



ConBRepro

XIII CONGRESSO BRASILEIRO DE ENGENHARIA DE PRODUÇÃO



IA nas Engenharias

29 nov. a 01
de dezembro 2023

Previsão de Preços por Meio de Dados Textuais: Uma Síntese da Literatura

Fernanda Krause
PPGEM - UFPR

Rúbia Maria Pereira
PPGMNE – UFPR

Cassius Tadeu Scarpin
PPGMNE - UFPR

Pablo Deivid Valle
PPGEM - UFPR

Resumo: A previsão automatizada de preços é um assunto já abordado por pesquisadores durante vários anos. No cenário atual, em que volumes massivos de dados estão disponíveis em tempo real e a interconexão global é mais profunda do que nunca, a capacidade de antecipar movimentos de preços tornou-se uma busca incessante para investidores e profissionais financeiros. Esse artigo apresenta uma revisão bibliográfica sobre esse assunto tendo analisado 40 artigos obtidos através da base de dados Scopus, com o objetivo de analisar a viabilidade de utilizar dados textuais de notícias online ou redes sociais para realizar a previsão de preços de forma geral. Após a aplicação da metodologia analítica ficou evidente que o subtema mais abordado na literatura é a precificação das ações, seguido do preço do petróleo. Os resultados do estudo indicam que o campo de previsão de preços com base em dados textuais é altamente promissor. Em todos os trabalhos analisados, a incorporação de textos de notícias ou dados de redes sociais demonstrou consistentemente melhorar o desempenho dos modelos de previsão, permitindo uma compreensão mais abrangente dos eventos oscilatórios nos preços dos ativos.

Palavras-chave: previsão, preço, revisão de literatura, texto, análise de sentimentos.

Predicting Prices through Textual Data: A Literature Synthesis

Abstract: The automated price forecasting has been a subject of research for several years. In the current landscape, where massive volumes of real-time data are available, and global interconnectivity is deeper than ever, the ability to anticipate price movements has become an ongoing quest for investors and financial professionals. This article presents a bibliographic review on this subject, having analyzed 40 articles obtained through the Scopus database, with the aim of assessing the feasibility of using textual data from online news or social media to perform price prediction in a general sense. Following the application of the analytical methodology, it became evident that the most frequently addressed subtopics in the literature are stock pricing, followed by oil prices. The results of the study indicate that the field of price forecasting based on textual data is highly promising. In all the analyzed works, the incorporation of news texts or social media data consistently demonstrated an improvement in the performance of prediction models, enabling a more comprehensive understanding of oscillatory events in asset prices.

Keywords: forecast, price, literature review, text, sentiment analysis.

1. Introdução

A previsão acurada dos preços, tornou-se crucial devido ao crescimento econômico e à competição entre empresas, fornecendo uma medida eficaz para a gestão empresarial e refletindo a situação econômica indiretamente (ZHANG et al., 2015). Todavia, desenvolver um modelo acurado de previsão de preço continua sendo uma questão desafiadora, principalmente com relação ao sentimento social gerado por meio de notícias e textos em redes sociais (KANSAL; KUMAR, 2019).

Atualmente, há muitos meios de obtenção de notícias relacionadas a diversos assuntos (GRUNDAMANN; FELDEN; POSPIECH, 2016). Isso se deve ao avanço das tecnologias atuais e, conseqüentemente, uma vasta quantidade de dados tornaram-se acessíveis para análise, como: 1) redes sociais; 2) notícias online; e 3) relatórios analíticos (Li et al., 2022).

No entanto, assim como na vida cotidiana, os preços podem ser altamente influenciados por notícias relevantes em várias áreas de negócios (CHAN; CHUI; KWOK, 2001). Sendo assim, a previsão de preço integrada a análise de notícias tem sido estudada de forma crescente pelos pesquisadores (Li et al., 2022).

Nesse contexto, o objetivo desta revisão sistemática é analisar a viabilidade de utilizar dados textuais de notícias online ou redes sociais para realizar a previsão de preços de forma geral. O estudo espera responder se variáveis textuais podem auxiliar nas previsões e quais oportunidades de trabalhos futuros estão disponíveis.

Para atingir este objetivo, o estudo propõe uma revisão sistemática de literatura utilizando a base de dados da Scopus como fonte de informações.

O artigo está dividido em 4 seções, 3 outras além desta; a segunda, chamada de Metodologia, contém a descrição da coleta de dados e os métodos de avaliação escolhidos; a terceira contém a maior densidade de conteúdo apresentando todos os aspectos quantitativos e qualitativos da revisão da literatura assim como seus resultados; e, por fim, a quarta seção contém as conclusões e sugestões de pesquisas futuras baseadas na análise realizada.

2. Metodologia

A metodologia utilizada no presente trabalho é chamada de Revisão Sistemática da Literatura (RSL). A RSL é uma abordagem de pesquisa recomendada para identificar, avaliar e interpretar todas as pesquisas pertinentes disponíveis relacionadas a uma questão de pesquisa específica, uma área temática ou um fenômeno de interesse.

Sendo assim, neste estudo foi utilizado a metodologia proposta por Kitchenham (2004). Essa metodologia consiste em três etapas principais, sendo elas, o planejamento, a condução e a apresentação dos resultados da revisão.

A fase inicial deve apresentar questões que serão respondidas ao final da revisão sistemática, juntamente com os critérios que determinarão quais estudos serão incluídos, excluídos. Desta forma, objetivo desta RSL é analisar a viabilidade de utilizar dados textuais de notícias online ou redes sociais para realizar a previsão de preços de forma geral. Foram incluídos artigos que abordam análises de texto relacionadas a previsões de preços e excluídos artigos que não haviam sido publicados em língua inglesa.

A segunda fase da pesquisa consiste na extração das publicações relevantes para a revisão sistemática e a escolha dos trabalhos de acordo com os critérios já estabelecidos

previamente. Desta forma, para a coleta dos trabalhos que foram alvo dessa revisão sistemática, utilizou-se a base Scopus.

As palavras-chaves escolhidas para a busca foram: “*price*” and “*forecast*” and “*text*” and “*analysis*”, sendo estas analisadas nos títulos, resumos e palavras-chaves dos artigos. Não foi utilizado nenhum filtro temporal, pois um dos pontos de análise desejados foi avaliar como as publicações sobre esse assunto cresceram ao longo dos anos. A busca inicial retornou 102 documentos relacionados com o tema. Todos os artigos haviam sido escritos em língua inglesa, desta forma, não houve a necessidade de exclusão por esse motivo.

A fase final da pesquisa consiste em apresentar os dados da pesquisa realiza. Os mesmos estão descritos no próximo tópico.

3. Resultados

Para realizar a análise quantitativa do trabalho foi utilizado a linguagem de programação R e a ferramenta bibliometrix. A Figura 1, apresenta um resumo geral das publicações utilizadas.

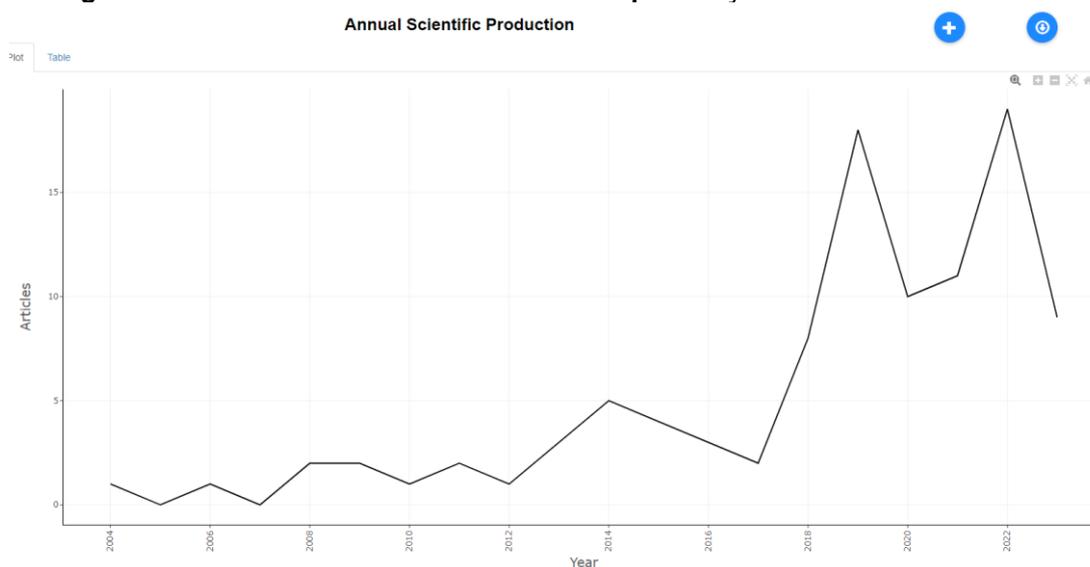
Figura 1 – Resumo geral das publicações utilizadas.



Fonte: Os autores (2023).

Com base na Figura 1, identificamos um total de 102 artigos a serem analisados, abrangendo o período entre os anos de 2004 e 2023. Observa-se ainda que o índice de crescimento anual das publicações desses documentos se situou em aproximadamente 12%, indicando que, ao longo do tempo, esse tema tem sido objeto de pesquisa crescente por parte dos acadêmicos. Este fato pode ser observado na Figura 2, que apresenta o total de publicação dos artigos por ano.

Figura 2 – Gráfico mostrando a tendência de publicações anuais sobre o tema.



Fonte: Os autores (2023).

Na Figura 2, observa-se que a partir do ano de 2019, houve um aumento significativo de publicações sobre o tema em questão, chegando a um aumento de 225% de publicações

em relação ao ano anterior. Isto pode ser explicado pelo crescimento expressivo da utilização de técnicas de Inteligência Artificial na área financeira nos anos anteriores (CHEN *et al.*, 2023). Ao analisar a quantidade de publicações nos últimos cinco anos, observa-se uma média de 13,6 publicações por ano, o que representa um aumento de mais de 500% em comparação com a média dos anos anteriores.

Figura 3 – Nuvem de palavras obtida.



Fonte: Os autores (2023).

Na Figura 3 tem-se a nuvem de palavras obtidas a partir dos textos. A palavra com maior destaque, como já era esperado, é “forecasting” indicando que os artigos analisados representam algum tipo de previsão. Entretanto, duas expressões chamam atenção nessa análise: *long short-term memory* e *sentiment analysis*. Essas expressões indicam dois métodos muito comuns para previsões utilizando informações textuais. A primeira representa uma técnica que tem o papel de evitar problemas inerentes às redes neurais para previsão de séries temporais (An *et al.*, 2023). A segunda uma técnica de análise textual utilizada em 80% dos textos analisados neste trabalho.

Após a análise conduzida através do bibliometrix, os artigos foram selecionados por análise de resumos considerando a sua relevância com relação ao objetivo desse estudo. Dos 102 artigos, 46 foram definidos como relevantes. Infelizmente, durante a pesquisa, o acesso gratuito foi obtido para apenas 40 deles, os quais foram de fato utilizados como objetos dessa pesquisa.

Na Tabela 1 apresenta-se a divisão desses 40 trabalhos conforme o índice que se propõe a prever.

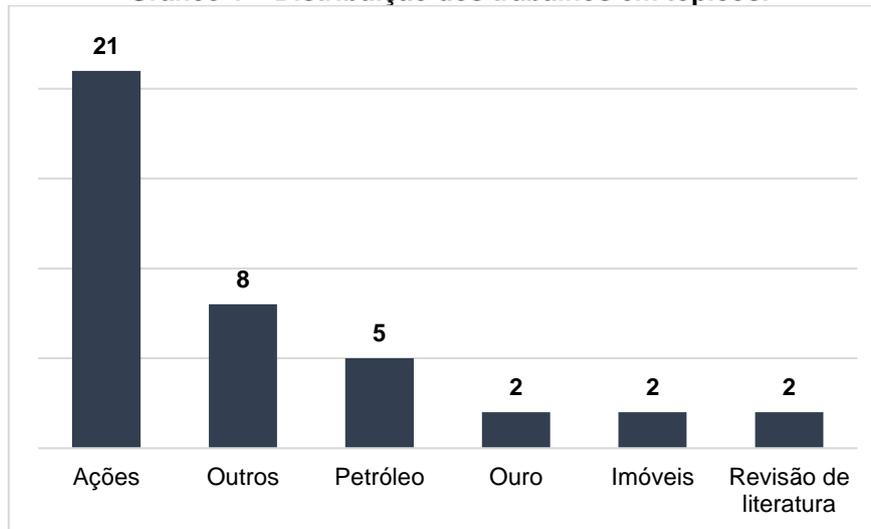
Tabela 1 – Quantidade de trabalhos divididos por assunto.

Objeto de estudo	Quantidade de trabalhos	Percentual
Preço de ação	21	52,5%
Preço de petróleo	5	12,5%
Preço de ouro	2	5,0%
Preço de imóvel	2	5,0%
Taxa de câmbio moeda indiana vs. Dólar	1	2,5%
Preços futuros agrícolas	1	2,5%
Preço de soja	1	2,5%
Preço de soja e milho	1	2,5%
Valor de mercado de banco	1	2,5%
Preço de metanol	1	2,5%
Preço de bitcoin	1	2,5%
Preço de carbono	1	2,5%
Revisão literatura (preço de petróleo)	1	2,5%
Revisão literatura (preço de ação)	1	2,5%

Fonte: Os autores (2023).

Com o intuito de facilitar a análise, os assuntos dos quais se teve acesso à apenas um trabalho foram agrupados na categoria “Outros”, sendo possível gerar o Gráfico 1 que apresenta a distribuição desses assuntos.

Gráfico 1 – Distribuição dos trabalhos em tópicos.



Fonte: Os autores (2023).

Dando sequência à análise, foram criadas as Tabelas 2, 3, 4, 5 e 6 apresentando, de forma resumida, as conclusões obtidas em cada um dos 40 artigos escolhidos. Todas estão apresentadas a seguir.

Tabela 2 – Conclusões sobre previsão de preços de ações utilizando dados de texto.

Título	Autores	Conclusão
Stock price prediction based on generative adversarial network	Li <i>et al.</i> (2022)	Utilizar análise de sentimento no modelo de previsão melhora sua performance comparado ao benchmark.
LSTM-based sentiment analysis for stock price forecast	Ko e Chang (2021)	Utilizar análise de sentimento em conjunto com o modelo univariado de séries temporais melhora sua performance preditora.
Analysis of stock market using text mining and natural language processing	Abdullah, Rahaman e Rahman (2013)	Utilizar quaisquer notícias pode guiar melhor a tomada de decisões com relação a tendências de preço, se comparado com dados cronológicos.
Quantitative analysis and sentiment analysis for stock price forecast: the case of colombo stock exchange	Pavithya <i>et al.</i> (2021)	É possível construir um modelo robusto de previsão da tendência do preço de ações utilizando notícias do dia anterior. Não é possível atingir uma performance excelente nessas previsões pois a previsão influenciará diretamente a decisão dos investidores, o que irá gerar uma alteração nova na tendência do preço.
Modeling public mood and emotion: blog and news sentiment and socio-economic phenomena	Chen e Chen (2019)	Utilizar a polaridade de sentimentos de conteúdos de blogs financeiros e artigos de notícias em previsões do preço de ações auxilia na melhora da performance dos modelos.
Financial news-based stock movement prediction using causality analysis of influence in the korean stock market	Nam e Seong (2019)	Utilizar notícias de empresas relacionadas e não se limitar à empresa objetivo melhora a performance do modelo de previsão.
ISSPM: a stock prediction model incorporating investor sentiment calculations based on fusedmax	Yang <i>et al.</i> (2021)	Utilizar uma rede neural LSTM em conjunto com análise de sentimentos pode prever com sucesso a tendência do valor de ações.

Tabela 2 – Conclusões sobre previsão de preços de ações utilizando dados de texto (continuação).

Título	Autores	Conclusão
Applying of a company's stock price prediction using data mining	Muthuselvi <i>et al.</i> (2019)	É possível utilizar métodos computacionais, como algoritmos e técnicas de mineração de dados, para analisar os padrões de comunicação de uma empresa e usar esses padrões para prever o desempenho da estrutura da empresa, como o desempenho das ações.
Stock market movement prediction using disparate text features with machine learning	Bouktif <i>et al.</i> (2019)	A alta e a queda nos preços das ações de uma empresa são afetados pelas opiniões públicas ou emoções expressas no Twitter. Os modelos baseados em árvore superam outros métodos em performance.
Sentiment analysis as a factor included in the forecasts of price changes in the stock exchange	Wojarnik (2021)	Utilizar sentimentos de textos postados em fóruns de discussão relacionados à bolsa de valores pode ser útil para analisar o comportamento dos preços das ações.
Research on stock price forecast based on news sentiment analysis—a case study of Alibaba	Zhang <i>et al.</i> (2018)	Utilizar sentimentos de notícias contribui para prever o preço mínimo, máximo e volume de transações de ações.
Stock price forecasting by combining news mining and time series analysis	Tang <i>et al.</i> (2009)	Utilizar relatórios de notícias pode fornecer informações adicionais para a previsão de preços de ações e melhorar as técnicas de previsão convencionais.
Using financial news sentiment for stock price direction prediction	Fazlija e Harder (2022)	Pontuações de sentimento calculadas por métodos de PLN de ponta podem auxiliar na previsão de preços de ações.
Forecasting the trends of stock price through social networks by fuzzy support vector machine with possibility measures	Hao <i>et al.</i> (2020)	A melhor taxa de precisão dos modelos preditivos é alcançada quando se utiliza tanto as informações textuais quanto as numéricas como entrada, evidenciando que, apesar do seu potencial, informações textuais de redes sociais não podem substituir inteiramente as informações de preços históricos em aplicações de previsão do mercado de ações.
Stock price forecasting based on BP neural network model of network public opinion	Yu <i>et al.</i> (2017)	Existe uma correlação significativa entre a opinião pública nas redes sociais e o preço das ações.
Using deep learning techniques in forecasting stock markets by hybrid data with multilingual sentiment analysis	Lin <i>et al.</i> (2022)	Utilizar dados híbridos - textos de redes sociais e dados técnicos de negociações - em modelos de previsão podem apresentar resultados melhores do que modelos com apenas uma natureza de dados. Modelos LSTM superam a performance de outros definidos como benchmark
BERT's sentiment score for portfolio optimization: a fine-tuned view in black and litterman model	Colasanto <i>et al.</i> (2022)	Utilizar notícias com um impacto específico nos mercados em modelos de previsão pode direcionar o comportamento de investidores.
NewsCATs: a news categorization and trading system	Knolmayer (2006)	Utilizar um dicionário específico para classificação automática de texto para prever as tendências de preços das ações gera um desempenho de modelo superior se comparado a utilizar dicionários mais genéricos.
Using financial news articles with minimal linguistic resources to forecast stock behavior	Linardos <i>et al.</i> (2015)	Utilizar análise de sentimento das notícias contribui significativamente para a previsão no valor das ações. Dentre os algoritmos disponíveis, o SVM desempenhou melhor a função.

Tabela 2 – Conclusões sobre previsão de preços de ações utilizando dados de texto (continuação).

Título	Autores	Conclusão
Exploiting financial news and social media opinions for stock market analysis using MCMC bayesian inference	Maragoudakis <i>et al.</i> (2015)	Os resultados apoiam a afirmação de que a incorporação de notícias financeiras como fonte de informação pode desempenhar um papel significativo na melhoria da previsão, enquanto isso não pareceu se validar no caso de opiniões públicas.
Optimized feature extraction based artificial intelligence technique for empirical analysis of stock market data	Kansal e Kumar (2019)	Utilizar análise de sentimento no modelo de previsão melhora sua performance comparado ao benchmark. O modelo projetado alcançou melhor precisão com a hibridização de CS e ANN (Redes Neurais Artificiais).

Fonte: Os autores (2023).

No caso dos preços de ações (Tabela 2), em resumo, os estudos analisados indicam que a incorporação da análise de sentimento melhora o desempenho dos modelos de previsão em relação aos *benchmarks*. A combinação de análise de sentimento com modelos de séries temporais, o uso de dados textuais de diversas fontes, como notícias e redes sociais, e a inclusão de dados de empresas relacionadas demonstraram consistentemente melhorar a capacidade de previsão. Modelos baseados em redes neurais, como LSTM, têm se destacado.

Conforme apresentado na Tabela 3, para o caso do preço do petróleo, os estudos também sugerem que o uso de textos de notícias é confiável na previsão da tendência dos preços. A inclusão de notícias traz informações significativas e melhora a precisão dos modelos de previsão de preços do petróleo bruto, demonstrando uma relação entre os retornos de preços do petróleo e notícias online. Um ponto de atenção é o trabalho de Zhao e Zeng (2019) no qual foi observado que a eficácia desses textos pode diminuir ao longo do tempo devido à atualidade das notícias.

Tabela 3 – Conclusões sobre previsão de preços de petróleo utilizando dados de texto.

Título	Autores	Conclusão
Analysis of timeliness of oil price news information based on SVM	Zhao e Zeng (2019)	O texto de notícias tem certa confiabilidade na previsão da tendência dos preços do petróleo. Entretanto, considerando a atualidade das notícias, é possível que a capacidade dos textos de notícias de prever tendências futuras de preços do petróleo diminuirá com o tempo.
Forecasting weekly crude oil using twitter sentiment of U.S. foreign policy and oil companies' data	Zaidi e Oussalah (2018)	Utilizar análise de sentimento no modelo de previsão de preço de petróleo melhora sua performance comparado ao benchmark.
Text-based crude oil price forecasting: a deep learning approach	Li <i>et al.</i> (2019)	O texto de notícias adiciona informações significativas à previsão do preço do petróleo bruto e trazem melhoras na precisão dos modelos.
Forecasting oil price volatility in the era of big data: a text mining for VaR approach	Zhao <i>et al.</i> (2019)	O aumento da precisão na previsão do modelo criado (VaR) prova que existe uma interação entre os retornos de preços do petróleo e notícias online relacionadas ao petróleo.
Forecasting oil price using web-based sentiment analysis	Zhao <i>et al.</i> (2019)	A relação entre o preço do petróleo e o sentimento do texto da web é quasi-linear. O uso de modelos de previsão altamente não lineares, como SVR e RF, não retorna um bom desempenho. Os melhores resultados obtidos por LASSO e Ridge indicam que modelos lineares ainda precisam ser considerados.

Fonte: Os autores (2023).

Na Tabela 4 estão as conclusões dos trabalhos a respeito do preço do ouro, a qual também demonstra um resultado positivo ao se utilizar dados textuais em modelos preditivos. A mesma inferência é válida ao analisarmos a Tabela 5, a qual trata de preços de imóveis.

Tabela 4 – Conclusões sobre previsão de preços de ouro utilizando dados de texto.

Título	Autores	Conclusão
The integration of artificial neural networks and text mining to forecast gold futures prices	Chen, Chen e Chiu (2016)	Utilizar análise de sentimento em conjunto com o modelo univariado de séries temporais melhora sua performance preditora de preços de ouro.
Using market sentiment analysis and genetic algorithm-based least squares support vector regression to predict gold prices	Yuan <i>et al.</i> (2020)	As opiniões obtidas através de mineração de texto aprimoraram a previsão do preço do ouro através do modelo GA-LSSVR.

Fonte: Os autores (2023).

Tabela 5 – Conclusões sobre previsão de preços de imóveis utilizando dados de texto.

Título	Autores	Conclusão
A text-based mining approach for real estate policy impact monitoring and analysis	Cao <i>et al.</i> (2021)	Considerar dados de opiniões públicas na internet em um framework de processamento de dados textuais contribuem para melhoras na previsão de tendências de preço de imóveis.
Can machine learning algorithms associated with text mining from internet data improve housing price prediction performance?	Guo <i>et al.</i> (2020)	O modelo de Random Forest é a melhor escolha como algoritmo para prever tendências de mercado imobiliário a partir de mineração de texto.

Fonte: Os autores (2023).

Por fim, na Tabela 6 estão os demais estudos, os quais destacam a eficácia da análise de sentimentos em dados de mídia social e informações textuais para prever preços e tendências em diversos mercados. A combinação de análise de sentimentos com técnicas de clusterização, *deep learning*, modelos híbridos e representação de séries temporais mostrou-se promissora para melhorar a acurácia das previsões em áreas que vão desde a Bitcoin até os preços de carbono, metanol, soja, milho e até mesmo o valor de mercado de bancos. Essas abordagens têm a capacidade de fornecer insights valiosos para pesquisadores e profissionais de previsão de preços, especialmente em contextos de oscilações abruptas e previsões de curto prazo.

Tabela 6 – Conclusões sobre previsão de preços dos demais objetos de estudo utilizando dados de texto.

Título	Autores	Conclusão
Analyzing BTC's trend during COVID-19 using a sentiment consensus clustering (SCC)	Ibrahim (2021)	Utilizar análise de sentimentos em tweets em conjunto com uma técnica de clusterização apresenta potencial para prever o preço da bitcoin em períodos de oscilação brusca.
Carbon price prediction models based on online news information analytics	Zhang e Xia (2022)	Utilizar informações de mídia social e aplicar uma abordagem de aprendizado profundo fornecem insights para pesquisadores de previsão de preços de carbono . O modelo criado foi performou melhor que outros escolhidos como benchmark.
Forecast the price of chemical products with multivariate data	Zhang <i>et al.</i> (2015)	Utilizar modelos híbridos com dados da web e experiência de especialistas maximiza a precisão da previsão dos preços de metanol .
Text-based soybean futures price forecasting: a two-stage deep learning approach	An <i>et al.</i> (2023)	Utilizar análise textual em mídias sociais no modelo de previsão dos preços da soja melhora sua performance comparado aos mais simples testados.

Tabela 6 – Conclusões sobre previsão de preços dos demais objetos de estudo utilizando dados de texto (continuação).

Título	Autores	Conclusão
On the enrichment of time series with textual data for forecasting agricultural commodity prices	Reis <i>et al.</i> (2022)	Modelos de representação de séries temporais que consideram informações textuais dificilmente terão um desempenho melhor em todas as etapas de previsão, no entanto, esse modelo se destaca ao prever oscilações abruptas em séries temporais. Comprovado nesse estudo considerando os preços da soja e do milho .
A novel text-based framework for forecasting agricultural futures using massive online news headlines	Li <i>et al.</i> (2022)	Utilizar variáveis qualitativas derivadas de sentimentos de notícias ajudam nas previsões dos preços dos futuros agrícolas anteriormente feitas apenas com informações numéricas.
An intelligent market capitalization predictive system using deep learning	Santhappan e Chokkalingam (2018)	A precisão dos modelos para previsão do valor de mercado de bancos foi aprimorada ao utilizar tweet para compor a previsão. O modelo apresenta maior acurácia para previsões semanais do que mensais.

Fonte: Os autores (2023).

4. Conclusões

Neste trabalho de revisão sistemática da literatura, foram analisados 40 artigos que se dedicaram à criação de modelos de previsão de preços utilizando alguma fonte de dados textual como entrada. A maioria desses artigos abordou a previsão de preços de ações, mas também incluíram trabalhos sobre preços de petróleo, ouro, imóveis e outros ativos financeiros. O objetivo principal desta revisão foi avaliar a possibilidade de usar dados textuais de notícias online e redes sociais para realizar previsões de preços e auxiliar na tomada de decisões de investimento.

Os resultados da pesquisa indicam que a utilização de dados textuais da web tem um grande potencial na previsão de preços de diversos índices do mercado financeiro. Todos os estudos analisados obtiveram resultados positivos, sugerindo que esse campo está em crescimento e oferece muitas oportunidades de pesquisa. No geral, fica evidente que as ferramentas de inteligência artificial para análise textual podem fornecer informações valiosas para o mercado financeiro.

No entanto, é importante notar que esta pesquisa teve suas limitações. Ela utilizou somente a base de dados Scopus, além disso, alguns artigos não estavam com acesso livre para análise completa. Futuras pesquisas sobre esse tema podem expandir a busca para outras bases de dados e utilizar ferramentas automáticas de análise textual para revisões sistemáticas, permitindo uma análise mais abrangente de estudos adicionais.

Referências

- ABDULLAH, S. S.; RAHAMAN, M. S.; RAHMAN, M. S. Analysis of stock market using text mining and natural language processing. In: International Conference on Informatics, Electronics and Vision (ICIEV), 2013/5/17-2013/5/18, [s. l.], Anais [...]. [S. l.]: IEEE, 2013.
- AN, W.; WANG, L.; ZENG, Y.-R. Text-based soybean futures price forecasting: A two-stage deep learning approach. **Journal of forecasting**, [s. l.], v. 42, n. 2, p. 312–330, 2023. Disponível em: <http://dx.doi.org/10.1002/for.2909>.
- BOUKTIF, S.; FIAZ, A.; AWAD, M. Stock market movement prediction using disparate text features with machine learning. In: Third International Conference on Intelligent Computing in Data Sciences (ICDS), 2019/10/28-2019/10/30, [s. l.], Anais [...]. [S. l.]: IEEE, 2019.

CAO, L.; XU, P.; SHANG, W. A text-based mining approach for real estate policy impact monitoring and analysis. In: IEEE International Conference on Big Data (Big Data), 2021/12/15-2021/12/18, [s. l.], Anais [...]. [S. l.]: IEEE, 2021.

CHAN, Yue-cheong; CHUI, Andy CW; KWOK, Chuck CY. The impact of salient political and economic news on the trading activity. **Pacific-Basin Finance Journal**, v. 9, n. 3, p. 195-217, 2001.

CHEN, M.-Y.; CHEN, T.-H. Modeling public mood and emotion: Blog and news sentiment and socio-economic phenomena. **Future generations computer systems: FGCS**, [s. l.], v. 96, p. 692–699, 2019. Disponível em: <http://dx.doi.org/10.1016/j.future.2017.10.028>.

CHEN, H.-H.; CHEN, M.; CHIU, C.-C. The integration of artificial neural networks and text mining to forecast gold futures prices. **Communications in statistics: Simulation and computation**, [s. l.], v. 45, n. 4, p. 1213–1225, 2016. Disponível em: <http://dx.doi.org/10.1080/03610918.2013.786780>.

CHEN, Xun-Qi et al. Explainable artificial intelligence in finance: A bibliometric review. **Finance Research Letters**, p. 104145, 2023.

COLASANTO, F.; GRILLI, L.; SANTORO, D.; VILLANI, G. BERT's sentiment score for portfolio optimization: a fine-tuned view in Black and Litterman model. **Neural computing & applications**, [s. l.], v. 34, n. 20, p. 17507–17521, 2022. Disponível em: <http://dx.doi.org/10.1007/s00521-022-07403-1>.

FAZLIJA, B.; HARDER, P. Using financial news sentiment for stock price direction prediction. **Mathematics**, [s. l.], v. 10, n. 13, p. 2156, 2022. Disponível em: <http://dx.doi.org/10.3390/math10132156>.

GRUNDMANN, Tina; FELDEN, Carsten; POSPIECH, Marco. Forecasting the natural gas price trend-evaluation of a sentiment analysis. In: 2016 IEEE 4th International Conference on Future Internet of Things and Cloud Workshops (FiCloudW). IEEE, 2016. p. 160-164.

GUO, J.-Q.; CHIANG, S.-H.; LIU, M.; YANG, C.-C.; GUO, K.-Y. Can machine learning algorithms associated with text mining from Internet data improve housing price prediction performance?. **International journal of strategic property management**, [s. l.], v. 24, n. 5, p. 300–312, 2020. Disponível em: <http://dx.doi.org/10.3846/ijspm.2020.12742>.

HAO, P.-Y.; OU, J.-B.; ZHUANG, K.-X. Forecasting the trends of stock price through social networks by fuzzy support vector machine with possibility measures. In: Proceedings of the 7th Multidisciplinary in International Social Networks Conference and the 3rd International Conference on Economics, Management and Technology, 2020, New York, NY, USA. Anais [...]. New York, NY, USA: ACM, 2020.

IBRAHIM, A. Analyzing BTC's trend during COVID-19 using A sentiment consensus clustering (SCC). In: IEEE 12th Annual Information Technology, Electronics and Mobile Communication Conference (IEMCON), 2021/10/27-2021/10/30, [s. l.], Anais [...]. [S. l.]: IEEE, 2021.

JOINER, D.; VEZEAU, A.; WONG, A.; HAINS, G.; KHMELEVSKY, Y. Algorithmic trading and short-term forecast for financial time series with machine learning models; State of the art and perspectives. In: IEEE International Conference on Recent Advances in Systems Science and Engineering (RASSE), 2022/11/7-2022/11/10, [s. l.], Anais [...]. [S. l.]: IEEE, 2022.

KANSAL, V.; KUMAR, R. Optimized feature extraction based artificial 1301 intelligence technique for empirical analysis of stock market data. **Int. J. Innov. Technol. Exploring Eng.**, v. 8, n. 10, 2019.

- KO, C.-R.; CHANG, H.-T. LSTM-based sentiment analysis for stock price forecast. **PeerJ. Computer science**, [s. l.], v. 7, n. e408, p. e408, 2021. Disponível em: <http://dx.doi.org/10.7717/peerj-cs.408>.
- KITCHENHAM, B. Procedures for performing systematic reviews. Keele, UK, Keele University, v. 33, n. 2004, p. 1-26, 2004.
- LI, Y.; CHENG, D.; HUANG, X.; LI, C. Stock price prediction Based on Generative Adversarial Network. In: International Conference on Big Data, Information and Computer Network (BDICN), 2022/1/20-2022/1/22, [s. l.], Anais [...]. [S. l.]: IEEE, 2022.
- LI, J.; LI, G.; LIU, M.; ZHU, X.; WEI, L. A novel text-based framework for forecasting agricultural futures using massive online news headlines. **International journal of forecasting**, [s. l.], v. 38, n. 1, p. 35–50, 2022. Disponível em: <http://dx.doi.org/10.1016/j.ijforecast.2020.02.002>.
- LI, X.; SHANG, W.; WANG, S. Text-based crude oil price forecasting: A deep learning approach. **International journal of forecasting**, [s. l.], v. 35, n. 4, p. 1548–1560, 2019. Disponível em: <http://dx.doi.org/10.1016/j.ijforecast.2018.07.006>.
- LIN, Y.-L.; LAI, C.-J.; PAI, P.-F. Using deep learning techniques in forecasting stock markets by hybrid data with multilingual sentiment analysis. **Electronics**, [s. l.], v. 11, n. 21, p. 3513, 2022. Disponível em: <http://dx.doi.org/10.3390/electronics11213513>.
- LINARDOS, E.; KERMANIDIS, K. L.; MARAGOUDAKIS, M. Using financial news articles with minimal linguistic resources to forecast stock behaviour. **International Journal of Data Mining Modelling and Management**, [s. l.], v. 7, n. 3, p. 185, 2015. Disponível em: <http://dx.doi.org/10.1504/ijdm.2015.071452>.
- MITTERMAYER, M.-A.; KNOLMAYER, G. NewsCATS: A news categorization and trading system. In: Sixth International Conference on Data Mining (ICDM'06), 2006/12/18-2006/12/22, [s. l.], Anais [...]. [S. l.]: IEEE, 2006.
- MUTHUSELVI, S.; RENGARAJAN, A.; SCINTIACLARINDA, S.; NITHYA, K.; Applying of a company's stock price prediction using data mining. **International Journal of Recent Technology and Engineering (IJRTE)**, [s. l.], v. 8, n. 2, p. 2847–2850, 2019. Disponível em: <http://dx.doi.org/10.35940/ijrte.b2039.078219>.
- NAM, K.; SEONG, N. Financial news-based stock movement prediction using causality analysis of influence in the Korean stock market. **Decision support systems**, [s. l.], v. 117, p. 100–112, 2019. Disponível em: <http://dx.doi.org/10.1016/j.dss.2018.11.004>.
- OUSSALAH, M.; ZAIDI, A. Forecasting weekly crude oil using twitter sentiment of U.s. foreign policy and oil companies data. In: IEEE International Conference on Information Reuse and Integration (IRI), 2018/7/6-2018/7/9, [s. l.], Anais [...]. [S. l.]: IEEE, 2018.
- PAVITHYA, M. B. D.; PERERA, G. S. D.; MUNASINGHE, S. L.; KARUNARATHNA, S. N. Quantitative analysis and sentiment analysis for stock price forecast: The case of Colombo stock exchange. In: 10th International Conference on Information and Automation for Sustainability (ICIAFS), 2021/8/11-2021/8/13, [s. l.], Anais [...]. [S. l.]: IEEE, 2021.
- REIS FILHO, I. J.; MARCACINI, R. M.; REZENDE, S. O. On the enrichment of time series with textual data for forecasting agricultural commodity prices. **MethodsX**, [s. l.], v. 9, n. 101758, p. 101758, 2022. Disponível em: <http://dx.doi.org/10.1016/j.mex.2022.101758>.
- SANTHAPPAN, J.; CHOKKALINGAM, P. An intelligent market capitalization predictive system using deep learning. In: International Conference on Advanced Computation and Telecommunication (ICACAT), 2018/12/28-2018/12/29, [s. l.], Anais [...]. [S. l.]: IEEE, 2018.

SANTOS, M. V.; MORGADO-DIAS, F.; SILVA, T. C. Oil sector and sentiment analysis—A review. **Energies**, [s. l.], v. 16, n. 12, p. 4824, 2023. Disponível em: <http://dx.doi.org/10.3390/en16124824>.

TADPHALE, A.; SARASWAT, H.; SONAWANE, O.; DESHMUKH, P. R. Impact of news sentiment on foreign exchange rate prediction. In: 3rd International Conference on Intelligent Technologies (CONIT), 2023/6/23-2023/6/25, [s. l.], Anais [...]. [S. l.]: IEEE, 2023.

TANG, X.; YANG, C.; ZHOU, J. Stock price forecasting by combining news mining and time series analysis. In: IEEE/WIC/ACM International Joint Conference on Web Intelligence and Intelligent Agent Technology, 2009/9/15-2009/9/18, [s. l.], Anais [...]. [S. l.]: IEEE, 2009.

WOJARNIK, G. Sentiment analysis as A factor included in the forecasts of price changes in the stock exchange. **Procedia computer science**, [s. l.], v. 192, p. 3176–3183, 2021. Disponível em: <http://dx.doi.org/10.1016/j.procs.2021.09.090>.

YANG, Y.; CHEN, Siting; CHEN, Z.; CHEN, Shaobo; LI, R.; CAI, Z.; GU, H.; YIN, H.; QUAN, Y. ISSPM: A stock prediction model incorporating investor sentiment calculations based on fusedmax. In: 2021 8th International Conference on Information, Cybernetics, and Computational Social Systems (ICCSS), 2021/12/10-2021/12/12, [s. l.], Anais [...]. [S. l.]: IEEE, 2021.

YU, Y.; WANG, S.; ZHANG, L. Stock price forecasting based on BP neural network model of network public opinion. In: 2nd International Conference on Image, Vision and Computing (ICIVC), 2017/6/2-2017/6/4, [s. l.], Anais [...]. [S. l.]: IEEE, 2017.

YUAN, F.-C.; LEE, C.-H.; CHIU, C. Using market sentiment analysis and genetic algorithm-based least squares support vector regression to predict gold prices. **International journal of computational intelligence systems**, [s. l.], v. 13, n. 1, p. 234, 2020. Disponível em: <http://dx.doi.org/10.2991/ijcis.d.200214.002>.

ZHANG, L.; FU, S.; LI, B. Research on stock price forecast based on news sentiment analysis - A case study of alibaba. In: Lecture Notes in Computer Science. Cham: Springer International Publishing, 2018. p. 429–442.

ZHANG, F.; XIA, Y. Carbon price prediction models based on online news information analytics. **Finance research letters**, [s. l.], v. 46, n. 102809, p. 102809, 2022. Disponível em: <http://dx.doi.org/10.1016/j.frl.2022.102809>.

ZHANG, Xia et al. Forecast the price of chemical products with multivariate data. In: 2015 International Conference on Behavioral, Economic and Socio-cultural Computing (BESC). IEEE, 2015. p. 76-82.

ZHANG, X.; YIN, H.; WANG, C.; WANG, J.; ZHANG, Y. Forecast the price of chemical products with multivariate data. In: International Conference on Behavioral, Economic and Socio-Cultural Computing (BESC), 2015/10/30-2015/11/1, [s. l.], Anais [...]. [S. l.]: IEEE, 2015.

ZHAO, L.-T.; LIU, L.-N.; WANG, Z.-J.; HE, L.-Y. Forecasting oil price volatility in the era of big data: A text mining for VaR approach. **Sustainability**, [s. l.], v. 11, n. 14, p. 3892, 2019a. Disponível em: <http://dx.doi.org/10.3390/su11143892>.

ZHAO, L.-T.; ZENG, G.-R. Analysis of timeliness of oil price news information based on SVM. **Energy procedia**, [s. l.], v. 158, p. 4123–4128, 2019. Disponível em: <http://dx.doi.org/10.1016/j.egypro.2019.01.821>.

ZHAO, L.-T.; ZENG, G.-R.; WANG, W.-J.; ZHANG, Z.-G. Forecasting oil price using web-based sentiment analysis. **Energies**, [s. l.], v. 12, n. 22, p. 4291, 2019b. Disponível em: <http://dx.doi.org/10.3390/en12224291>.