

# Análise de Técnicas de Classificação Hierárquica usando Seleção de Atributos para Previsão da Função de Proteínas

Mestrado

Leonardo Henrique Pereira, Júlio Cesar Nievola

**Contexto:** Com o rápido avanço de pesquisas nas áreas de genoma e proteoma, o crescimento de bases com conhecimento biológico foi inevitável, tornando a análise destes dados uma tarefa muito complexa quando feita manualmente, foi necessário, então, a intervenção de campos da informática para o maior proveito da manipulação desses dados. A aplicação de técnicas de informática para a análise da informação biológica se denomina Bioinformática. **Problema:** Nesse sentido, um dos problemas desta área é a classificação de proteínas, que não é tão comum pelo fato da análise ser muito trabalhosa e complexa de se tratar, principalmente quando existe classes com hierarquia, ou seja, suas classes organizadas em superclasses e subclasses, formando estruturas de arvores ou grafos acíclicos direcionados. **Objetivo Geral:** A proposta aqui apresentada é a análise de técnicas de classificação hierárquica usando seleção de atributos para prever funções proteicas. **Método de Pesquisa:** Este trabalho explora a aplicação das abordagens de classificação hierárquica por nó, por nó pai e por camadas em conjunto à redução de dimensionalidade de atributos das relações utilizadas. As etapas do método de pesquisa seguirão as seguintes: Inicialmente o pré-processamento dos dados, seguindo com a aplicação das abordagens de classificação hierárquicas, então será feita a redução de dimensionalidade, continuando com a classificação classes geradas em cada abordagem de classificação hierárquica, finalizando com a análise dos resultados.

**Palavras-Chave:** Classificação Hierárquica de Proteínas, Predição de Função de Proteína, Seleção de Atributos, Bioinformática.