



O que é a classificação de dados?

O que é a classificação de dados?

Classificação de dados - Mineração de dados



Dentro do âmbito da mineração de dados, a classificação é uma das tarefas mais conhecidas e utilizadas porque através dela é possível rotular dados segundo suas características e estabelecer relações que permitam identificar o rótulo de outros dados, estes desconhecidos.

O que é a classificação de dados?

O que é a mineração de dados?



A mineração de dados é o processo de realizar análises em grandes conjuntos de dados para extrair novos conhecimentos, através dos seus padrões de comportamento, o que é realizado aproveitando métodos computacionais.

Então, quando se utiliza a tarefa de classificação em um grande conjunto de dados significa que queremos analisar os dados disponíveis para entender as relações existentes entre eles e criar um modelo que descreva as relações identificadas.

O que é a classificação de dados?

Qual o papel da mineração e classificação de dados?

- Explorar ao máximo as informações disponíveis em bases de dados;
- Apoiar o processo de tomada de decisão de gestores;
- Criar novas informações e conhecimento que signifiquem vantagem de mercado.

O que é a classificação de dados?

De que forma a mineração de dados apoia uma empresa?

- A mineração de dados é uma das etapas do processo de KDD (*Knowledge Database Discovery* - Descoberta de conhecimento em base de dados);
- O KDD é um processo dividido em etapas e projetado para explorar ao máximo o potencial de geração de conhecimento de bases de dados;
- As bases de dados guardam muitas informações que não são percebidas pelo olhar humano;
- O processo é iterativo e permite que as bases sejam exploradas de diferentes formas dentro de uma empresa.

O que é a classificação de dados?

O processo de KDD

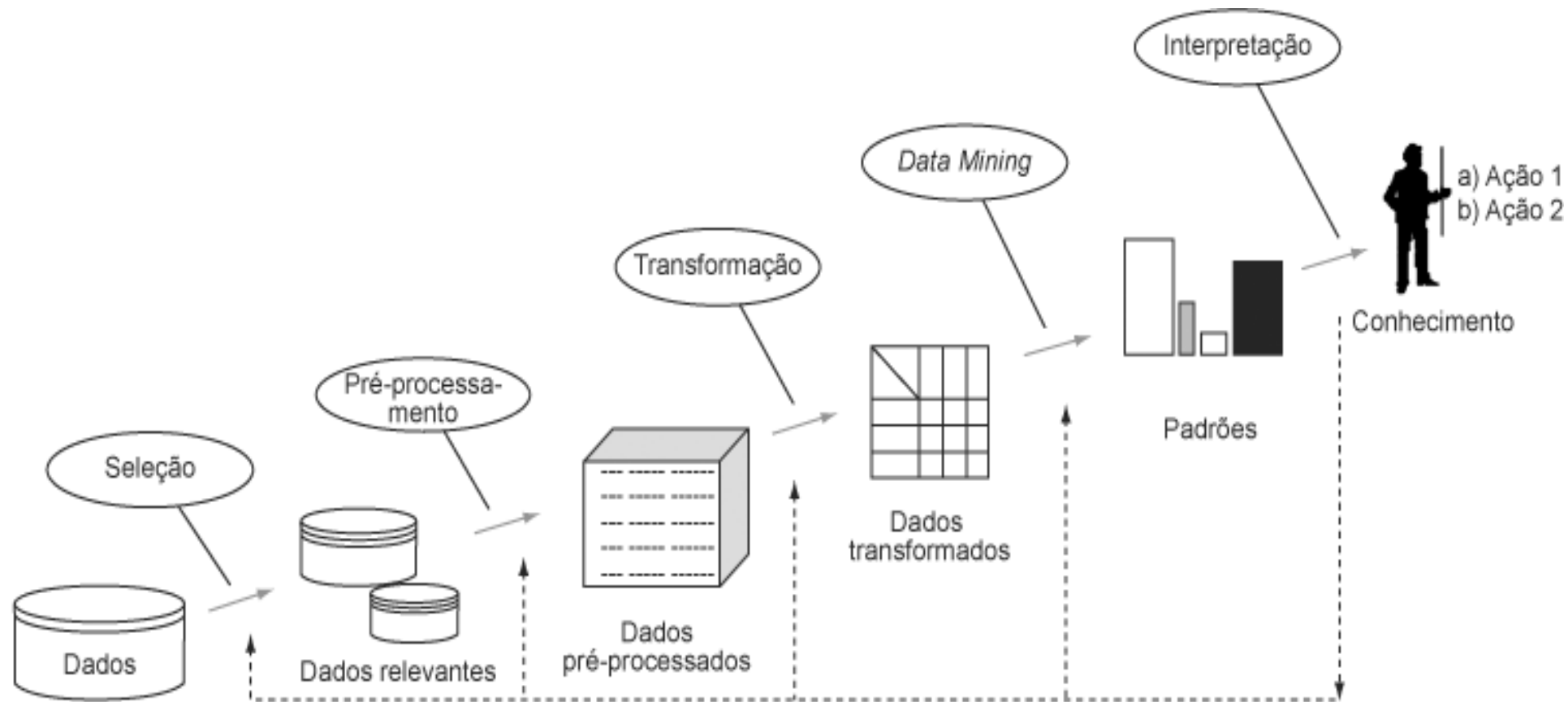


Figura 1. Etapas do processo *KDD* (Fayyad et al. (1996)).

O que é a classificação de dados?

A classificação de dados



A classificação de dados, portanto, faz parte de um processo amplo e diz respeito a exploração direta das bases de dados disponibilizadas para a construção de modelos que irão facilitar, posteriormente, a extração de novos conhecimentos, analisando dados desconhecidos e classificando-os.

O que é a classificação de dados?

Referências

Fayyad, U., Shapiro, G. P., and Smyth, P. “From data mining to knowledge discovery in databases”. AI Magazine, p. 37–54, 1996.



Obrigada!

Ana Laurentino