O que é mineração de dados?

A mineração de dados ou *data mining* é uma estratégia que busca extrair as informações mais relevantes como tendências, relações e anomalias úteis de uma ou mais bases de dados. Para esse processo, as empresas utilizam ferramentas especializadas que exploram um volume grande de informações para encontrar padrões ou tirar *insights* que irão ajudar nas tomadas de decisão dos gestores. Tudo isso com auxílio de inteligência artificial, machine learning e estatística.

Qual a importância da mineração de dados?

A mineração de dados tem enorme importância estratégica para uma empresa, permitindo que entenda de modo mais contextualizado e preciso os comportamentos dos consumidores e os movimentos de mercado.

Ao compreender os aspectos que influenciam o comportamento e as decisões do cliente, é possível se adaptar rapidamente, aprimorando suas ofertas de produtos ou serviços para angariar resultados positivos.

Além disso, é possível ter em mãos insights que ajudam você a analisar seu banco de dados, entendendo-o e permitindo que você se livre de todas as informações desnecessárias.

Assim, é possível melhorar continuamente seu processo de tomada de decisão.

Fases do KDD:

Seleção: decidir quais conjuntos de dados serão escolhidos para extrair informações úteis.

Limpeza: onde ocorre uma espécie de filtragem de dados, onde os dados ausentes, errôneos ou inconsistentes nas bases de dados devem ser corrigidas de forma a não comprometer a qualidade dos dados, e dados inválidos serão excluídos.

Enriquecimento: onde pode-se melhorar a qualidade de dados com informações adicionais de outras fontes.

Transformação ou Codificação de Dados: é a organização dos dados de modo a ficar mais conveniente realizar a mineração de dados.

Data mining (mineração de dados):

interpretação/avaliação: esta fase consiste em interpretar os dados gerados e verificar se possuem informações úteis como padrões e relações e incorporá-las no sistema.

Data warehouse:

Data warehouse é um termo do inglês que significa repositório de dados, ele centraliza e consolida grandes quantidades de dados que fluem de sistemas transnacionais, bancos de dados (pode haver vários) entre outras várias fontes para o data warehouse. Ele funciona como uma coleção de dados que foi acumulado ao longo do tempo, podendo ter tanto dados estruturados quanto dados não estruturados.

Relação data warehouse e data mining:

Como o data warehouse é um armazenador de dados que centraliza grandes quantidades de dados de diversas fontes, é mais fácil aplicar as técnicas de data mining para conseguir bons resultados e informações relevantes, que podem auxiliar na tomada de decisões e ajudar a ter expectativas para o futuro.

Pesquisar os seguintes tópicos:

* ~~O que é mineração de dados?~~
* ~~Qual a importância da mineração de dados?~~
* ~~Tipos de dados~~
* ~~Data mining~~
* ~~Data warehouse~~
* ~~Relação Data mining e Data warehouse~~
* ~~Fases do KDD~~
* ~~Resultado da Mineração~~
* ~~Aplicações~~
* ~~Objetivos~~
* ~~Conhecimento~~

Fazer uma aplicação prática

Bibliografia:

* <https://blog.betrybe.com/tecnologia/mineracao-de-dados/>
* <https://blog.unyleya.edu.br/inicie-sua-carreira/mineracao-de-dados/#O_que_e_mineracao_de_dados>
* <https://netsupport.com.br/mineracao-de-dados/>
* <https://blog.cedrotech.com/mineracao-de-dados-o-que-e>
* <http://www.ufopa.edu.br/lsd/index.php?option=com_content&view=article&id=8:beginners&catid=19&Itemid=260>
* <https://www.totvs.com/blog/negocios/mineracao-de-dados/#:~:text=todas%20as%20aplica%C3%A7%C3%B5es.-,Qual%20a%20import%C3%A2ncia%20da%20minera%C3%A7%C3%A3o%20de%20dados%3F,e%20os%20movimentos%20de%20mercado>.
* <https://distrito.me/blog/data-mining/>
* <https://danielteofilo.wordpress.com/2015/02/16/kdd-knowlegde-discovery-in-database/>
* https://www.trinapse.com.br/blog/utilizar-data-warehouse-e-data-mining/
* https://aws.amazon.com/pt/data-warehouse/