Aplicação de Técnicas de Mineração de Dados e Aprendizagem de Máquina no Mercado de Ações: Uma Revisão Sistemática

Daniele Gonçalves de Toledo Luchetta Raminelli, Bruno Samways dos Santos

Resumo: O mercado de ações de grandes nações possui um volume de operações consideravelmente alto, o que justifica o interesse de investidores e instituições financeiras em aplicar técnicas de mineração para analisar corretamente os dados de forma a extrair conhecimento, prever tendências, preços futuros de ações e meios de obter os maiores lucros possíveis. Este artigo teve como objetivo realizar uma revisão sistemática sobre as técnicas de mineração de dados e aprendizagem de máquina aplicadas no mercado de ações até o ano de 2018 usando as bases de dados *Scopus*, *Web of Science* e *ScienceDirect*. Quanto às tarefas de mineração de dados e aprendizagem de máquina utilizadas nas aplicações, a predição foi a que mais se destacou, o que é de se esperar visto que uma das principais vantagens competitivas almejadas no mercado de ações é o ato de prever os preços futuros. A tarefa de associação também obteve destaque, indicando a busca por correlações existentes entre o preço de ações e fatores externos. Quanto às técnicas e ferramentas utilizadas para manipular os dados, houve um relevante destaque da aplicação de *Neural Networks* (ou redes neurais). Notou-se também o crescente uso de *Deep Learning* (ou aprendizado profundo). O trabalho cumpriu o objetivo proposto, mostrando o cenário atual de utilização de técnicas de mineração de dados e aprendizagem para problemas relacionados ao mercado de ações.
Palavras chave: Pesquisa Operacional, Inteligência Computacional, Mineração de Dados, Aprendizagem de Máquina, Mercado de Ações.

Application of Data Mining and Machine Learning Techniques in the Stock Market: A Systematic Review

Abstract: The stock market of large nations has a considerably high trading volume, which justifies the interest of investors and financial institutions to apply mining techniques to correctly analyze data to extract knowledge, forecast trends, future stock prices and means, to get the highest profits possible. This paper aimed to conduct a systematic review of the data mining and machine learning techniques applied in the stock market by 2018 using the Scopus, Web of Science and ScienceDirect databases. As for the data mining and machine learning tasks used in the applications, the prediction was the one that stood out, which is to be expected since one of the main competitive advantages sought by the stock market is the act of predicting future prices. The association task was also highlighted, indicating the search for existing correlations between stock price and external factors. Regarding the techniques and tools used to manipulate the data, there was a relevant highlight of the application of Neural Networks. There has also been a growing use of Deep Learning. The work met the proposed objective by showing the current scenario of using data mining techniques and learning for stock market issues.
Key-words: Operational Research, Computational Intelligence, Data Mining, Machine Learning, Stock Market.

1. Introdução

As técnicas de mineração de dados associadas a algoritmos de aprendizagem de máquina estão desempenhando papéis cada vez mais importantes em diversas aplicações e setores, sobretudo àqueles com grande volume de dados. Algumas das aplicações podem ser

encontradas no setor público, serviços financeiros, área da saúde, manufatura, telecomunicações, varejo, entre outras indústrias (KUMAR, 2015).

Como exemplo do setor financeiro, o mercado de ações de grandes nações possui um volume de operações consideravelmente alto, o que justifica o interesse de investidores e instituições financeiras em aplicar técnicas de mineração para analisar corretamente os dados de forma a extrair conhecimento, prever tendências, preços futuros de ações e meios de obter os maiores lucros possíveis.

A previsão do mercado de ações é considerada uma tarefa desafiadora visto os preços das ações reagem à fatores econômicos, financeiros, políticos e comportamentais. Entretanto, com o avanço da inteligência artificial e capacidade de processamento de dados das últimas décadas, tornou-se possível prever os movimentos dinâmicos do mercado financeiro seja com modelos de regressão, e até mesmo de classificação (VALENCIA et al., 2019). Sendo assim, a análise do mercado de ações tem sido um tema atraente para diversas pesquisas de aplicação de técnicas computacionais, bem como para servir de auxílio à tomada de decisão.

Visto que técnicas de mineração de dados e aprendizagem de máquinas podem analisar eficientemente grandes bancos de dados, torna-se relevante verificar como a literatura científica tem aplicado tais abordagens no mercado financeiro para descobrir informações relevantes e novos conhecimentos.

O artigo tem como objetivo:

a) Analisar o número de trabalhos publicados e citações (até o ano de 2018);

b) Determinar quais técnicas e ferramentas estão sendo utilizadas na aplicação de mineração de dados e aprendizagem de máquina no mercado de ações;

c) Constatar a importância científica do tema e fornecer um panorama atual.

A estrutura do trabalho está organizada da seguinte maneira:

- A seção 2 fornece uma breve explicação sobre o que a mineração de dados, a aprendizagem de máquina e sobre a bolsa de valores e o mercado de ações;
- A seção 3 descreve o método utilizado para executar a revisão sistemática, as etapas realizadas e o detalhamento de cada uma delas;
- Na seção 4 tem-se os dados das análises dos artigos, gráficos que ilustram os resultados obtidos e uma breve discussão para a conclusão da pesquisa realizada.

2. Referencial teórico

O referencial teórico deste artigo aborda, brevemente, os conceitos de mineração de dados, aprendizagem de máquina e o mercado de ações.

2.1 Mineração de dados

Segundo Witten e Frank (2005, p. 4, tradução livre), estamos sobrecarregados com dados. A quantidade de dados no mundo, em nossas vidas, parece continuar aumentando - e não há fim à vista. Torna-se então necessário o uso de teorias e ferramentas para a extração de informações úteis de dados. Tais teorias e ferramentas são o objeto de estudo da Descoberta de Conhecimento em dados ou KDD (de *Knowledge Discovery of Databases*), que se ocupa com o desenvolvimento de métodos e técnicas para compreender os dados (FAYYAD et al., 1996).

Diversas definições de mineração de dados ou DM (de *Data Mining*) podem ser encontradas na literatura. Berry e Linoff (2004, p. 7), descrevem a mineração de dados como a exploração e análise de grandes quantidades de dados para descobrir padrões e regras interessantes. Para Fayyad et al. (1996, p. 39), o KDD consiste no processo geral de descoberta de conhecimento útil a partir de dados e a mineração de dados consiste em uma etapa específica deste processo. Pode-se dizer então que o objetivo principal da mineração de dados é obter informações não-triviais de um banco de dados por meio de técnicas e algoritmos.

De acordo com Larose D e Larose C (2014, p. 8) as seis tarefas mais comuns de mineração de dados são:

- Descrição;
- Estimativa;
- Predição;
- Classificação;
- Agrupamento;
- Associação.

Cada tarefa (ou resultado que se almeja alcançar) possui técnicas indicadas para manejar os dados. Cabe ressaltar que as técnicas de mineração de dados não estão restritas a apenas uma tarefa cada, o que resulta em uma boa quantidade de sobreposição entre técnicas e tarefas de mineração de dados.

2.2 Aprendizagem de Máquina

A aprendizagem de máquina ou ML (de *Machine Learning*) é considerada um ramo da inteligência artificial e possui inúmeros campos de aplicação. O conceito inicial surgiu em 1959 por Arthur Samuel. O autor afirma que: “aprendizado de máquina é um campo de estudo que dá a computadores a capacidade de aprender sem ser programado de forma explícita” (SAMUEL, 1959). Para Mitchell (1997), o objetivo da aprendizagem de máquina é construir modelos computacionais que podem adaptar-se e aprender a partir da experiência contida nos dados.

Cuocolo et al. (2019, p. 2) relata que os algoritmos de ML podem ser classificados em três tipos diferentes:

- Aprendizado supervisionado (depende da rotulagem dos dados do treinamento antes do processo de aprendizagem);
- Aprendizado não supervisionado (caracterizado pela ausência de divisão humana preliminar de dados em categorias);
- Aprendizado por reforço (na qual o algoritmo aprende com seus erros e sucessos devido um processo contínuo de feedback).

Em suma, programa-se o computador para que, por meio de dados de experiências anteriores, otimize-se um critério de desempenho. É necessária a aplicação de aprendizagem de máquina nos casos em que não podemos expressar direta ou facilmente como um programa poderia ser resolvido por algoritmos, entretanto podemos identificar conjuntos de exemplos que ilustram a solução (LEE et al., 2017)

2.3 Bolsa de valores e o mercado de ações

Pinheiro (2005) ressalta que o mercado de capitais é um conjunto de instituições que negociam com títulos e valores, com o objetivo de canalizar os recursos dos agentes

compradores para os agentes vendedores, viabilizando a capitalização de empresas.
A bolsa de valores é o local em que empresas que possuem capital aberto negociam suas ações (ou ativos).

"Compre na baixa, venda na alta" é uma das frases mais utilizadas para evidenciar que é possível obter lucro com a negociação de tais ações. Entretanto, a frase que parece simples, não demonstra a grande complexidade de um mercado em que o futuro é praticamente imprevisível. Sendo assim, métodos que auxiliam na solução de problemas de previsibilidade e outras correlações no mercado de ações provocam grande interesse por pesquisadores, instituições financeiras e investidores. Analisaremos estes métodos e técnicas adiante.

3. Metodologia de pesquisa

A questão de pesquisa que norteou este trabalho foi: Qual é o cenário atual de utilização de técnicas de mineração de dados e aprendizagem de máquina para problemas relacionados ao mercado de ações?

Esta revisão sistemática seguiu a metodologia delineada em Tranfield et al. (2003) com adaptações, consistindo em sete etapas:

- Etapa 1: Escopo do estudo;
- Etapa 2: Identificação de termos de pesquisa;
- Etapa 3: Identificação de fontes de dados;
- Etapa 4: Coleta de artigos;
- Etapa 5: Filtragem de artigos;
- Etapa 6: Avaliação de conteúdo;
- Etapa 7: Síntese e resultados.

A seguir tem-se a descrição detalhada de cada etapa.

3.1 Escopo do estudo

Este estudo enfoca a aplicação de técnicas de mineração de dados e aprendizagem de máquina no mercado de ações.

3.2 Identificação de termos de pesquisa

Para enquadrar o escopo do estudo, identificou-se as palavras-chave usadas como termos de pesquisa para capturar artigos relevantes. Foram incluídos termos de pesquisa relacionados às técnicas desejadas como "*data mining*" ou "*machine learning*" combinados com qualquer um dos termos de pesquisa relacionados à área de aplicação desejada do mercado de ações, como "*stock market*" ou "*stock prices*". Essas palavras-chave e suas combinações foram consideradas para representar um conjunto de termos de pesquisa para desvendar a literatura científica pertinente.

3.3 Identificação de fontes de dados

Nossas fontes de dados consistiram em:

a) *Scopus*;

b) *Web of Science*;

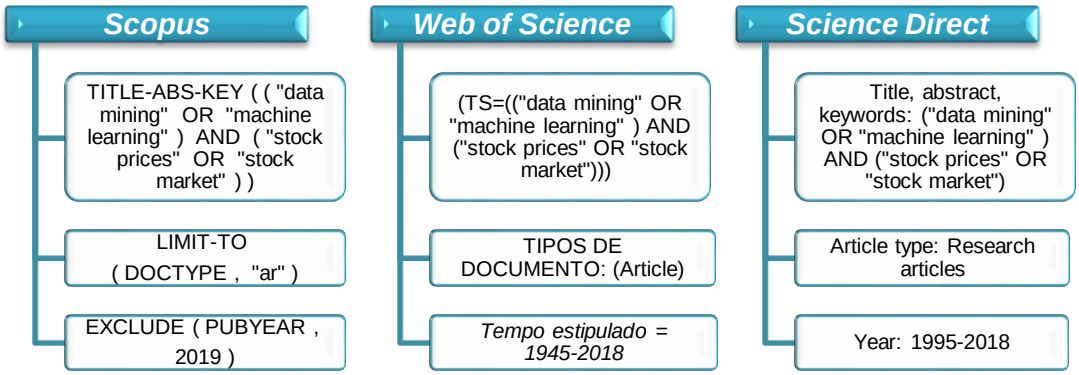
c) *ScienceDirect.*

Considerou-se que estes canais representam uma coleção abrangente de fontes literárias que

projetaram uma rede ampla o suficiente para cobrir pesquisas relevantes para o escopo do estudo.

3.4 Coleta de artigos

Foi pesquisada a literatura sobre aplicações de mineração de dados e aprendizagem de máquinas no mercado de ações usando as combinações dos termos de pesquisa especificados na Etapa 2, sem restrição de tempo ou de saída nas múltiplas fontes eletrônicas. As *strings* de busca podem ser vistas na Figura 1.



Fonte: Autoria própria.

Figura 1 – *Strings* de busca utilizadas nas bases de dados

3.5. Filtragem de artigos

Um processo de remoção dos artigos duplicados entre as bases foi realizado utilizando o software *Mendeley*.

Em seguida um processo manual de inspeção e filtragem foi realizado pelos autores para incluir apenas artigos que satisfazem os seguintes critérios de inclusão:

a) descrever uma aplicação específica de DM ou ML no mercado de ações;

b) descrever explicitamente as técnicas de DM ou ML que foram utilizadas;

c) tais critérios serem visíveis por meio da leitura do abstract.

Todos os outros artigos que não descreveram uma aplicação concreta de mineração de dados ou aprendizagem de máquinas, como artigos interpretativos, comentários e revisões de literatura, ou que não forneceram detalhes suficientes para satisfazer os critérios de inclusão foram excluídos.

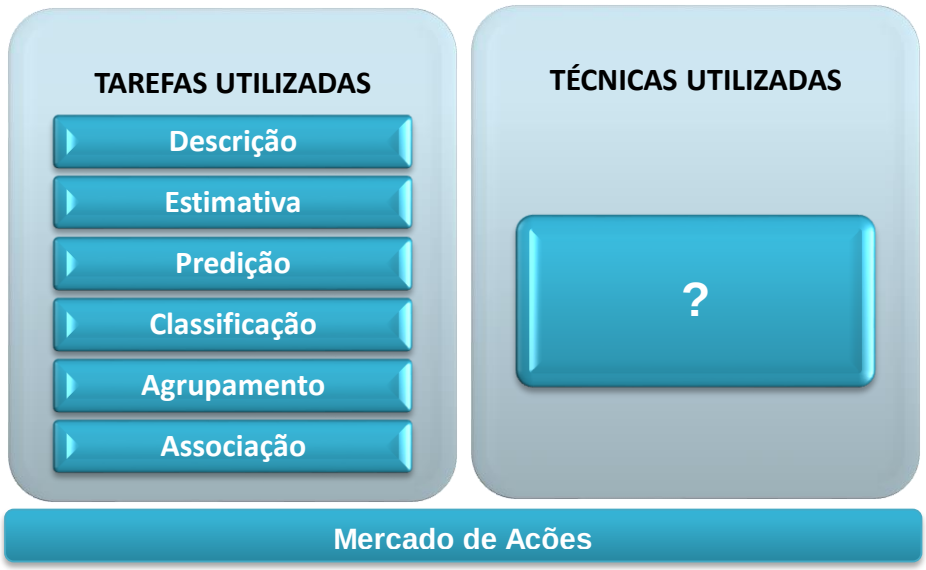
3.6. Avaliação de conteúdo

Foi utilizado um formulário de extração de dados para capturar:

- detalhes bibliográficos (incluindo autor(es), data de publicação, título);
- palavras-chaves do artigo;
- tarefa(s) da mineração de dados (descrição, estimativa, predição, classificação, agrupamento, associação);
- técnica(s) de mineração de dados e aprendizagem de máquina (por exemplo, regressão, redes neurais, árvores de decisão, máquinas de vetores de suporte).

3.7 Síntese e resultados

O objetivo de pesquisa foi capturar a literatura sobre o maior número possível de aplicações de DM e ML no mercado de ações e identificar a natureza e as principais técnicas aplicadas. A Figura 2 ilustra a classificação proposta. Por fim, os artigos mais citados nas bases pesquisadas foram analisados mais a fundo.

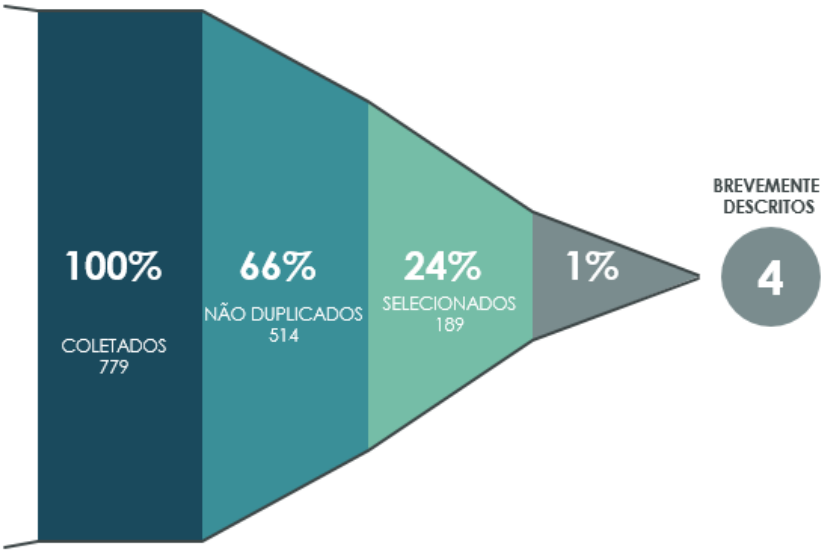


Fonte: Autoria própria.

Figura 2 – Classificação proposta

4. Análises e resultados

A revisão sistemática seguiu a metodologia proposta, e após as etapas obteve os resultados que podem ser vistos na Figura 3.



Fonte: Autoria própria.

Figura 3 – Resultados obtidos

As etapas detalhas da pesquisa são descritas a seguir.

Para realizar a pesquisa nas bases definidas foram utilizados combinadores booleanos identificando as palavras-chave desejadas. Os resultados são apresentados na Tabela 1.

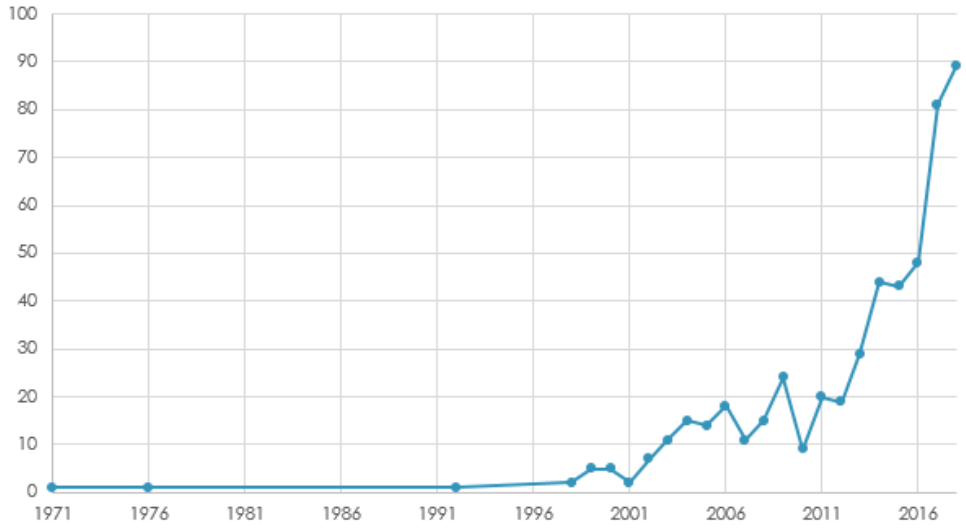
Palavras-chave	*Scopus*	*Web of Science*	*Science Direct*
"Machine Learning" OR "Data Mining" AND "Stock Market" OR "Stock Prices"	423	234	122

Fonte: Autoria própria.

Tabela 1 – Resultados da pesquisa por termos de busca

Em seguida, foi utilizado o *software Mendeley* para remover os artigos duplicados, resultando na quantidade de 514 artigos únicos.

A Figura 4 apresenta o quantitativo de publicações até 2018, considerando as bases pesquisadas.



Fonte: Autoria própria.

Figura 4 – Publicações sobre a área de estudo entre 1971 e 2018

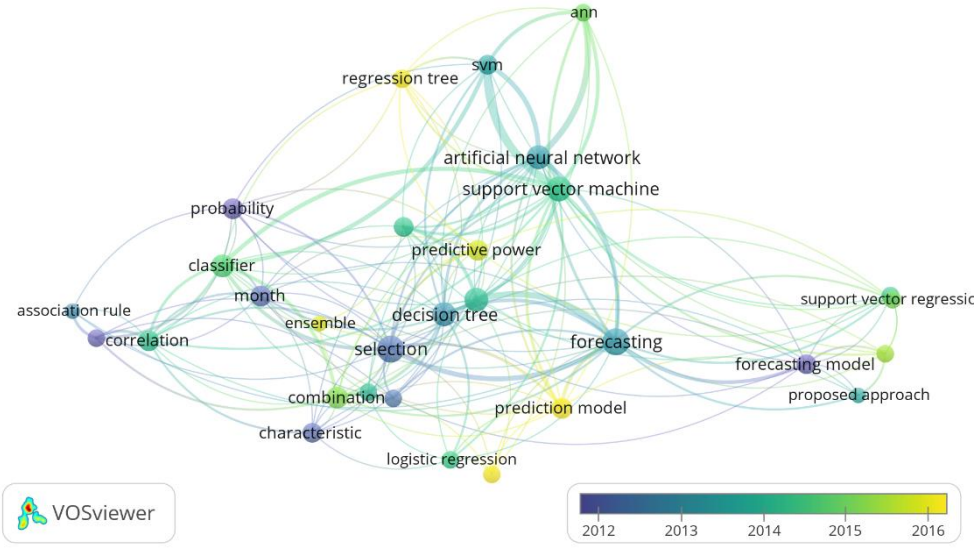
As publicações realizadas com o tema do estudo pesquisado têm crescido substancialmente. Tal fato demonstrado na Figura 4 reflete o interesse dos pesquisadores nesse assunto.

Gerou-se um gráfico do tipo “nuvem de palavras” para obter uma visualização mais nítida das palavras-chave que mais ocorreram nos artigos selecionados. De acordo com a Figura 5 foram evidenciadas palavras como: “*Forecasting*”, “Analysis” e “*Learning*”.

Fonte: Autoria própria.

Figura 5 – Palavras-chave e suas ocorrências

Foi então utilizado o software *VOSviewer* na versão 1.6.13 para elaboração de uma visualização das palavras principais relacionadas à mineração de dados e aprendizagem de máquinas que estavam contidas no título e no resumo dos artigos selecionados e suas correlações. O resultado pode ser visto na Figura 6, onde consta uma escala de cores para ilustrar o decorrer dos anos.



Fonte: Autoria própria (gerado no *VOSviewer*).

Figura 6 – Visualização das palavras mais comuns no título e no resumo

Foi efetuado o processo manual de inspeção e filtragem satisfazendo os critérios de inclusão:

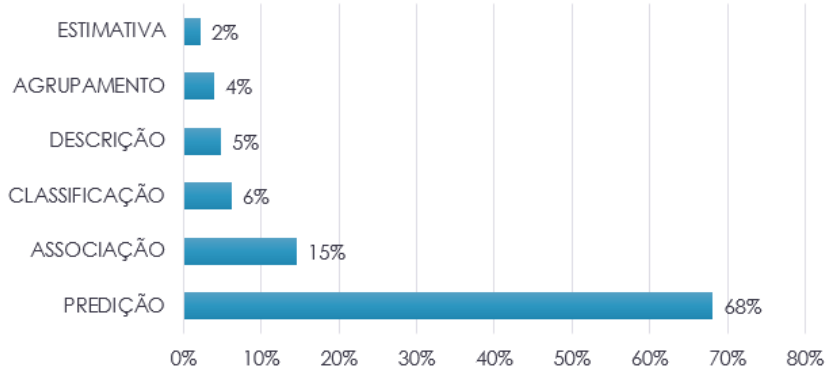
a) descrever uma aplicação específica de DM ou ML no mercado de ações;

b) descrever explicitamente as técnicas de DM ou ML que foram utilizadas;

c) tais critérios serem visíveis por meio da leitura do abstract.

Um total de 189 artigos satisfez os critérios de inclusão.

De acordo com as tarefas de mineração de dados anteriormente definidas, foi possível observar a distribuição conforme a Figura 7.



Fonte: Autoria própria.

Figura 7 – Tarefas de mineração utilizadas nos artigos selecionados

A fim de identificar quais as principais técnicas utilizadas para tratar os dados e métodos de suporte à utilização destas técnicas foi gerada uma matriz de técnicas mencionadas por ano, juntamente com uma escala de cores na forma de mapa de calor. O mapa de calor utiliza cores mais quentes para ressaltar maior frequência, e cores mais frias para demonstrar menor frequência. O resultado pode ser observado no Figura 8.

Técnicas	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	Total Geral
Ant Colony	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	1	1	3
Association Rules	0	0	0	0	0	1	1	0	1	0	0	1	0	1	1	0	1	2	0	0	1	10
Backpropagation	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	4	0	0	0	5
Decision Trees	1	0	0	0	0	0	1	0	1	0	1	1	1	2	0	0	0	2	0	1	0	11
Deep Learning	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	3	7	11
Ensemble	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	2	0	1	1	2	7
Fourier transformation	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	1
Genetic Algorithm	0	1	0	0	0	0	1	1	2	0	2	0	2	0	1	0	1	1	1	2	1	16
K-means	0	0	0	0	0	0	0	1	0	0	1	1	1	0	0	0	0	0	0	1	0	5
K-NN	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	1	0	1	0	3
Logistic Regression	0	0	0	0	0	0	0	1	1	0	0	0	0	0	0	0	1	0	1	0	0	4
Model-Free Estimators	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1
Naïve Bayes	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	1	2
Neural NetworK	1	0	0	1	2	1	3	4	5	4	2	3	2	2	0	3	5	5	9	3	5	60
Nonlinear correlation	0	0	1	0	1	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	2	5
Statistical analysis	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	1
SVR	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	1
Time Series Analisys	0	0	0	0	0	1	2	1	0	2	0	0	1	2	2	1	0	2	0	0	0	14
Suport Vector Machine	0	0	0	0	0	0	0	3	1	1	1	7	1	0	0	1	1	1	4	4	0	25
Multi-scale	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1
Random forests	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	1	2

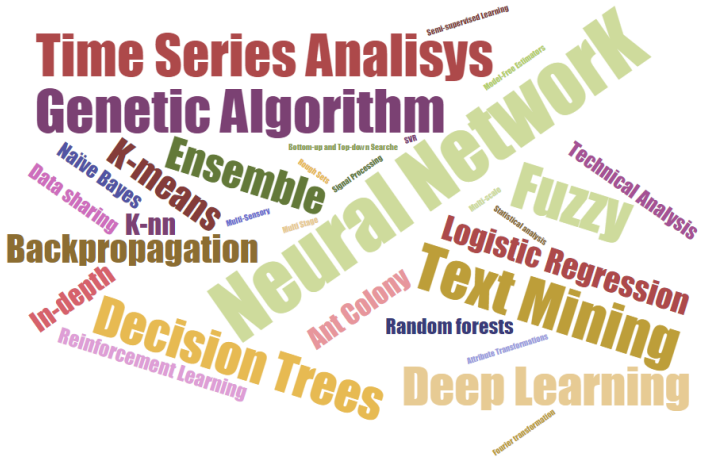
Fonte: Autoria própria.

Figura 8 – Técnicas de mineração utilizadas nos artigos selecionados

Pela matriz, observou-se que a técnica de mineração de dados e aprendizagem e máquina mais utilizada para analisar dados relacionados ao mercado de ações é a aplicação de *Neural Networks* (ou rede neurais). Além disso, vem crescendo substancialmente o uso de *Deep Learning* (ou aprendizado profundo). Tais tendências são explicadas devido ao fato de que a previsão de preços futuros e o entendimento matemático das correlações do mercado de ações envolvem uma série de fatores. Por não se tratar de uma tarefa simples, o uso de

métodos mais complexos de resolução é justificável.

Uma nova nuvem de palavras foi gerada apenas com as técnicas de mineração de dados e alguns métodos de apoio ao tratamento dos dados que foram utilizadas nos artigos selecionados, conforme Figura 9. Foram evidenciados os termos: "*Neural Network*", "Genetic Algotithm" e "*Decision Trees*".



Fonte: Autoria própria.
Figura 9 – Técnicas de mineração e de apoio utilizadas nos artigos selecionados

Por fim, foram selecionados os 4 artigos mais citados da plataforma com mais resultados (*Scopus*) por considerar que são os de maior relevância para o tema. As informações dos 4 artigos podem ser vistas na Figura 10.

CITAÇÕES	ANO	AUTORES	TÍTULO
312	2009	Schumaker, R.P., Chen, H.	Textual analysis of stock market prediction using breaking financial news: The AZFin text system
217	2005	Enke, D., Thawornwong, S.	The use of data mining and neural networks for forecasting stock market returns
197	2005	Boginski, V., Butenko, S., Pardalos, P.M.	Statistical analysis of financial networks
167	2009	Huang, W.-Q., Zhuang, X.-T., Yao, S.	A network analysis of the Chinese stock market

Figura 10 – Resultados os artigos mais citados

Temos uma breve descrição dos artigos a seguir.

Os autores Schumaker e Chen (2009) descrevem suas descobertas experimentais e

explicações de um projeto preditivo que aborda aprendizagem de máquina. É feita a análise de textos de artigos financeiros usando "*bag of words*" e "*proper nouns*" para determinar as palavras-chave, e algoritmos de aprendizagem identificam correlações com os impactos no mercado de ações. É verificada a eficácia da previsão usando notícias e a combinação de técnicas mais valiosa.

Já os autores Enke e Thawornwong (2005) reforçam a não linearidade do mercado de ações. Utilizam o aprendizado de máquina para avaliar as relações preditivas entre variáveis financeiras e econômicas. Foi usada a técnica de redes neurais e desenvolvido um modelo.

Boginski et al. (2005) analisaram estatisticamente os índices do mercado de ações e estudaram as características que representam o gráfico do mercado dos Estados Unidos. Foi aplicada técnica de programação linear inteira nos dados e a análise forneceu uma abordagem alternativa de mineração de dados.

Por fim, Huang et al. (2009) analisaram as correlações existentes no mercado da china. Apresentaram um estudo detalhado da rede que representa o conjunto de ações negociadas na china.

4. Conclusão

O presente artigo teve como objetivo realizar uma revisão da literatura para analisar as técnicas de mineração de dados e aprendizagem de máquina mais utilizadas no mercado de ações. Seguindo o protocolo de busca obteve-se a aprovação da seleção de 189 artigos alinhados ao tema de pesquisa.

Quanto às tarefas de mineração de dados e aprendizagem de máquina utilizadas nas aplicações, a predição foi a que mais se destacou, o que é de se esperar visto que uma das principais vantagens competitivas almejadas no mercado de ações é o ato de prever os preços futuros. A tarefa de associação também obteve destaque, indicando a busca por correlações existentes entre o preço de ações e fatores externos.

Quanto às técnicas e ferramentas utilizadas para manipular os dados, houve um relevante destaque da aplicação de *Neural Networks* (ou redes neurais). Notou-se também o crescente uso de *Deep Learning* (ou aprendizado profundo).

Desta forma, o trabalho cumpriu o objetivo proposto, mostrando o cenário atual de utilização de técnicas de mineração de dados e aprendizagem para problemas relacionados ao mercado de ações e a relevância do tema em questão.

Referências

BERRY, M. J. A.; LINOFF, G. **Data mining techniques : for marketing, sales, and customer relationship management**. 2nd. ed. [S.l.]: Wiley Publishing, Inc, 2004.

CUOCOLO, R. et al. Machine learning applications in prostate cancer magnetic resonance imaging. **European Radiology Experimental**, v. 3, n. 1, p. 35, 2019. Disponível em: <https://eurradiolexp.springeropen.com/articles/10.1186/s41747-019-0109-2>.

ENKE, D., THAWORNWONG, S. The Use of Data Mining and Neural Networks for Forecasting Stock Market Returns. Expert Systems with Applications, 29, 927-940, 2005. https://doi.org/10.1016/j.eswa.2005.06.024

FAYYAD, U. M.; PIATETSKY-SHAPIRO, G.; SMYTH, P. From data mining to knowledge discovery: an overview. **Advances in Knowledge Discovery and Data Mining**, v. 17, n. 3, p.

37–54, 1996.

HUANG, W.Q.; ZHUANG, X.T.; YAO, S. A network analysis of the Chinese stock market. Physica A.,388, 2956–2964, 2009.

KUMAR, R. R. Visualizing Big Data Mining: Challenges, Problems and Opportunities. v. 6, n. 4, p. 3933–3937, 2015.

LAROSE, D. T.; LAROSE, C. D. **Discovering Knowledge in Data: An Introduction to Data Mining**. 2nd. ed. Hoboken, New Jersey: John Wiley & Sons, Inc, 2014.

LEE, A.; TAYLOR, P.; KALPATHY-CRAMER, J. Machine Learning Has Arrived! **Ophthalmology**, v. 124, n. 12, p. 1726–1728, 2017. Disponível em: <https://doi.org/10.1016/j.ophtha.2017.08.046>.

MENDELEY. Página inicial. Disponível em: <https://www.mendeley.com/>. Acesso em: 15 de jun. de 2019.

MITCHELL, T. M. **Machine Learning**. [S.l.]: McGraw-Hill Science/Engineering/Math, 1997.

PINHEIRO, J. L. Mercado de capitais (3a ed.). São Paulo: Atlas, 2005.

SAMUEL, A. L. Some Studies in Machine Learning Using the Game of Checkers. **IBM Journal of Research and Development**, v. 3, n. 3, p. 210–229, 1959. Disponível em: <http://ieeexplore.ieee.org/document/5392560/>.

SCHUMAKER, R. P., CHEN, H. Textual analysis of stock market prediction using breaking financial news: The AZFin text system, **ACM**, 2009.

BOGINSKI V., BUTENKO S., PARDALOS P.M. Statistical analysis of financial networks. **Comput. Stat. Data Anal.**, 48 (2005), pp. 431-443

TRANFIELD, D.; DENYER, D.; SMART, P. Towards a Methodology for Developing Evidence-Informed Management Knowledge by Means of Systematic Review* Introduction: the need for an evidence- informed approach. **British Journal of Management**, v. 14, p. 207–222, 2003.

VALENCIA, F.; GÓMEZ-ESPINOSA, A.; VALDÉS-AGUIRRE, B. Price movement prediction of cryptocurrencies using sentiment analysis and machine learning. **Entropy**, v. 21, n. 6, 2019.

VOSVIEWER. Página inicial. Disponível em: < https://www.vosviewer.com/>. Acesso em: 8 de jul. de 2019.

WITTEN, I. H.; FRANK, E. **Data Mining: Practical Machine Learning Tools and Techniques**. [S.l: s.n.], 2005. Disponível em: <http://books.google.com/books?id=bDtLM8CODsQC&pgis=1>.