

PRINCIPAIS CONCEITOS SOBRE DATA MINING E SUAS TENDÊNCIAS PARA O FUTURO

MAIN CONCEPTS ABOUT DATA MINING AND ITS TENDENCIES FOR THE FUTURE

Danilo Lima Almeida – danilolima345@hotmail.com
 Faculdade de Tecnologia de Taquaritinga (Fatec) –Taquaritinga –SP –Brasil

Giuliano Sombatti Pinto – giuliano.pinto@fatec.sp.gov.br
 Faculdade de Tecnologia de Taquaritinga (Fatec) –Taquaritinga –SP –Brasil

DOI: 10.31510/infa.v22i1.2244
 Data de submissão: 10/04/2025
 Data do aceite: 26/06/2025
 Data da publicação: 30/06/2025

RESUMO

A mineração de dados é um processo crucial para extrair conhecimento útil de grandes quantidades de dados. Ao usar técnicas estatísticas, *machine learning* e inteligência artificial, a mineração de dados permite a identificação de padrões, tendências e relacionamentos ocultos em dados. O objetivo deste artigo é verificar as principais ideias de mineração de dados, cobrindo as fases, funções e tendências. A metodologia utilizada foi a revisão bibliográfica. Os resultados mostraram que o uso da mineração de dados é uma parte fundamental da empresa, principalmente no auxílio de captação de clientes e também na prevenção de fraudes e segurança cibernética. Conclui-se que sem sua utilização, as empresas não conseguiram agir de forma rápida para atender as necessidades dos clientes e aprimorar como as pessoas interagem com as informações, com inovação e tomada de decisões estratégicas.

Palavras-chave: Mineração de dados. *Machine learning*. Inteligência artificial.

ABSTRACT

Data mining is a crucial process for extracting useful knowledge from large amounts of data. By using statistical techniques, machine learning, and artificial intelligence, data mining allows the identification of patterns, trends, and relationships hidden in data. The objective of this article is to examine the main ideas of data mining, covering its phases, functions, and trends. A methodology was used for the literature review. The results demonstrated that the use of data mining is a fundamental part of the company, mainly in helping to attract customers and also in fraud prevention and cybersecurity. It is concluded that without its use, companies would not be able to act quickly to meet customer needs and improve how people interact with information, with innovation and strategic decision-making.

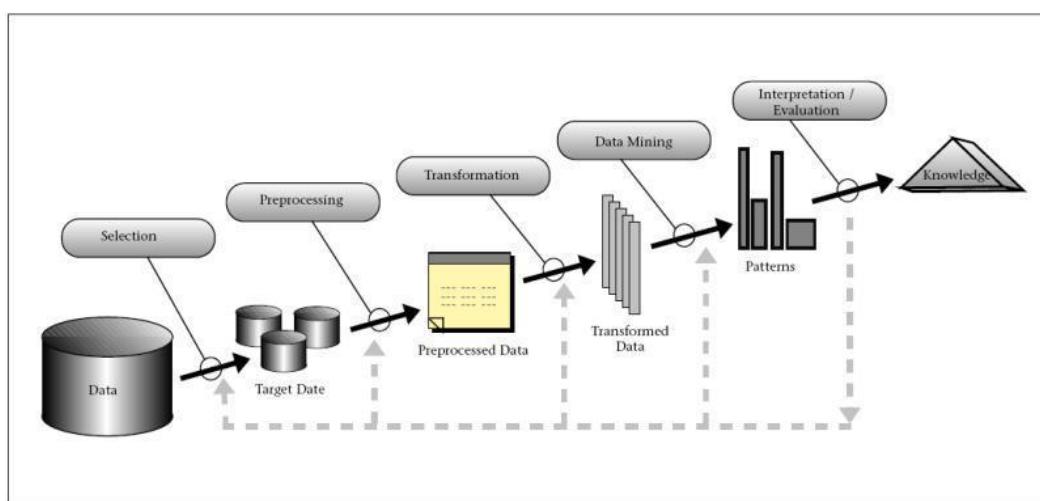
Keywords: Data mining. Machine learning. Artificial intelligence.

1. INTRODUÇÃO

Nos últimos anos, a quantidade de dados gerados por pessoas, empresas e dispositivos conectados à internet cresceu exponencialmente. A digitalização de processos, o avanço da Internet das Coisas (IoT) e o aumento significativo de interações em plataformas online resultaram na geração de trilhões de gigabytes de informação por dia. No entanto, um acúmulo de dados não é suficiente. Essas informações devem ser comprovadas, interpretadas e transformadas em conhecimento acessível. Para isso, surgem métodos avançados de análise de dados como o *Data Mining* ou Mineração de Dados, uma abordagem utilizada para a descoberta de padrões, tendências e *insights* ocultos em grandes volumes de informação.

O *Data Mining*, de forma multidisciplinar, combina inteligência artificial, estatística, aprendizado de máquinas e banco de dados para relacionar variações entre prever comportamentos futuros e também oferecer suporte para decisões baseado nos dados obtidos. Ele se diferencia de abordagens tradicionais de análise ao automatizar a descoberta de conhecimento permitida e escalável. Dessa forma, ele desempenha um papel vital em diversas áreas do conhecimento e setores econômicos (Han; Kamber; Pei, 2011). A Figura 1 apresenta as etapas do KDD (Descoberta de Conhecimento em Base de Dados).

Figura 1 – Etapas do KDD



Fonte: Han; Kamber (2000)

Segundo Holdsworth (2024), o *data mining* é aplicado nas mais variadas áreas. Na saúde, por exemplo, a mineração de dados se aplica em prever o surgimento de doenças, em personalizar tratamentos e também na detecção de anomalias em exames e diagnósticos.

médicos. No setor financeiro, algoritmos avançados são usados na detecção de fraudes, análise de risco de crédito e na predição de tendências para ações do mercado. No marketing digital e varejo, a mineração de dados oferece a possibilidade de entendimento do comportamento de consumo e segmentação de mercado, bem como na personalização de recomendações considerando históricos de compra e interações do usuário. Tais técnicas possibilitam ótimos resultados em outras áreas como segurança cibernética, educação, manufatura e logística.

De acordo com Goldschmidt, Passos e Bezerra (2015), o processo de mineração de dados consiste em várias etapas conectadas entre si, desde a coleta dos dados, sua limpeza e preparação até a aplicação de algoritmos complexos para extrair conhecimento e organizar resultados. Redes neurais artificiais, algoritmos de decisão, algoritmos de clusterização e regras de associação são alguns dos métodos geralmente empregados para encontrar informações significativas que poderiam passar despercebidas apenas pela observação manual de dados.

Ainda para Goldschmidt, Passos e Bezerra (2015), apesar de ter muitas vantagens, há também grandes desafios a serem enfrentados pelo *data mining*. Questões como qualidade de dados, privacidade e ética no uso da informação, complexidade computacional e a necessidade de pessoas qualificadas são questões críticas a serem consideradas. Além disso, legislações como a Lei Geral de Proteção de Dados (LGPD) do Brasil e o Regulamento Geral de Proteção de Dados (GDPR) da União Europeia, oferecem diretrizes para o uso responsável dos dados extraídos, garantindo segurança e transparência no tratamento de dados sensíveis.

Diante disso, o objetivo deste artigo é explorar e examinar as principais ideias da mineração de dados, abordando suas fases, aplicações e tendências futuras. Ao longo do artigo, será possível entender o quanto essa abordagem é relevante para várias áreas de atuação, apoiando a inovação, a competitividade e a transformação digital em diversos setores. Além disso, serão discutidos os desafios e as tendências da mineração de dados, incluindo as regulamentações emergentes e os avanços tecnológicos.

2. MINERAÇÃO DE DADOS

Segundo Han, Kamber e Pei (2011), a mineração de dados é o processo de exploração e análise de grandes volumes de dados, com o intuito de descobrir padrões, relacionamentos e tendências ocultos. Este método permite transformar dados brutos em informações úteis, que poderão ser usadas para decisões estratégicas em vários setores. Combinando muitas disciplinas, incluindo estatística, aprendizado de máquina, inteligência artificial e bancos de

dados, a mineração de dados está se tornando uma ferramenta valiosa para as empresas que competem usando dados.

A mineração de dados vai além da recuperação direta de informações. Este processo utiliza-se de técnicas sofisticadas para prever tendências futuras, descobrir padrões entre variáveis e automatizar a descoberta de conhecimento. Com essa abordagem, empresas e instituições podem compreender melhor seus mercados, otimizar processos e inovar na criação de bens e serviços.

Ainda para Han, Kamber e Pei (2011), os principais objetivos do *Data Mining* incluem:

- **Descoberta de Conhecimento** – Extrair padrões e informações úteis a partir de grandes volumes de dados.
- **Previsão de tendências (predição)**: Utilizando modelos estatísticos e algoritmos para prever eventos futuros com base em dados históricos.
- **Classificação e categorização**: Melhorando a compreensão e a análise, agrupando conjuntos de dados semelhantes.
- **Análise de anomalias**: Encontrar desvios e irregularidades que podem apontar para fraudes ou problemas operacionais.

2.1 Aplicações do *Data Mining*

Conforme Holdsworth (2024), a mineração de dados é utilizada em diversos campos, promovendo inovações e melhorando o processo de tomada de decisão. Alguns dos campos mais importantes em que a mineração de dados é empregada são:

2.1.1 Saúde

Na medicina, a prática da mineração de dados visa explorar grandes quantidades de dados médicos e epidemiológicos, possibilitando a identificação das doenças em seus estágios iniciais, a personalização do tratamento e a otimização da gestão hospitalar. Os modelos preditivos favorecem a detecção de anormalidades em diagnósticos clínicos, otimizando o diagnóstico e melhorando a precisão do tratamento.

A aplicação de algoritmos mais sofisticados permite também elaborar análises de dados genéticos, locais de origem das predisposições genéticas e do desenvolvimento de tratamentos personalizados.

2.1.2 Finanças

O setor financeiro utiliza mineração de dados para a análise de crédito, detecção de fraudes e previsão de tendências econômicas. Além de atuar na prevenção de crimes financeiros e na identificação de transações, algoritmos sofisticados são utilizados para personalizar ofertas de produtos e serviços bancários, com base no comportamento do cliente.

Além disso, técnicas de análise preditiva como apoio na modelagem do risco e na construção de estratégias de investimento.

2.1.3 Varejo e Marketing

Empresas que lidam com vendas e marketing utilizam a mineração de dados para compreender o comportamento dos consumidores e melhorar suas estratégias de vendas. Os algoritmos de recomendação estudam o histórico de compras e navegação do consumidor, de modo a personalizá-las, aumentando assim a fidelidade do consumidor.

A mineração de dados é também utilizada para prever tendências do consumidor e melhorar campanhas de marketing digital, orientando análises baseadas nas preferências e interesses do usuário.

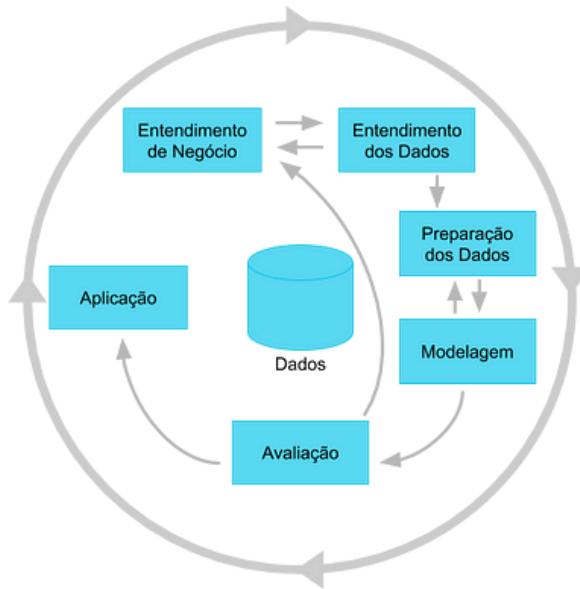
2.1.4 Segurança Cibernética

A mineração de dados colabora na detecção de atividades maliciosas e na proteção contra ataques cibernéticos, especificamente no caso da segurança cibernética. Sistemas com aprendizado de máquina analisam quantidades massivas de tráfego da rede para detectar atividades maliciosas e bloquear sites maliciosos em tempo real.

As técnicas como a análise de log e a detecção de anomalias são fundamentais para a segurança do sistema corporativo e evitam vazamentos de dados confidenciais.

A Figura 2 apresenta os passos da mineração de dados de modo geral.

Figura 2 - Passo a passo da mineração de dados



Fonte: *The CRISP-DM Consortium (2000)*

2.2 Desafios do *Data Mining*

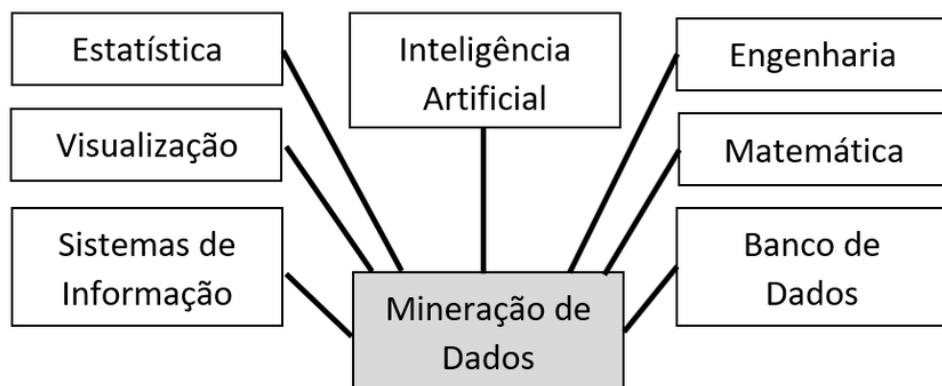
Apesar de seus inúmeros benefícios, a aplicação do *Data Mining* enfrenta diversos desafios, incluindo:

- **Qualidade dos dados:** A qualidade dos dados analisados afeta diretamente a precisão dos resultados. Dados inconsistentes, incompletos ou redundantes podem complicar as análises e levar a conclusões ambíguas, exigindo técnicas eficazes de limpeza e preparação (Objective, 2024).
- **Privacidade e ética:** A coleta e utilização massiva de dados levantam preocupações relativas à privacidade e proteção de informações sensíveis, exigindo a conformidade com legislações como GDPR e LGPD. O uso indevido desses dados pode ter efeitos danosos para pessoas e organizações (Objective, 2024).
- **Complexidade computacional:** Algoritmos de mineração de dados demandam um alto consumo de recursos computacionais e por isso muitas vezes necessitam de infraestruturas sofisticadas, como processamento distribuído e computação em nuvem, para que possam lidar com a grande quantidade de dados de forma eficaz (Objective, 2024).

- **Necessidade de profissionais especializados:** A interpretação dos resultados exige conhecimento técnico baseado em estatística, aprendizado de máquina e ciência de dados, o que é fundamental para a formação contínua dos profissionais da área (Objective, 2024).

Dessa forma, superar esses desafios é fundamental para assegurar a eficácia e a responsabilidade da Mineração de Dados, promovendo a inovação sustentável orientada por dados. A Figura 3 apresenta a multidisciplinaridade da Mineração de Dados, combinando técnicas de análise de dados com conhecimentos de áreas diversas.

Figura 3 – A multidisciplinaridade de Mineração de Dados



Fonte: Ramos (2016)

3. METODOLOGIA

A metodologia utilizada no presente artigo foi a revisão bibliográfica. Para esse estudo, foram utilizadas consultas em livros, artigos, monografias, websites e outros materiais acadêmicos a fim de levantar informações e sintetizar a compreensão dos principais conceitos sobre o *Data Mining* e suas tendências para o futuro.

Através de pesquisa exploratória, foi possível identificar os materiais mais relevantes de acordo com o tema proposto através de buscas no Google Acadêmico. Para a pesquisa, foram utilizados os termos “principais conceitos sobre mineração de dados”, “mineração de dados e tendências”, “o futuro da mineração de dados”, “empresas que usam mineração de dados”. Em seguida, os materiais encontrados foram analisados e separados os tópicos de maior relevância para o prosseguimento do artigo, bem como empresas renomadas no mercado como Amazon, que utilizam a mineração de dados. O total de trabalhos encontrados foi de 18 e, após esse processo de análise, selecionou-se 4 artigos que correspondiam com a proposta do presente

trabalho. Posteriormente, após o levantamento, foram realizados breves resumos dos tópicos para ajudar na divisão e correlação do tema.

Por fim, após a análise completa em todo o material selecionado, foram elaborados os resultados em consonância com a fundamentação teórica, bem como as considerações finais.

4. RESULTADOS E DISCUSSÃO

O desenvolvimento da mineração de dados permitiu descobertas e inovações em uma variedade de campos, possibilitando conseguir *insights* valiosos no âmbito da tecnologia, saúde, segurança e negócios.

4.1 Tendências para o futuro do *Data Mining*

Na visão de Goldschmidt, Passos e Bezerra (2015), o futuro do *Data Mining* está intimamente relacionado a tecnologias emergentes como inteligência artificial (IA), computação quântica e *big data analytics*. Com o aumento da complexidade dos dados produzidos, novos algoritmos estão sendo desenvolvidos para processar melhor esses dados, reduzindo o tempo de processamento e aumentando assim a precisão da análise.

Uma das crescentes tendências futuras do *Data Mining* é a análise dos dados em tempo real. À medida que a Internet das Coisas (IoT) continua a crescer e a quantidade de dados gerados continuamente aumenta, será necessário dispor de algoritmos que analisem dados instantaneamente, em setores como segurança cibernética, atendimento ao cliente e monitoramento industrial. A mineração localiza anormalidades rapidamente, o que permite uma resposta rápida a ataques, falhas operacionais, assim como a mudanças de comportamento do consumidor.

Segundo Witten, Frank e Hall (2016), o aprendizado profundo e técnicas de *machine learning* têm revolucionado a mineração de dados e superado modelos mais precisos e eficientes. As redes neurais profundas podem lidar com dados não estruturados como imagens, vídeos e texto, permitindo análises mais sofisticadas exigidas em áreas como diagnóstico médico, processamento de linguagem natural e visão computacional. Além disso, a automação na mineração de dados diminuirá a necessidade de intervenção humana, tornando os procedimentos analíticos mais acessíveis e mais eficientes em termos de idade.

A popularidade da IA explicável (*Explainable AI*) está crescendo à medida que as regulamentações relativas à proteção de dados se tornam mais rígidas. Essa abordagem busca desenvolver modelos de mineração de dados mais interpretáveis e compreensíveis para os

usuários, bem como assegurar a audibilidade e a justificativa das decisões algorítmicas. A tendência tem sido especialmente relevante em áreas regulamentadas como finanças e saúde, onde a transparência é um elemento central das tomadas de decisões.

Quanto à computação quantitativa, é provável que uma revolução na mineração de dados poderá ser realizada, na medida em que ela permitirá processar grandes quantidades de dados, de modo significativamente mais rápido do que pelos métodos tradicionais. Os algoritmos quantitativos podem lidar com problemas complexos de otimização e análise de padrões, tornando a mineração de dados ainda mais poderosa.

Apesar do fato dessa tecnologia ainda estar em sua fase de pesquisa, espera-se que, em um futuro próximo, sua sinergia com a mineração de dados forneça um progresso considerável para modelagem de dados e predição.

O crescimento da mineração de dados traz também desafios em termos de regulamentação e ética. A utilização de grandes volumes de dados pessoais gera preocupações quanto à segurança da informação, consentimento e privacidade. Regulamentações, como o GDPR na Europa e a LGPD no Brasil, impõem restrições ao uso de dados sensíveis e exigem das organizações um maior nível de transparência e regulamentação. Novas regulamentações provavelmente irão surgir no futuro, buscando proporcionar um equilíbrio entre inovação e proteção dos direitos dos usuários.

Coerente com Holdsworth (2024), no âmbito do marketing digital e varejo, o futuro da mineração de dados está associado a recomendações super personalizadas. Algoritmos complexos serão capazes de explorar dados contextuais e comportamentais de maneira muito mais apurada, oferecendo experiências personalizadas em tempo real. As empresas serão capazes de prever as necessidades dos consumidores e fornecer produtos e serviços altamente focados utilizando análise preditiva.

A capacidade de lidar com grandes quantidades de informação de maneira confiável e eficiente transforma a capacidade das empresas de conseguir tomar decisões estratégicas (Han; Kamber; Pei, 2011).

4.2 Perspectivas para o futuro do *Data Mining*

Entretanto, apesar de inúmeras aplicações e vantagens, a mineração de dados ainda enfrenta desafios significativos. Questões como privacidade, segurança de dados e transparência de algoritmos ainda estão sendo abordadas de forma mais profunda. A criação de

métodos mais rigorosos de regulamentação e mais transparentes e auditáveis é um dos componentes para garantir a utilização da tecnologia de forma ética (Amazon, 2025).

O crescimento contínuo do *Data Mining* e sua integração com tecnologias emergentes, incluindo inteligência artificial e computação quântica, indicam um futuro otimista para o domínio. A busca por maior eficiência e eficácia da análise de dados continuará a promover inovações, assegurando que a mineração de dados se torne uma ferramenta cada vez mais essencial para a sociedade e para os negócios (Witten; Frank; Hall, 2016).

Empresas como Netflix, Amazon e Uber implementam com sucesso a mineração de dados em suas estratégias de negócio. Sua implementação tem etapas como definição de objetivos, organização e coleta de dados, escolhas assertivas nas técnicas, interpretação dos resultados e implementação de suas ações conforme suas necessidades. São utilizadas técnicas e ferramentas adequadas para cada negócio de forma eficaz como linguagem de programação, ferramentas de visualização, algoritmos de aprendizado de máquinas e também o uso do processamento de *Big Data*. Com o uso da mineração de dados, essas empresas conseguem se destacar em seus respectivos mercados de forma objetiva, aproveitando os benefícios dessa prática para impulsionar seu sucesso e crescimento (Awari, 2023).

5. CONSIDERAÇÕES FINAIS

A mineração de dados surgiu como um instrumento decisivo para a extração de conhecimento a partir de conjuntos massivos de dados, levando a avanços substanciais em vários campos de pesquisa científica e setores econômicos. À medida que a inteligência artificial, computação quântica e mineração de dados continuem a progredir em tempo real, suas aplicações se tornarão mais eficazes e precisas.

Entretanto, juntamente com essas oportunidades também aparecem significativos desafios, tais como privacidade de dados, conformidade regulamentar e interpretações de modelos de análise. Portanto, é vital que empresas e instituições possuam agentes de endereçamento responsáveis e inovadores se desejarem que a mineração de dados seja utilizada de maneira ética e eficaz.

O futuro da mineração de dados promete, ainda mais, modificar nossa interação com as informações, transformando-a em um elemento essencial para a inovação e a tomada de decisões estratégicas em um mundo cada vez mais impulsionado por dados.

REFERÊNCIAS

- AWARI. **Empresas que usam Mineração de Dados: Exemplos de Sucesso.** 30 jul. 2023. Disponível em: <https://awari.com.br/empresas-que-usam-mineracao-de-dados-exemplos-de-sucesso/#:~:text=Empresas%20como%20Amazon%2C%20Netflix%2C%20Uber,em%20suas%20estrat%C3%A9gias%20de%20neg%C3%B3cios.> . Acesso em 31 mar 2025.
- AMAZON. **O que é mineração de dados?** Disponível em: <https://aws.amazon.com/pt/what-is/data-mining/> . Acesso em 31 mar. 2025.
- GOLDSCHMIDT, R.; PASSOS, E.; BEZERRA, E. **Data mining: conceitos, técnicas, algoritmos, orientações e aplicações.** 15 mai. 2015. Rio de Janeiro: Elsevier, 2015.
- HAN, J.; KAMBER, M.; PEI, J. **Data mining: concepts and techniques.** 6 jul. 2011. 3. ed. Waltham: Morgan Kaufmann, 2011.
- HOLDSWORTH, J. **O que é Mineração de Dados?** 28 jun. 2024. Disponível em: <https://www.ibm.com/br-pt/topics/data-mining> . Acesso em 31 mar 2025.
- OBJECTIVE. **Data Mining: como melhorar a experiência do cliente com análise de dados.** 14 nov. 2024. Disponível em: <https://www.objective.com.br/insights/data-mining/> . Acesso em 11 mar 2025.
- WITTEN, I. H.; FRANK, E., HALL, M. A. **Data Mining: Practical Machine Learning Tools and Techniques.** 17 nov. 2016. Morgan Kaufmann Publishers, 2016.