

Data Mining:

Conceitos e Aplicações

Apresentação

- **Alexandra Miguel Raibolt da Silva**
 - <Laboratório Nacional de Computação Científica>
 - Bolsista de Iniciação Científica.
 - <FAETERJ>
 - Cursando Tecnologia da Informação & Comunicação.
 - <Estácio de Sá>
 - Cursando Sistemas da Informação.
- **Contato:** alexandra.raibolt@gmail.com

Sumário

- Conceitos;
- Knowledge Discovery in Databases (KDD);
 - Etapas KDD;
- Como implementar?;
 - Redes Neurais;
- Aplicações potenciais;
- Cases de Sucesso;
- Ferramentas.

Conceitos

“A mineração de dados é o processo de busca de dados para padrões anteriormente desconhecidos e muitas vezes usando esses padrões para prever os resultados futuros” (JONAS; HARPER, 2006, p. 1).

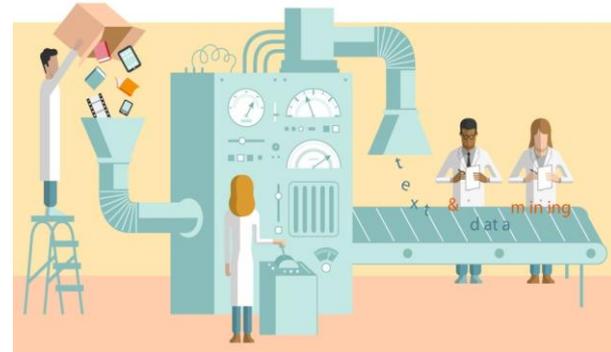
- A expressão Data Mining se refere a minerar/extrair conhecimento de um conjunto de dados;
- É o processo de explorar e analisar grandes quantidades de dados que permite compreender o comportamento dos dados.

Conceitos

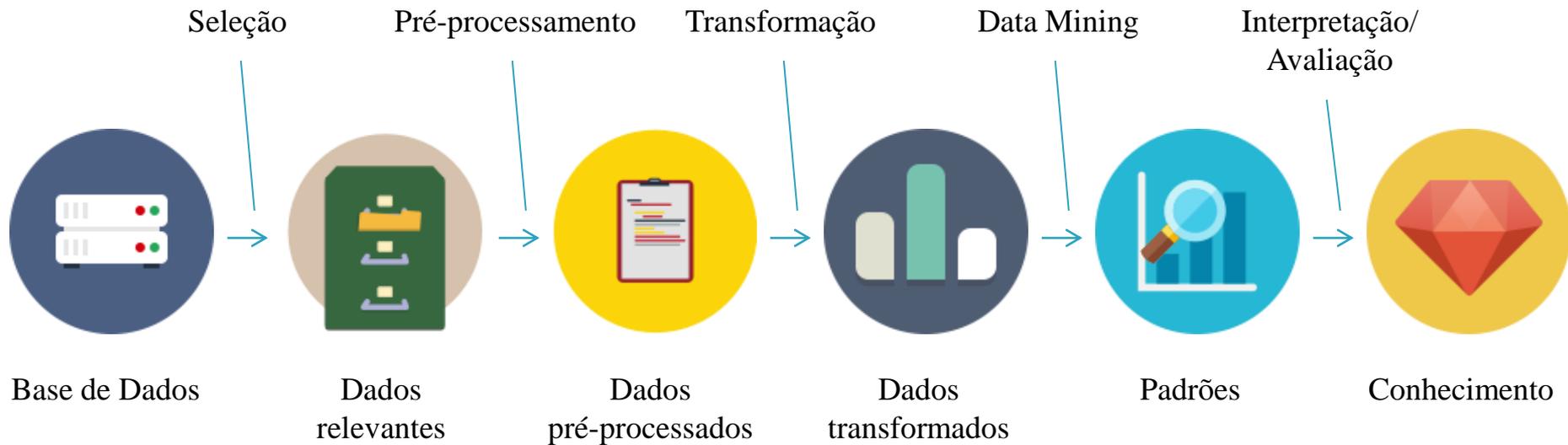
- Tem o objetivo de descobrir regras, padrões consistentes, associações, anomalias, mudanças, identificar fatores e tendências-chave para auxiliar na tomada de decisões sobre estratégias e vantagens competitivas;
- Pode ser aplicado em diversas áreas de negocio, pois as maiorias das empresas utilizam-se de ferramentas que geram dados;
- Aplica-se técnicas e algoritmos para descobrir novos conhecimentos, como por exemplo: Estatística, Redes neurais, Reconhecimento de padrões, entre outros.

Knowledge Discovery in Databases

- KDD (Knowledge Discovery in Databases) - Descoberta de Conhecimento em Bases de Dado;
- É empregado para descrever todo o processo de extração de conhecimentos dos dados, possuindo várias etapas interligadas;
- Apresenta uma definição parecida ao Data Mining, porém o Data Mining é a etapa de análise do KDD;
- **Resumindo:** O KDD compreende todo o ciclo que o dado percorre até virar informação.

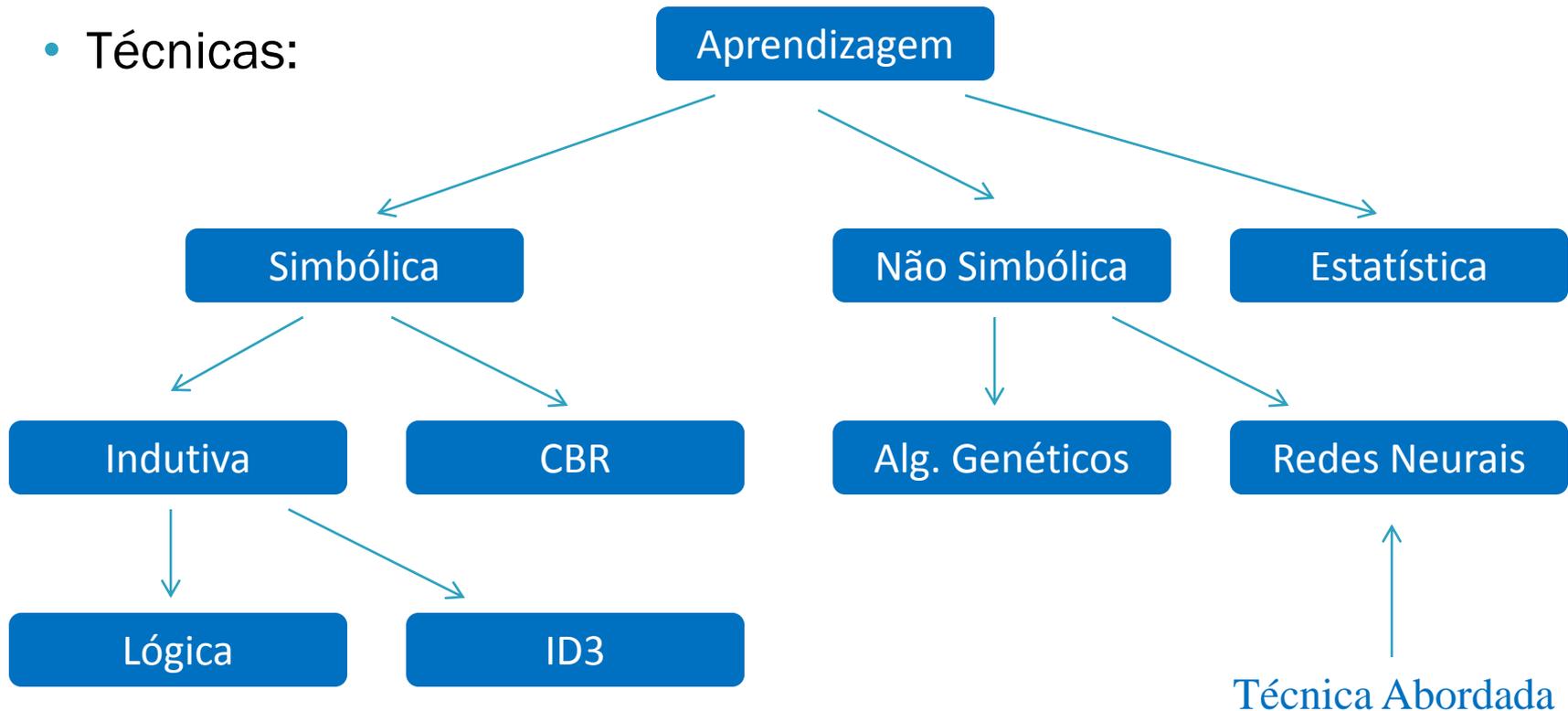


Etapas KDD



Como implementar?

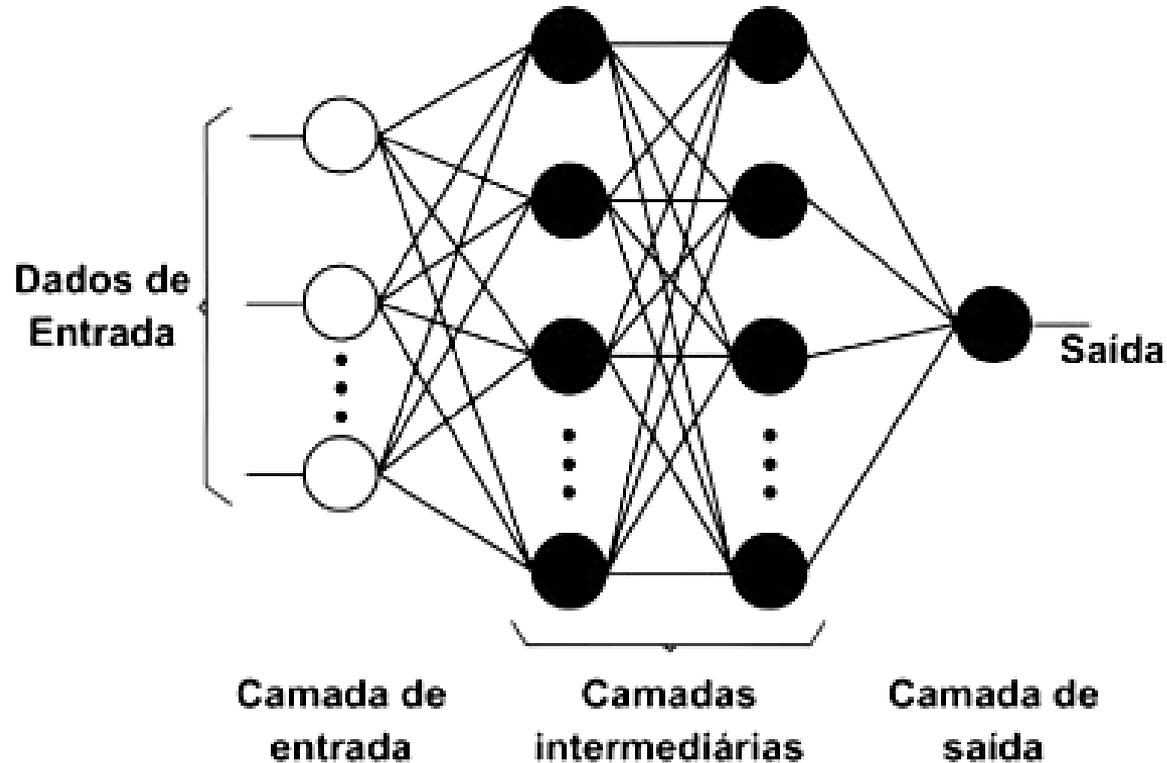
- Técnicas:



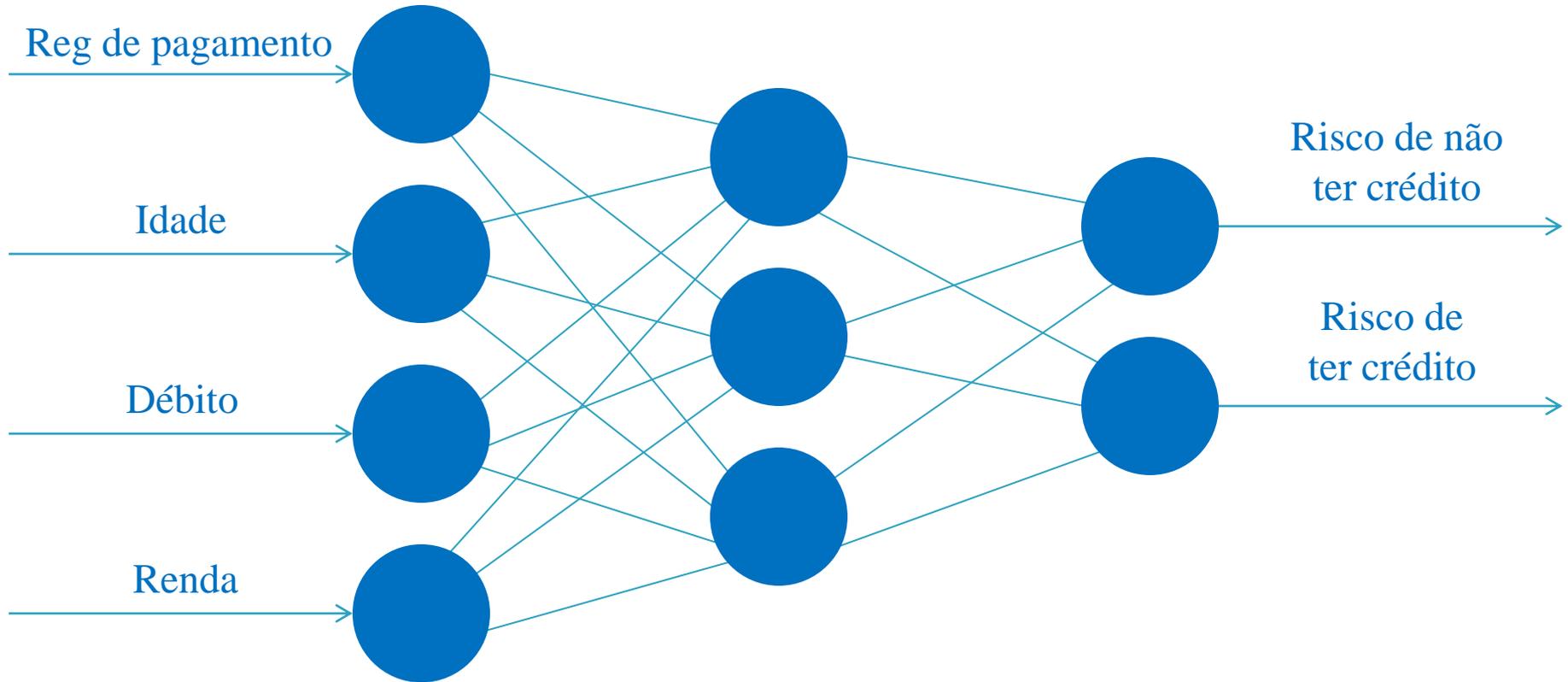
Redes Neurais

- Rede Neural (RN), ou Rede Neural Artificial (RNA) é uma técnica computacional que apresenta um modelo matemático inspirado no sistema nervoso central que é capaz de aprender (aprendizado de máquina);
- Sua principal característica é a grande capacidade de reconhecer padrões e também sua eficiência para classificações de padrões;
- Consiste em um número de elementos interconectados (neurônios/nós), que possuem entrada, processamento e saída;
- Uma grande rede neural artificial pode ter centenas ou milhares de unidades de processamento.

Redes Neurais



Redes Neurais



Aplicações potenciais

- **Análise de mercado:**
 - Descobrir fatores associados a perda de clientes;
 - Determinar o padrão de consumo;
 - Identificar associação entre produtos;
 - Traçar perfil dos clientes;
 - Achar grupos que compartilham as mesmas características ou interesses;
 - Identificar potenciais clientes para determinados produtos;
 - Identificar consumidores “leais”.

Aplicações potenciais

- **Análise de riscos:**
 - Planejamento financeiro;
 - Planejamento de recursos;
 - Monitorar a direção do mercado.
- **Detecção de fraude:**
 - Serviços de cartão de crédito;
 - Seguro;
 - Transações de dinheiro suspeitas;
 - Plano de saúde.

Cases de Sucesso

- **Amazon.com:** Melhoria da customização da interface com o usuário (melhoria de vendas por indicação), eliminação de fraudes;
- **1-800-FLOWERS.com:** Compreensão e antecipação de comportamento de clientes;
- **Harrah's Entertainment Inc. (Cassino):** Dobrou lucros usando informações de cartões de “jogadores frequentes” , identificando grupo de jogadores que gastavam pouco mas geravam muito lucro, criou promoções diferenciadas;
- **Japan Credit Bureau (Empresa de cartão de crédito):** Melhoria da resposta a campanhas de marketing, retenção de clientes, identificação de novos segmentos de mercado.

Ferramentas

- Existem diversas ferramentas (gratuitas e pagas) de software no mercado para desenvolvimento de modelos de mineração de dados;
 - KEEL;
 - KNIME;
 - Neural Designer;
 - OpenNN;
 - Orange;
 - SAS;
 - R;
 - SPSS;
 - Weka.
- Uma das ferramentas mais utilizada hoje é o R.



Referências

- JONAS, Jeff; HARPER, Jim. **EFFECTIVE COUNTERTERRORISM AND THE LIMITED ROLE OF PREDICTIVE DATA MINING.** Disponível em: <<https://object.cato.org/sites/cato.org/files/pubs/pdf/pa584.pdf>>. Acesso em: 25 de junho de 2017.
- THOMÉ, Antonio Carlos Gay. **REDES NEURAIS - UMA FERRAMENTA PARA KDD E DATA MINING.** Disponível em: <http://equipe.nce.ufrj.br/thome/grad/nn/mat_didatico/apostila_kdd_mbi.pdf>. Acesso em: 27 de junho de 2017.
- MARTINS, Francilene Luzia; SILVA, Laise do Nascimento; FONTES, Viviane Gonçalves. **REDES NEURAIS APLICADO A DATA MINING.** Disponível em: <<http://eitecpicos.com/novo/files/EITEC-III/Redes%20Neurais%20Aplicado%20a%20Data%20Mining.pdf>>. Acesso em: 27 de junho de 2017.
- JÚNIOR, Fernando Silvera Goulart; FIDALGO, Robson do Nascimento. **DATA MINING.** Disponível em: <https://pt.slideshare.net/JeffersonRicardo/data-mining-10221271?qid=8db8921b-a7da-4d99-b4b0-797aca3ff208&v=&b=&from_search=7>. Acesso em: 28 de junho de 2017.