



T0828

ORGANIZAÇÃO, CLASSIFICAÇÃO E PREPARAÇÃO DOS DADOS DE MATRIZES DE SUÍNOS PARA PROCESSO DE MINERAÇÃO DE DADOS

Rafael Viola (Bolsista PIBIC/CNPq) e Prof. Dr. Luiz Henrique Antunes Rodrigues (Orientador), Faculdade de Engenharia Agrícola - FEAGRI, UNICAMP

Empresas detentoras de tecnologias de melhoramento genético produtoras de matrizes de suínos possuem bancos de dados relacionais e estruturados com dados diariamente atualizados, passível de obtenção de conhecimento implícito nessas bases de dados. Essa descoberta do conhecimento é possível através da aplicação de técnicas de KDD (Knowledge Discovery in Databases) em bases de dados. Para isso, é necessária uma fase de pré-processamento, que abrange a seleção, pré-processamento e transformação de dados para aplicação de técnica de Mineração de Dados. O objetivo desse projeto foi realizar o pré-processamento dos dados da base de dados de matrizes de suínos através da determinação dos objetivos do estudo, da preparação dos dados através da análise de todas as variáveis envolvidas e seu relacionamento com as demais variáveis e da identificação das fontes de ruídos e respectivo tratamento. No processo de pré-processamento dos dados foram identificadas diversas fontes de ruídos, sendo que, para cada situação foi realizado um estudo juntamente com profissionais da área para tratamento desses dados. Como resultado desse projeto obteve-se uma base de dados transformada e pronta para a realização da técnica de Mineração de Dados.

Preparação de dados - Pré-processamento - Mineração de dados