

Aplicação de ID3 ou C4.5 ao Diagnóstico de Parkinson

Objetivo

A partir de um conjunto de exemplos, é possível aprender (ou derivar) as regras de classificação para o domínio em análise. O objetivo deste projeto é a determinação da árvore de decisão que traduz essas regras no diagnóstico de Parkinson.

Descrição

Um conjunto de exemplos (amostras) são usados na aprendizagem das regras de classificação para os elementos (população) do domínio em análise. Os elementos do domínio estão definidos por um conjunto de atributos (ou variáveis), que irão constituir os parâmetros das regras de classificação a derivar após o processo de aprendizagem. Um primeiro passo, muito importante, é a identificação dos atributos que melhor definem o domínio em análise.

O programa deve aplicar o algoritmo ID3/C4.5 para, com base no conjunto de dados disponibilizado [aqui](#), determinar uma Árvore de Decisão que traduz as regras de classificação desse conjunto de dados. O conjunto de dados deve ser cuidadosamente analisado de forma a verificar a eventual necessidade de pré-processamento. O modelo obtido deve poder depois ser utilizado na classificação de novos casos.

O projeto inclui os seguintes procedimentos:

- Implementação/aplicação do algoritmo ID3/C4.5.
- Apresentação das regras de classificação, que são retiradas da Árvore de Decisão resultante.
- Medição detalhada de resultados nos dados de treino e de teste.

Metodologias

Algoritmos de Aprendizagem ID3 e C4.5.

Ferramentas

Linguagem C++, Java

Bibliografia

Apontamentos das Aulas; "Artificial Intelligence: A Modern Approach"