

Protótipo de um Sistema de Apoio à Decisão utilizando técnicas de Data Mining voltadas para a tarefa de Previsão.

Aluno: Ricardo Kremer

Orientador: Maurício Capobianco Lopes

Roteiro

- Introdução;
- Sistemas de Informação;
- Data Warehouse;
- Data Mining;
- Redes Neurais;
- Desenvolvimento do SAD;
- Conclusões;
- Sugestões.

Introdução

- **Motivação**

- informações são cada vez mais acumuladas em Data Warehouses;
- necessidades dos executivos de tomarem decisões mais acertadas.

- **Objetivo**

Auxiliar o processo de tomada de decisões de uma empresa, através de um Sistema de Apoio à Decisão utilizando técnicas de Data Mining, mais especificamente para efetuar previsões genéricas.

Sistemas de Informação

- **Conceito**

Sistema de Informação é uma combinação das formas de trabalho, informações, pessoas, e tecnologias de informação organizadas para alcançar metas em uma Organização [ALT92].

- **Formas de utilização**

- análise e investigação de tendências;
- mensuração e rastreamento de indicadores de fatores críticos;
- análise prospectiva;
- monitoramento de problemas;
- análise da concorrência.

Data Warehouse

- **Conceito**

O Data Warehouse é um "depósito" de dados orientado a assunto, alimentado pelos vários sistemas transacionais da empresa, no qual se podem buscar informações para o conhecimento do negócio. Nele os dados estão agrupados e, dessa forma, é fácil a visualização dos mesmos [DAL99].

- **Componentes**

- aquisição dos dados;
- armazenamento dos dados;
- acesso aos dados.

- **Vantagens**

- dados resumidos e não-voláteis.

Data Mining

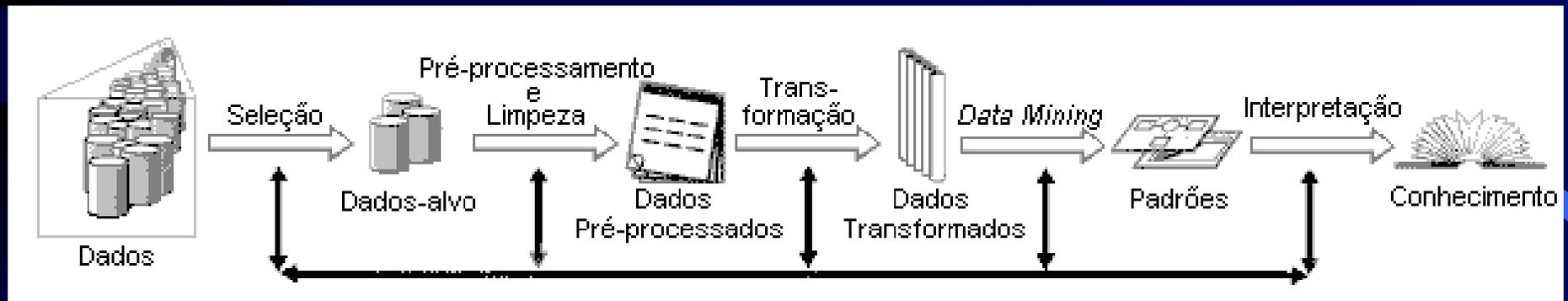
- **Conceito**

É a exploração e análise, por meios automáticos ou semi-automáticos, de uma grande quantidade de dados para descobrir padrões e regras significativas [BER97].

Data Mining (continuação)

- **KDD (Prospecção de conhecimento em bases de dados)**

Processo que envolve a automação da identificação e do reconhecimento de padrões em um banco de dados.

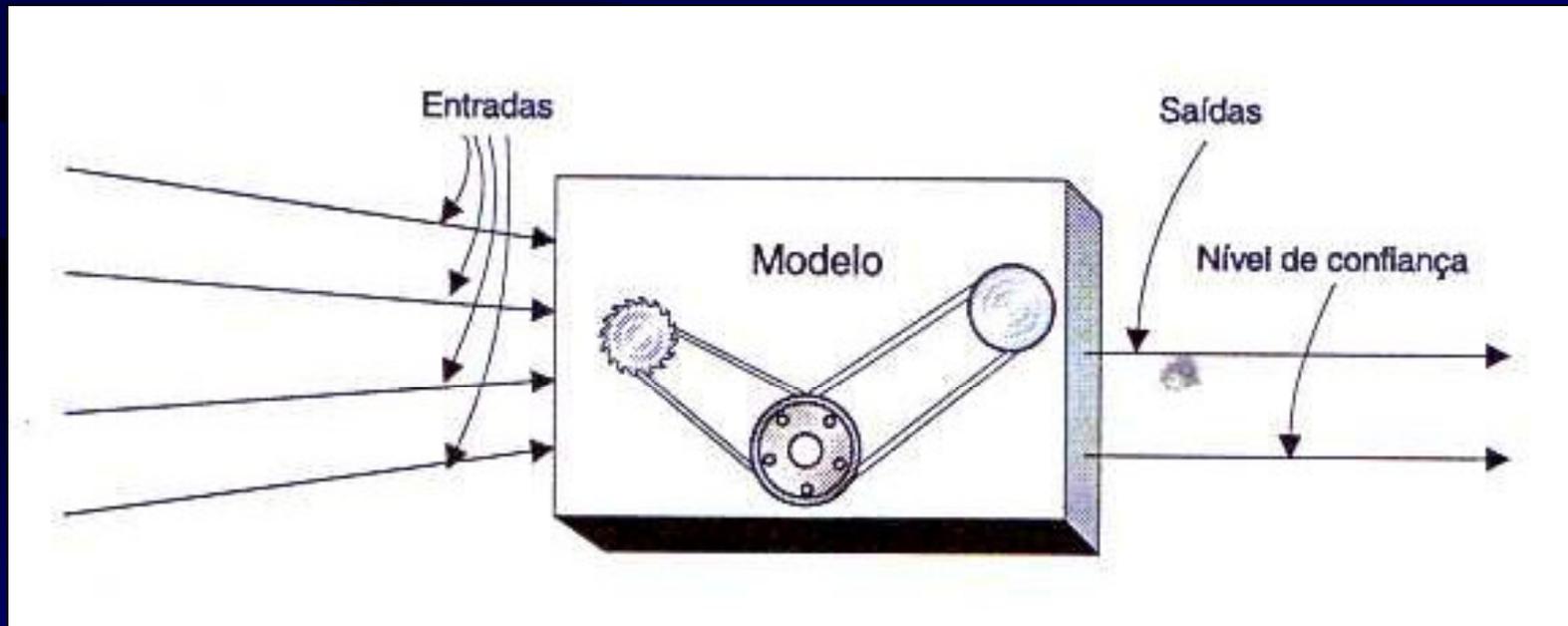


Passos do processo de KDD

Data Mining (continuação)

- Modelos

Produz um ou mais valores a partir de um dado conjunto de entradas.



Data Mining (continuação)

- **Utilidades**

- classificação;
- estimativa;
- agrupamento por afinidade;
- previsão;
- segmentação.

Redes Neurais

- **Conceito**

Redes neurais artificiais são sistemas computacionais de implementação em hardware ou software, que imitam as habilidades computacionais do sistema nervoso biológico, usando um grande número de simples neurônios artificiais interconectados [LOE96].

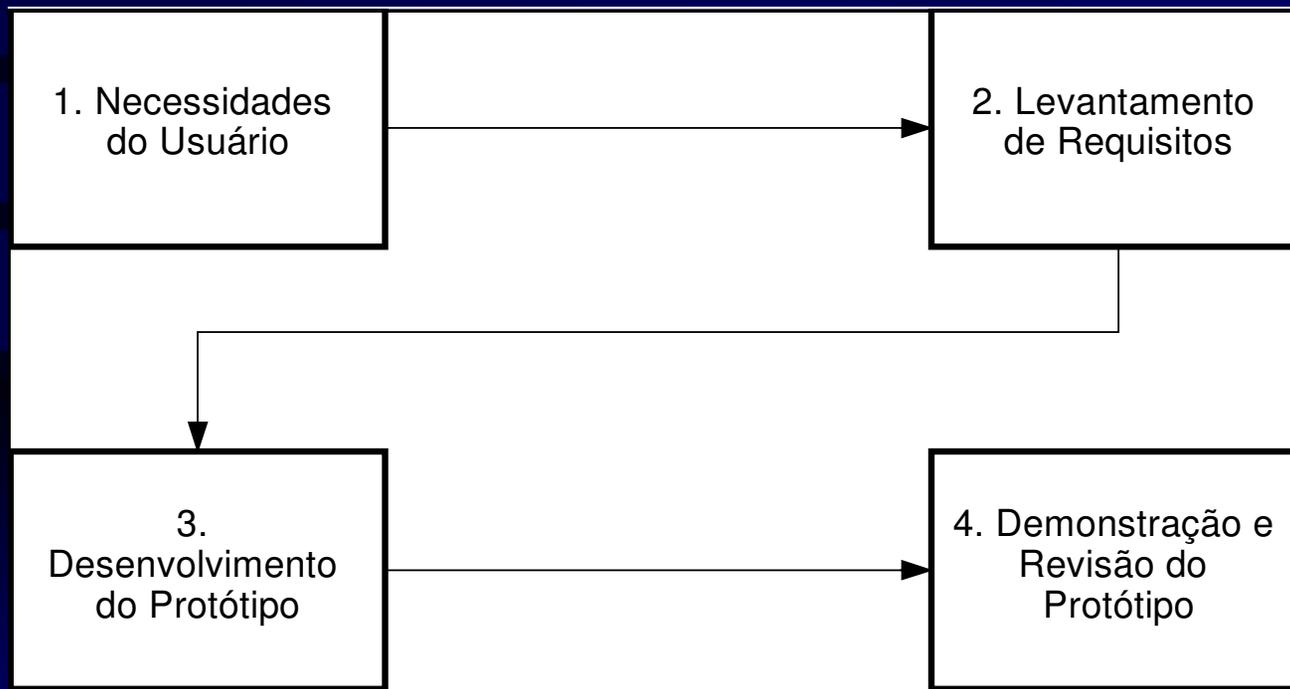
- **Treinamento**

- supervisionado;
- não-supervisionado.

- **Revocação**

Desenvolvimento do SAD

Metodologia de Prototipação Fundamental [MEL90]



Adaptação das Etapas da Metodologia

Desenvolvimento do SAD (continuação)

Necessidades do Usuário

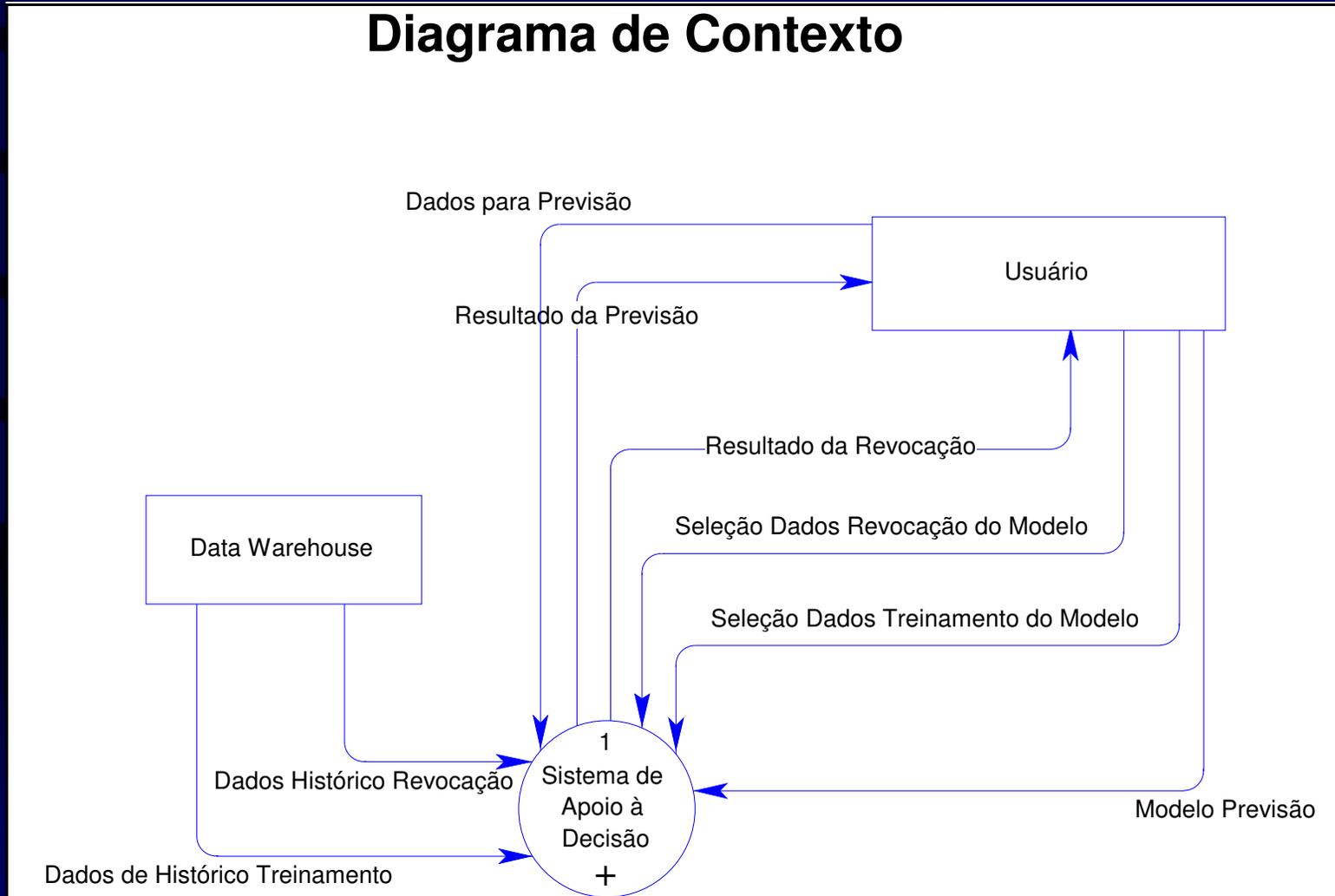
- Executivos necessitam medir os resultados das empresas e tomar decisões mais rapidamente;
- Acumulo de informações nas organizações;
- Um SAD utilizando Data Mining que fosse parametrizável e que possibilitasse prever ações futuras baseado em registros históricos seria de muita utilidade.

Levantamento de Requisitos

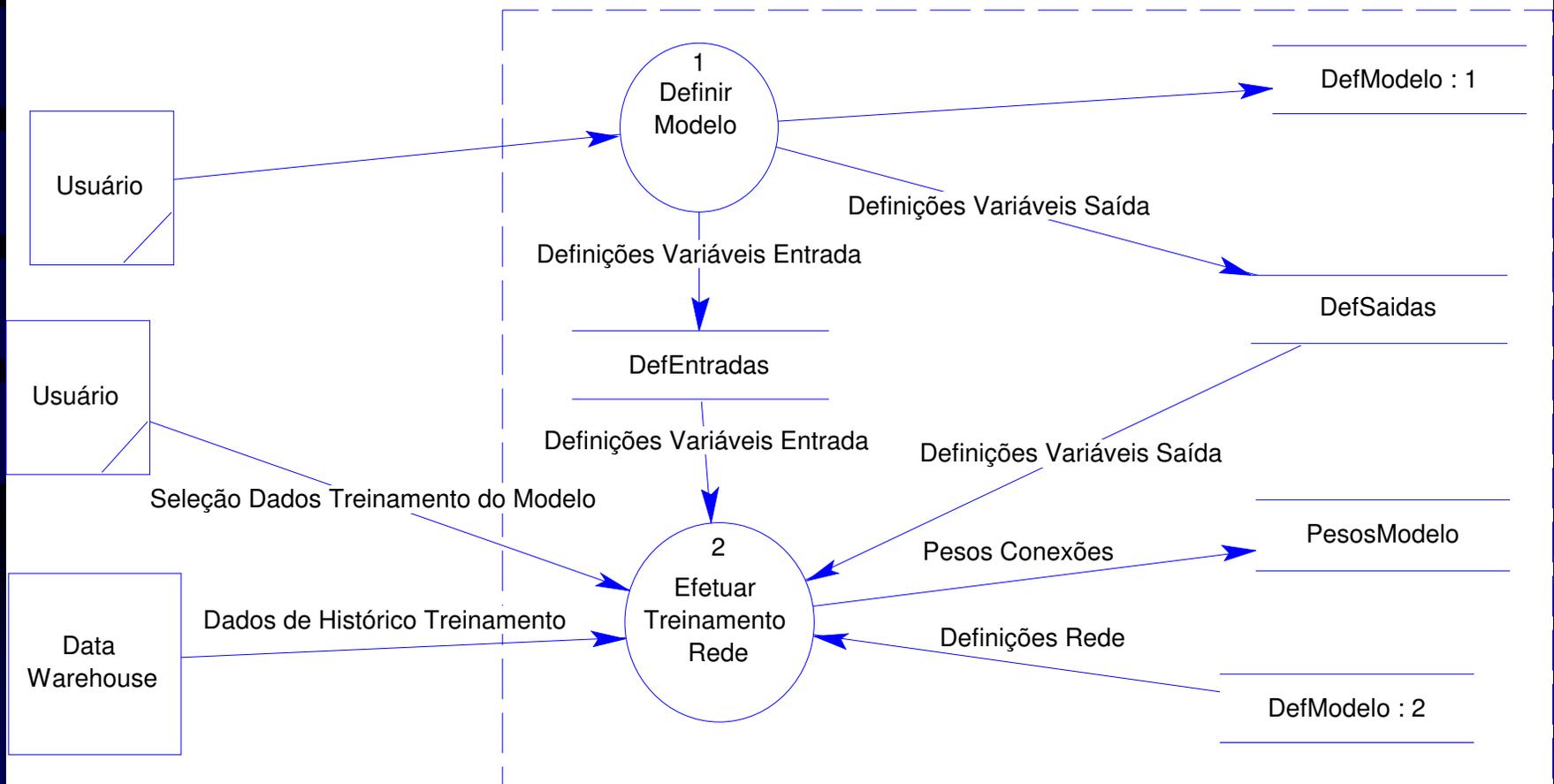
- Rede Neural modelo Feedforward / Backpropagation;
- Linguagem que possibilite desenvolver um Sistema amigável;
- Banco de Dados.

Desenvolvimento do SAD (continuação)

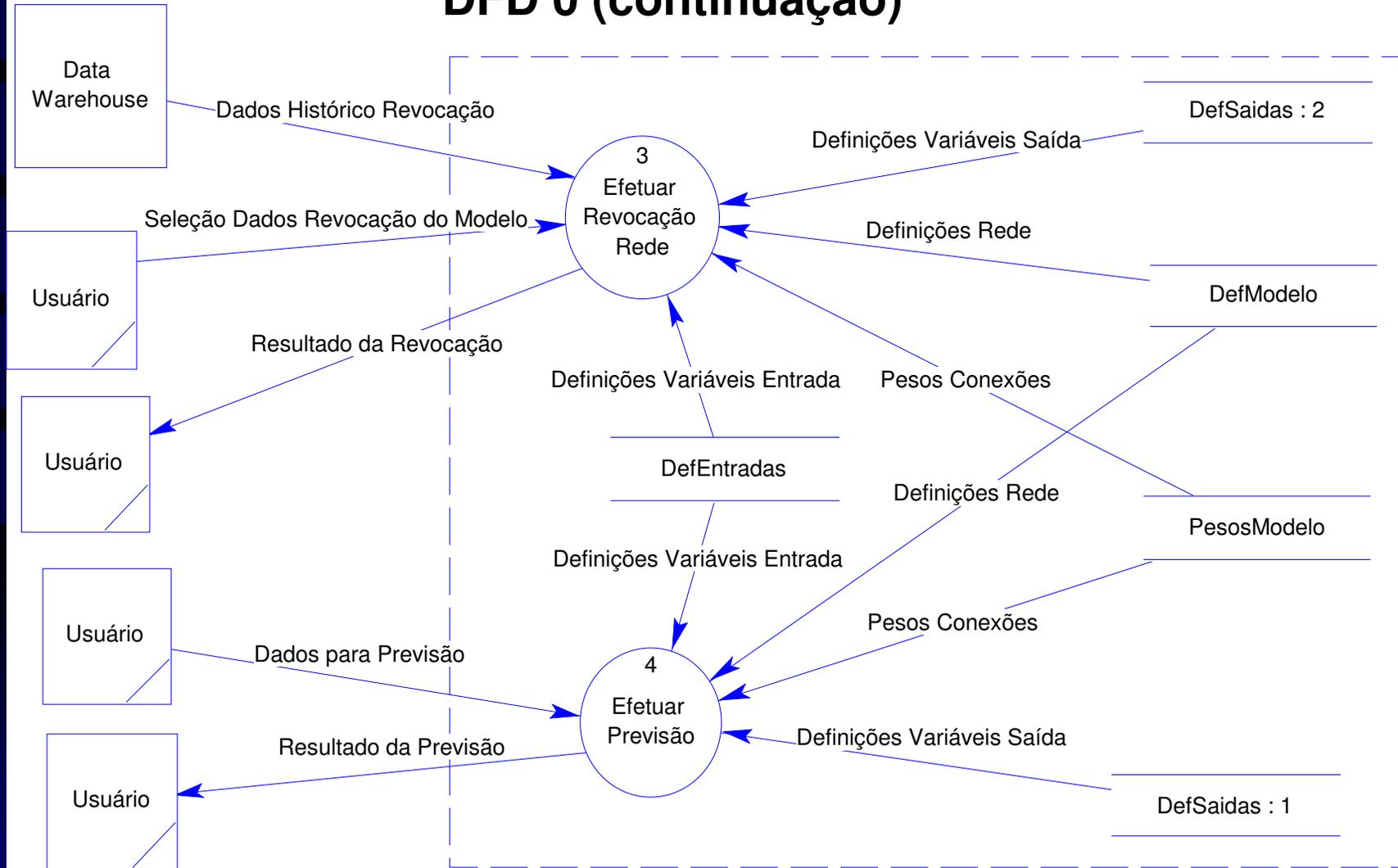
Diagrama de Contexto



DFD 0

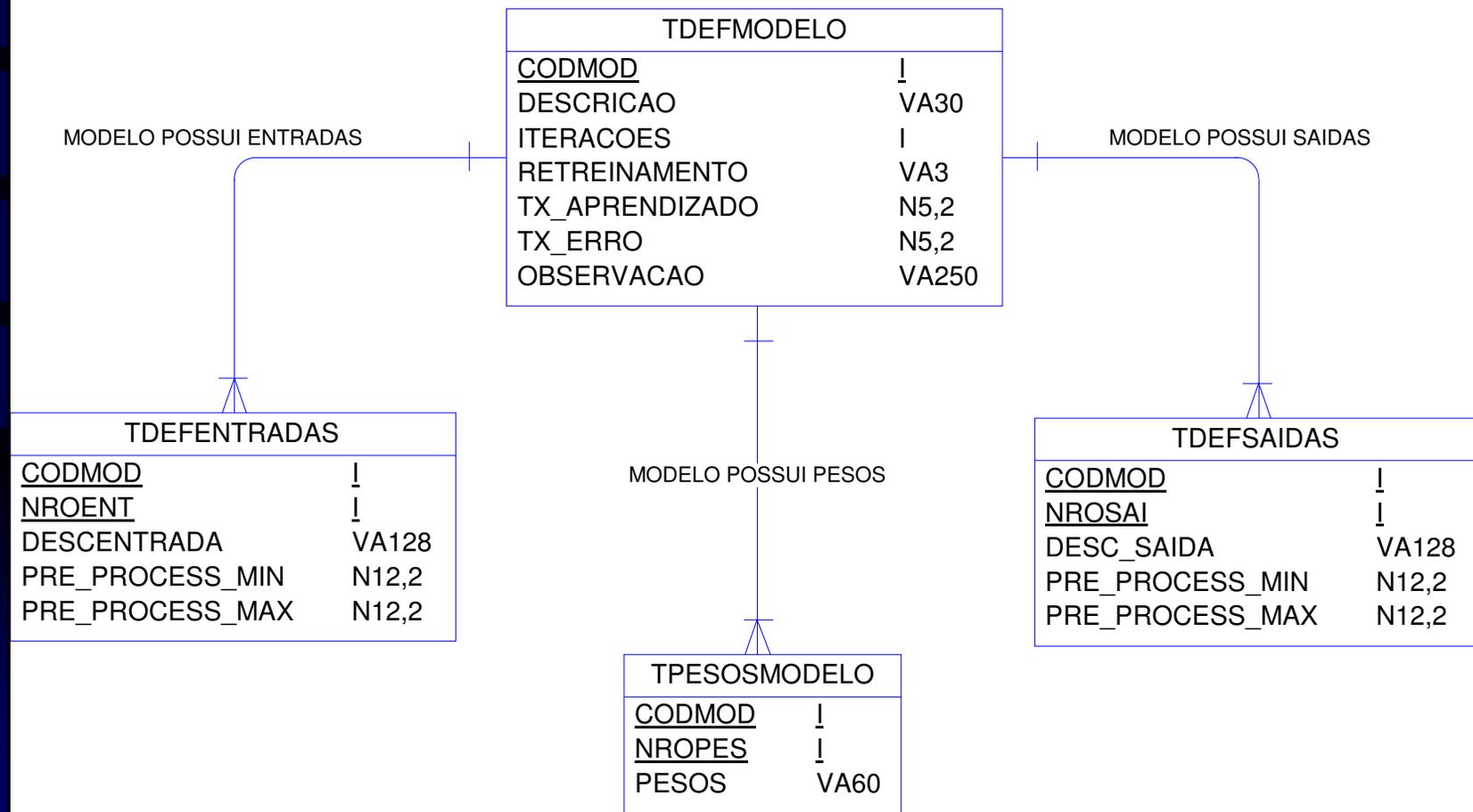


DFD 0 (continuação)



Desenvolvimento do SAD (continuação)

MER - Sistema de Apoio Decisão

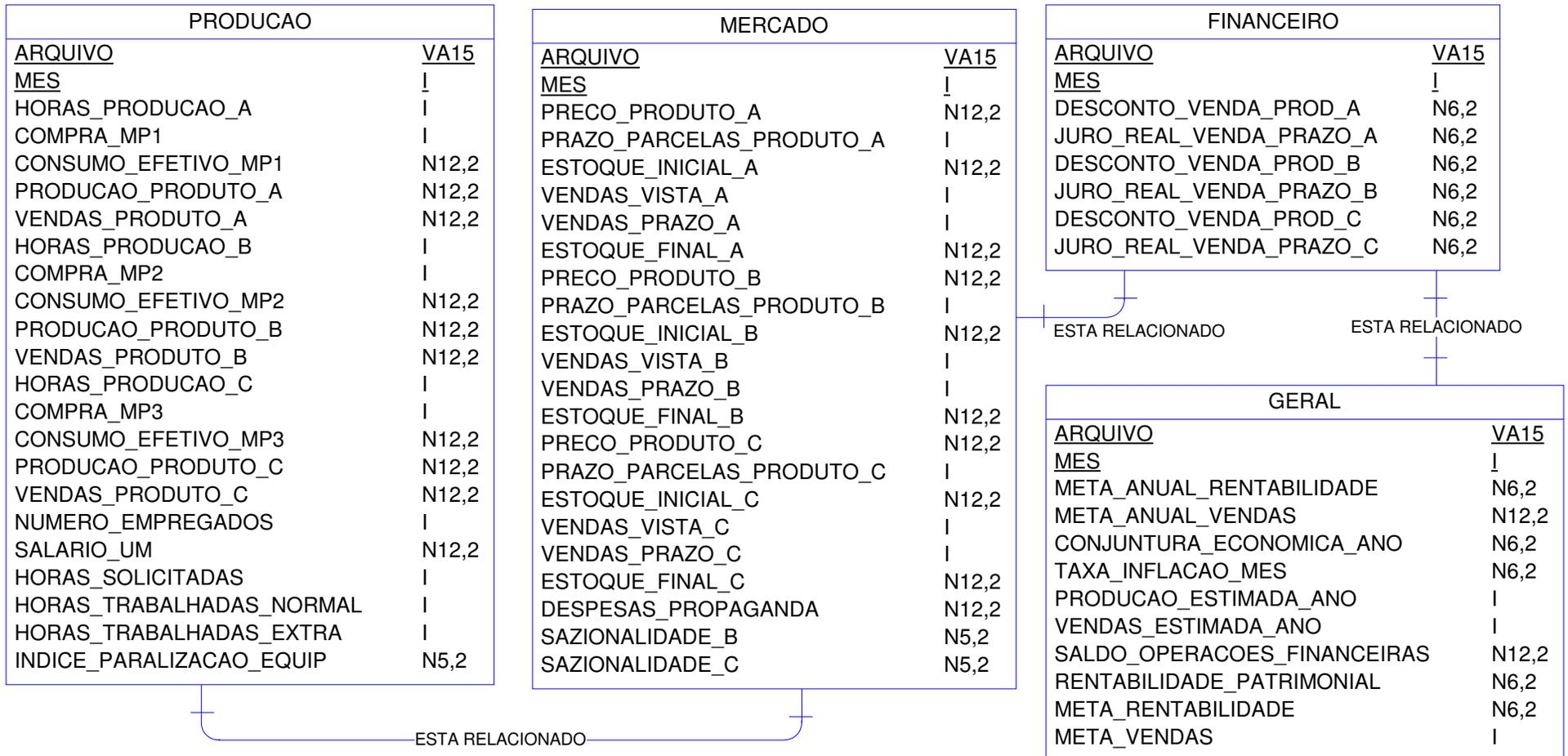


Desenvolvimento do SAD (continuação)

Desenvolvimento do Protótipo

- Componente de Rede Neural [VAL97]
- Aquisição dos Dados
 - JE Virtual;
 - Cenários e Variáveis e sua forma de armazenamento;
 - Rotina de Integração do JE Virtual para o Data Warehouse.
- Armazenamento dos Dados
 - Sybase SQL Anywhere;
 - Modelagem feita a partir dos dados Catalogados no JE Virtual.

MER - Data Warehouse



Desenvolvimento do SAD (continuação)

- **Desenvolvimento do Protótipo**

- Acesso aos Dados

- Utilização do SAD especificado seguindo os processos de KDD:

- Domínio da Aplicação;
 - Seleção dos Dados;
 - Pré-processamento e limpeza;
 - Data Mining;
 - Interpretação do Conhecimento.

Conclusões

- O Data Mining devolve informações que são induzidas dos dados;
- O Data Mining juntamente com as etapas de KDD se mostrou bastante eficiente para o desenvolvimento do Sistema;
- Testes com o JEVirtual mostraram a eficiência para a construção de modelos;
- Desvantagens do uso de Redes Neurais;
- Os objetivos do trabalho foram atingidos.

Conclusões (continuação)

- Limitações
 - Regras de pré-processamento limitadas;
 - Fonte de dados externa fixa.
- Dificuldades
 - Bibliografia;
 - Componente.

Sugestões

- Aplicação do Data Mining em outras tarefas/técnicas, como Classificação com o uso de Estatística padrão;
- Possibilidade de se escolher entre mais fontes de dados;
- Acesso aos dados através de outros Bancos de Dados.