) = 1 se o CCI do dia N for superior a 200 ou menor do que o CCI do dia N-1. Caso contrário, iguale-se a 0.
- Momento = 1 se o momento do dia N for positivo. Caso seja 0 ou negativo, iguale-se a 0.

No quinto experimento, elaborou-se modelos híbridos LSTM-GRU e GRU-LSTM, alimentados com as mesmas variáveis de entrada do terceiro experimento, a fim de verificar se estes modelos apresentam melhor performance do que os anteriores.

Ainda, a fim de averiguar o funcionamento dos modelos em diferentes períodos de tempo, os 5 experimentos foram executados utilizando-se 3 períodos de tempo: de 2010, 2015 e 2020 até 2023.

Assim, desenvolveu-se e testou modelos de aprendizado de máquina dos tipos *Gaussian Naive-Bayes* (GNB), *Bernoulli Naive-Bayes* (BNB), *Random Forest* (RF), *Support Vector Machine* (SVM), *Gradient Boosting* (GB), *X Gradient Boosting* (XGB), *Light Gradient Boosting* (LGB), *Multilayer Perceptron* (MLP), *Long Short Term Memory* (LSTM) e *Gated Recurrent Unit* (GRU), além de híbridos LSTM-GRU e GRU-LSTM.

Os parâmetros dos modelos, com exceção de LSTM e GRU, foram escolhidos por meio de parametrização em pesquisa de grade (*grid search*). Os outros 2 modelos tiveram seus parâmetros escolhidos por experimentação.

Tendo como objetivo lucrar apenas com posições compradas do índice, jamais apostando em sua queda, como base de comparação, além da acurácia (Equação 1), utilizou-se a medida de precisão (Equação 2), que mede a probabilidade de um modelo acertar ao prever que o Ibovespa se valorizará no dia seguinte (JUNIOR, 2021).

$$Acurácia = \frac{VP+VN}{VP+VN+FP+FN} = \frac{Acertos}{Acertos + Erros} \quad (1)$$

$$Precisão = \frac{VP}{VP+FP} \quad (2)$$

Para um modelo ser considerado satisfatório, sua precisão deve superar o percentual de dias em que o Ibovespa se valorizou no período de teste, vencendo assim, um modelo que prevê a valorização do índice todos os dias, sendo este o modelo base de comparação dos experimentos.

A coleta dos dados ocorreu a partir do site de finanças e investimentos Yahoo Finance, por meio da biblioteca Yfinance. Importou-se dados a respeito do valor diário máximo, mínimo e de fechamento do Ibovespa entre janeiro de 2010 e setembro de 2023.

O algoritmo foi implementado com uso da linguagem de programação Python, por meio do ambiente Google Colab. Para a manipulação de dados, utilizou-se a biblioteca Pandas. Seaborn e Matplotlib foram utilizadas para a exibição das séries

históricas dos índices e de um mapa de calor da correlação linear entre eles. Numpy foi aplicada para operações matemáticas. E Scikit-learn, Keras e TensorFlow possibilitaram a elaboração dos modelos de aprendizado de máquina.

Desta forma, este trabalho se trata de uma pesquisa exploratória e quantitativa, pois busca quantificar, por meio de modelos estatísticos, o comportamento de uma amostra, neste caso o índice da bolsa de valores brasileira, o Ibovespa.

4 ANÁLISE DOS DADOS E INTERPRETAÇÃO DOS RESULTADOS

Quatro experimentos iniciais foram realizados a fim de comparar a performance dos modelos de aprendizado de máquina desenvolvidos, quando alimentados por diferentes variáveis e em diferentes períodos de tempo (início em 2010, 2015 e 2020).

4.1 PRIMEIRO EXPERIMENTO

Quadro 3 - Acurácia dos modelos no primeiro experimento.

Modelo	2010	2015	2020	Média
Bernoulli Naive Bayes	0,4778	0,4754	0,5059	0,4864
Gated Recurrent Unit	0,5109	0,5205	0,5378	0,5231
Gaussian Naive Bayes	0,4886	0,5042	0,5059	0,4996
Gradient Boosting	0,5157	0,4975	0,4743	0,4958
Light Gradient Boosting	0,5244	0,5059	0,4585	0,4963
Logistic Regression	0,4962	0,5178	0,4743	0,4961
Long-Short Term Memory	0,5228	0,5256	0,5458	0,5314
Multilayer Perceptron	0,5027	0,5178	0,4901	0,5035
Random Forest	0,4973	0,5144	0,4704	0,4940
Support Vector Machine	0,5244	0,5178	0,4901	0,5108
X Gradient Boosting	0,4930	0,4805	0,4901	0,4879

Fonte: Autoria própria.

Quadro 4 - Precisão dos modelos no primeiro experimento.

Modelo	2010	2015	2020	Média
Bernoulli Naive Bayes	0,5023	0,4937	0,4959	0,4973
Gated Recurrent Unit	0,5533	0,5337	0,5274	0,5381
Gaussian Naive Bayes	0,5186	0,5355	0,4815	0,5119
Gradient Boosting	0,5391	0,5098	0,4651	0,5047
Light Gradient Boosting	0,5484	0,5180	0,4620	0,5095
Logistic Regression	0,5195	0,5296	0,4685	0,5059
Long-Short Term Memory	0,5460	0,5293	0,5357	0,5370
Multilayer Perceptron	0,5239	0,5178	0,4901	0,5106
Random Forest	0,5209	0,5241	0,4742	0,5064
Support Vector Machine	0,5244	0,5178	0,4901	0,5108
X Gradient Boosting	0,5177	0,4987	0,4793	0,4986
Modelo base	0,5244	0,5178	0,4901	0,5108

Fonte: Autoria própria.

Com base nos resultados obtidos no primeiro experimento, observa-se que a acurácia de 7 modelos caiu com a redução do período de dados, enquanto que em apenas 4 ela subiu, entre eles os 2 modelos de maior acurácia média: GRU e LSTM, com 0,5231 e 0,5314 respectivamente.

Quanto à precisão, observa-se que todos os modelos tiveram uma perda de performance com a redução do período de dados. Os modelos de melhor performance média foram, novamente, GRU e LSTM com 0,5381 e 0,5370 respectivamente, além do Naive Bayes Gaussiano que também superou o modelo base médio de 0,5108.

4.2 SEGUNDO EXPERIMENTO

Quadro 5 - Acurácia dos modelos no segundo experimento.

Modelo	2010	2015	2020	Média
Bernoulli Naive Bayes	0,4826	0,4813	0,5099	0,4913
Gated Recurrent Unit	0,5321	0,5350	0,5480	0,5384

Gaussian Naive Bayes	0,4794	0,4898	0,5336	0,5009
Gradient Boosting	0,4913	0,5204	0,4625	0,4914
Light Gradient Boosting	0,4935	0,5272	0,5138	0,5115
Logistic Regression	0,5054	0,5051	0,4743	0,4949
Long-Short Term Memory	0,5288	0,5214	0,5200	0,5234
Multilayer Perceptron	0,4967	0,5051	0,5059	0,5026
Random Forest	0,4913	0,5153	0,4822	0,4963
Support Vector Machine	0,4946	0,5136	0,4901	0,4994
X Gradient Boosting	0,4816	0,4949	0,5099	0,4955

Fonte: Autoria própria.

Quadro 6 - Precisão dos modelos no segundo experimento.

Modelo	2010	2015	2020	Média
Bernoulli Naive Bayes	0,5064	0,4956	0,5000	0,5007
Gated Recurrent Unit	0,5494	0,5592	0,5542	0,5543
Gaussian Naive Bayes	0,5029	0,5026	0,5455	0,5170
Gradient Boosting	0,5128	0,5212	0,4483	0,4941
Light Gradient Boosting	0,5165	0,5343	0,5039	0,5182
Logistic Regression	0,5288	0,5138	0,4767	0,5064
Long-Short Term Memory	0,5617	0,5312	0,5167	0,5365
Multilayer Perceptron	0,5200	0,5139	0,4972	0,5104
Random Forest	0,5161	0,5228	0,4795	0,5061
Support Vector Machine	0,5198	0,5136	0,4901	0,5078
X Gradient Boosting	0,5056	0,5079	0,5000	0,5045
Modelo base	0,5239	0,5136	0,4901	0,5092

Fonte: Autoria própria.

Assim como no primeiro experimento, os modelos do tipo LSTM e GRU se destacaram frente aos demais, com acurácias médias iguais a 0,5234 e 0,5384, seguidos por LGB com 0,5115.

Quanto à precisão média, 5 modelos superaram o modelo base de 0,5092: GRU, LSTM, LGB, GNB e MLP, com destaque para o primeiro com 0,5543, uma melhora em relação ao primeiro experimento.

4.3 TERCEIRO EXPERIMENTO

Quadro 7 - Acurácia dos modelos no terceiro experimento.

Modelo	2010	2015	2020	Média
Bernoulli Naive Bayes	0,5015	0,4923	0,4893	0,4944
Gated Recurrent Unit	0,4892	0,5317	0,5487	0,5232
Gaussian Naive Bayes	0,5083	0,4938	0,5036	0,5019
Gradient Boosting	0,5034	0,4923	0,4464	0,4807
Light Gradient Boosting	0,5152	0,4938	0,4429	0,4840
Logistic Regression	0,5083	0,5031	0,4679	0,4931
Long-Short Term Memory	0,4921	0,5085	0,5596	0,5201
Multilayer Perceptron	0,5113	0,5046	0,5393	0,5184
Random Forest	0,5328	0,5138	0,5357	0,5274
Support Vector Machine	0,5113	0,5046	0,4786	0,4982
X Gradient Boosting	0,5348	0,4862	0,4750	0,4987

Fonte: Autoria própria.

Quadro 8 - Precisão dos modelos no terceiro experimento.

Modelo	2010	2015	2020	Média
Bernoulli Naive Bayes	0,5134	0,4974	0,4740	0,4949
Gated Recurrent Unit	0,5556	0,6154	0,5833	0,5848
Gaussian Naive Bayes	0,5099	0,4983	0,4675	0,4919
Gradient Boosting	0,5333	0,4979	0,4266	0,4859
Light Gradient Boosting	0,5248	0,4991	0,4247	0,4829
Logistic Regression	0,5130	0,5047	0,4590	0,4922
Long-Short Term Memory	0,6190	0,5228	0,6571	0,5996
Multilayer Perceptron	0,5113	0,5046	0,5373	0,5177

Random Forest	0,5315	0,5166	0,5130	0,5204
Support Vector Machine	0,5113	0,5046	0,4786	0,4982
X Gradient Boosting	0,5477	0,4927	0,4532	0,4979
Modelo base	0,5113	0,5046	0,4786	0,4982

Fonte: Autoria própria.

No terceiro experimento, quanto à acurácia, observou-se a melhora do modelo do tipo RF que, além de superar o modelo base pela primeira vez, apresentou acurácia superior aos demais modelos.

Quanto à precisão, nota-se o grande destaque dos modelos LSTM e GRU em relação aos demais, com precisões médias iguais a 0,5996 e 0,5848 respectivamente.

4.4 QUARTO EXPERIMENTO

Quadro 9 - Acurácia dos modelos no quarto experimento.

Modelo	2010	2015	2020	Média
Bernoulli Naive Bayes	0,5059	0,4946	0,4875	0,4960
Gated Recurrent Unit	0,4912	0,5116	0,5254	0,5094
Gaussian Naive Bayes	0,5069	0,4915	0,4910	0,4965
Gradient Boosting	0,5088	0,4961	0,4659	0,4903
Light Gradient Boosting	0,5216	0,4961	0,5448	0,5208
Logistic Regression	0,5157	0,5039	0,4982	0,5059
Long-Short Term Memory	0,4951	0,5209	0,4964	0,5041
Multilayer Perceptron	0,4980	0,5023	0,5125	0,5043
Random Forest	0,5176	0,4761	0,5054	0,4997
Support Vector Machine	0,5265	0,4900	0,4624	0,4930
X Gradient Boosting	0,5029	0,4961	0,5090	0,5027

Fonte: Autoria própria.

Quadro 10 - Precisão dos modelos no quarto experimento.

Modelo	2010	2015	2020	Média
Bernoulli Naive Bayes	0,5179	0,5000	0,4694	0,4958
Gated Recurrent Unit	0,5345	0,5135	0,5070	0,5183
Gaussian Naive Bayes	0,5190	0,4971	0,4726	0,4962
Gradient Boosting	0,5169	0,5010	0,4615	0,4931
Light Gradient Boosting	0,5294	0,5010	0,5185	0,5163
Logistic Regression	0,5245	0,5071	0,4833	0,5050
Long-Short Term Memory	0,5870	0,5188	0,4624	0,5227
Multilayer Perceptron	0,5092	0,5062	0,4948	0,5034
Random Forest	0,5229	0,4830	0,4891	0,4983
Support Vector Machine	0,5334	0,4961	0,4474	0,4923
X Gradient Boosting	0,5120	0,5013	0,4915	0,5016
Modelo base	0,5118	0,5054	0,4803	0,4992

Fonte: Autoria própria.

Neste experimento, primeiramente, observou-se a redução do tempo de execução da parametrização, treino e teste dos modelos em relação ao terceiro experimento.

Quanto à acurácia dos modelos, nota-se uma perda de performance em relação ao experimento anterior, principalmente dos modelos RF, GRU e LSTM. O mesmo ocorre para a precisão.

4.5 QUINTO EXPERIMENTO

Com a análise dos quatro experimentos iniciais, notou-se, no terceiro, a discrepante performance dos modelos LSTM e GRU, tanto quanto à acurácia quanto à precisão, em relação aos demais e em relação a eles mesmos nos outros 3 experimentos. Desta forma, decidiu-se elaborar modelos híbridos LSTM-GRU e GRU-LSTM, aplicá-los nos dados do terceiro experimento e averiguar as suas performances de acurácia (Quadro 11) e precisão (Quadro 12).

Quadro 11 - Acurácia dos modelos híbridos no quinto experimento.

Modelo	2010	2015	2020	Média
LSTM-GRU	0,5231	0,5077	0,5271	0,5193
GRU-LSTM	0,4956	0,5123	0,5271	0,5117
LSTM	0,4921	0,5085	0,5596	0,5201
GRU	0,4892	0,5317	0,5487	0,5232

Fonte: Autoria própria.

Quadro 12 - Precisão dos modelos híbridos no quinto experimento.

Modelo	2010	2015	2020	Média
LSTM-GRU	0,5779	0,5144	0,6250	0,5724
GRU-LSTM	0,5352	0,6000	0,5333	0,5562
LSTM	0,6190	0,5228	0,6571	0,5996
GRU	0,5556	0,6154	0,5833	0,5848

Fonte: Autoria própria.

Com base nos resultados obtidos, nota-se que, infelizmente, a hibridização dos modelos LSTM e GRU não apresentou melhoras, possuindo um desempenho inferior aos modelos aplicados individualmente tanto em relação à acurácia quanto à precisão. Ainda, tem-se que a hibridização LSTM-GRU apresentou melhores resultados do que GRU-LSTM.

5 CONSIDERAÇÕES FINAIS

Com a realização deste trabalho, cumpriu-se o objetivo estabelecido inicialmente: o desenvolvimento e comparação de múltiplos modelos de aprendizado de máquina para a predição da tendência de desempenho do Ibovespa.

Como resultado final, constatou-se que os modelos que melhor performaram foram os baseados em Redes Neurais Recorrentes (RNNs), LSTM e GRU, com resultados próximos e alternando qual apresentava a melhor performance nos experimentos realizados. Também foi proposta a hibridização destes dois modelos,

no entanto, os seus resultados de acurácia e precisão foram inferiores aos dos modelos aplicados individualmente.

Para trabalhos futuros, há a possibilidade de testar outras variáveis de entrada, períodos de tempo diferentes, modelos de aprendizado de máquina mais avançados, como as Redes Neurais Convolucionais (CNN), ou outros modelos híbridos. Além disso, pode-se trabalhar com a previsão da tendência do Ibovespa para um prazo mais longo, como meses ou anos, ao invés de apenas 1 dia.

REFERÊNCIAS

ALVES, J. V. S. **Um modelo preditivo de cotação de ações de empresas estatais brasileiras utilizando redes neurais artificiais no ambiente MATLAB**. 2019. 138 f. Dissertação (Mestrado) - Curso de Gestão Organizacional, Faculdade de Gestão e Negócios, Universidade Federal de Uberlândia, Uberlândia, 2019.

BLUVOL, Leonardo Mizrahy. **Análise de algoritmos de machine learning e redes neurais para previsão de preços de ações do Ibovespa**. 2022. 44 f. Dissertação (Mestrado) - Curso de Mestrado Profissional em Economia Empresarial e Finanças, Escola Brasileira de Economia e Finanças, Fundação Getúlio Vargas, Rio de Janeiro, 2022.

BROWNLEE, Jason. **A Gentle Introduction to SARIMA for Time Series Forecasting in Python**. 2018. Disponível em: <https://machinelearningmastery.com/sarima-for-time-series-forecasting-in-python/>. Acesso em: 24 ago. 2023.

BROWNLEE, Jason. **How to Develop a Naive Bayes Classifier from Scratch in Python**. 2020. Disponível em: <https://machinelearningmastery.com/classification-as-conditional-probability-and-the-naive-bayes-algorithm/>. Acesso em: 27 set. 2023.

GÉRON, Aurélien. **Hands-on Machine Learning with Scikit-Learn, Keras, and TensorFlow: concepts, tools, and techniques to build intelligent systems**. 2. ed. Sebastopol: O'reilly Media, 2019. 279 p.

GOODFELLOW, Ian; BENGIO, Yoshua; COURVILLE, Aaron. **Deep Learning**. Cambridge: The MIT Press, 2016. 801 p.

HARDESTY, L. **Explained: Neural networks**: Ballyhooed artificial-intelligence technique known as “deep learning” revives 70-year-old idea. 2017. Disponível em: <https://news.mit.edu/2017/explained-neural-networks-deep-learning-0414>. Acesso em: 25 nov. 2022.

JEONG, Yujin; KIM, Sunhye; YOON, Byungun. An Algorithm for Supporting Decision Making in Stock Investment through Opinion Mining and Machine Learning. **2018 Portland International Conference On Management Of Engineering And Technology (PICMET)**, [S.L.], p. 1-10, ago. 2018. IEEE.

KARA, Yakup; BOYACIOGLU, Melek Acar; BAYKAN, Ömer Kaan. Predicting direction of stock price index movement using artificial neural networks and support vector machines: the sample of the istanbul stock exchange. **Expert Systems With Applications**, [S.L.], v. 38, n. 5, p. 5311-5319, maio 2011. Elsevier BV. <http://dx.doi.org/10.1016/j.eswa.2010.10.027>.

KUMAR, Deepak et al. A systematic review of stock market prediction using machine learning and statistical techniques. **Materials Today: Proceedings**, [S.L.], v. 49, n. 8, p. 3187-3191, jan. 2021. Elsevier BV.

MOHRI, Mehryar et al. **Foundations of Machine Learning**. 2. ed. Cambridge: The Mit Press, 2018. 505 p.

NABIPOUR, M. et al. Deep Learning for Stock Market Prediction. **Entropy**, [S.L.], v. 22, n. 8, p. 840, 30 jul. 2020. MDPI AG.

OBTHONG, Mehtabhorn; TANTISANTIWONG, Nongnuch; JEAMWATTHANACHAI, Watthanasak; WILLS, Gary. A Survey on Machine Learning for Stock Price Prediction: algorithms and techniques. **Proceedings Of The 2Nd International Conference On Finance, Economics, Management And It Business**, [S.L.], p. 1-12, 2020. SCITEPRESS - Science and Technology Publications.

OLIVEIRA JÚNIOR, Clébio de. **Indo Além da Acurácia: Entendo a Acurácia Balanceada, Precisão, Recall e F1 score**. 2021. Disponível em: <https://medium.com/data-hackers/indo-al%C3%A9m-da-acur%C3%A1cia-entendo-a-acur%C3%A1cia-balanceada-precis%C3%A3o-recall-e-f1-score-c895e55a9753>. Acesso em: 1 out. 2023.

PATEL, Jigar; SHAH, Sahil; THAKKAR, Priyank; KOTTECHA, K. Predicting stock and stock price index movement using Trend Deterministic Data Preparation and machine learning techniques. **Expert Systems With Applications**, [S.L.], v. 42, n. 1, p. 259-268, jan. 2015. Elsevier BV. <http://dx.doi.org/10.1016/j.eswa.2014.07.040>.

SILVA, José Erasmo; SÁTIRO, Renato Máximo. O poder preditivo dos modelos boosting de machine learning no mercado brasileiro de ações. **Brazilian Journal Of Quantitative Methods Applied To Accounting**. Monte Carmelo, p. 52-68. out. 2022.

TSAL, C F; WANG, S P. Stock Price Forecasting by Hybrid Machine Learning Techniques. In: INTERNATIONAL MULTICONFERENCE OF ENGINEERS AND COMPUTER SCIENTISTS, 2009, Hong Kong. **Proceedings [...]**. Hong Kong: Imecs, 2009. p. 755-760.

WU, X. et al. Top 10 algorithms in data mining. **Knowledge And Information Systems**, [S.L.], v. 14, n. 1, p. 1-37, 4 dez. 2007. Springer Science and Business Media LLC.