

## MINERAÇÃO DE DADOS (KNOWLEDGE DISCOVERY IN DATABASES - KDD)

Possível mineração de dados e sobre os estudos, que dados analisados dentro do processo KDD a Mineração de Dados possui o papel principal, o qual define a utilidade. De acordo com Castro e Ferrari, o avanço da tecnologia relacionados ao KDD está ligado a automatização de sistemas. Tendo vista que a mineração de dados abre espaço para o crescimento de novas ferramentas para gerenciar, analisar e, eventualmente, obter vantagem sobre a inundação de dados que a sociedade moderna enfrenta. A competitividade de mercado por grandes corporações e o auxílio na tomada de decisões empresariais, cada vez mais presentes na sociedade atual. Conhecer e entender a mineração de dados e sua aplicabilidade na geração de conhecimento relevante, através da aplicação processo de KDD (Descoberta de Conhecimento em Bases de Dados). Além disso, apresentamos nossa perspectiva de uma estrutura unificada centrada no processo para o KDD. O objetivo é fornecer uma visão geral da variedade de atividades neste campo multidisciplinar e como eles se encaixam juntos. O KDD inclui atividades multidisciplinares. Isso abrange armazenamento e acesso a dados, escalonamento de algoritmos para conjuntos de dados massivos e interpretação de resultados. A limpeza de dados e o processo de acesso aos dados incluídos no data warehousing facilitam o processo do KDD. A inteligência artificial também suporta o KDD, descobrindo leis empíricas de experimentação e observações. Os padrões reconhecidos nos dados devem ser válidos em novos dados e possuem algum grau de certeza. Esses padrões são considerados novos conhecimentos. As etapas envolvidas em todo o processo do KDD são: 1. Selecionar dados - Nesta fase é necessário definir o domínio da aplicação, que inclui de preferência um nível elevado de conhecimento prévio como também identificar o objetivo do processo do KDD a partir da perspectiva do cliente. A fase de seleção dos dados é quando se inicia o processo de descobrimento de informação. É selecionado o conjunto de dados, pertencente a um domínio, contendo todas as possíveis variáveis e registros que farão parte da análise. Normalmente a escolha dos dados fica a critério de um especialista do domínio. Este processo de seleção é bastante complexo, pois os dados podem vir de uma série de fontes diferentes e podem estar nos mais diversos formatos. Por isso, esse passo é fundamental e de grande impacto sobre a qualidade do resultado do processo. Por mais que dados sejam gerados em quantidades cada vez mais exponenciais, esses mesmos dados sem conhecimento pertinente são inúteis, por isso faz-se tão importante a aplicação de cada passo da metodologia do KDD na descoberta de informações, lançando mão de algoritmos de Data Mining a fim de encontrar padrões nos dados. Dentro do processo de KDD a Mineração de Dados possui o papel principal, o qual define a utilidade dos dados analisados. O KDD abre espaço para o crescimento de novas ferramentas para gerenciar, analisar e, eventualmente, obter vantagem sobre a inundação de dados que a sociedade moderna enfrenta. O crescimento dos estudos relacionados ao KDD está ligado a automatização de sistemas, o avanço no uso da inteligência artificial, a competitividade de mercado por grandes corporações e o auxílio na tomada de decisões empresariais, cada vez mais presentes na sociedade atual. Como por exemplo, a captação de dados de um usuário de redes sociais para divulgação de publicidade de acordo com suas preferências, pode ser considerado uma forma de descoberta de conhecimento através de bases de dados.

**Referências:**

**LIFSCHITZ S.; CÔRTEZ, S.; PORCARO, R. Mineração de Dados, Funcionalidades, Técnicas e Abordagens. INSS, PUC-Rio 2002**

**OLIVEIRA, D. Data mining ganha espaço na estratégia empresarial. 2012. Disponível em: < <http://computerworld.com.br/tecnologia/2012/03/16/data-mining-ganha-espaco-na-estrategia-empresarial/>> Acesso em: 15 de maio de 2019.**

**GOLDSCHMIDT, R. e BEZARRA, E., 2015. Data Mining. Elsevier. Rio de Janeiro. Brasil.**

**MORIK, K. The Representation Race. Reprocessing for Handling Time Phenomeno. Proceeding of the European Conference on Machine Learning 2000, Lecture Notes in Artificial Intelligence 1810. Springer-Verlag Berlin Heidelberg, 2000.**

**CASTRO, L. N. e FERRARI, D. G., 2016. Introdução a Mineração de Dados. Editora Saraiva. São Paulo. Brasil.**

**FAYYAD, U.; PIATETSKY-SHAPIRO, G.; SMYTH, P. From Data Mining to Knowledge Discovery in Databases. American Association for Artificial Intelligence, 1996.**

**Descritores:**Anderson Russyel Silva Batista; Gabriel de Lima Sousa; Marcelo Rodrigues de Lima.