

SELEÇÃO DE ATRIBUTOS PARA CLASSIFICAÇÃO HIERÁRQUICA

Doutorado

SANTOS, Maria Angela Moscalewski Roveredo dos, NIEVOLA, Julio Cesar

Contexto: Segundo Santoro, os métodos indutivos do aprendizado de máquina aprendem a expressão do conceito a partir de um conjunto de treinamento. Conjuntos de treinamento, normalmente, são compostos por instâncias descritas por pares atributo-valor e uma classe associada. O conjunto de atributos usado para descrever as instâncias de treinamento tem um forte impacto na expressão induzida do conceito. As técnicas para a seleção de subconjuntos de atributos no contexto de aprendizado de máquina objetivam identificar os atributos que efetivamente contribuem para a caracterização da classe de uma instância. A seleção de atributos é uma das tarefas que podem ser realizadas durante a fase de pré-processamento de dados que serão posteriormente minerados. Sua importância caracteriza-se no fato de, na maioria dos casos, os dados serem coletados para propósitos diferentes da classificação. A seleção de atributos pode ser vista como um processo de busca onde o algoritmo usado deve encontrar o menor subconjunto de atributos com a melhor acurácia de classificação. Um tipo especial de classificação, denominado classificação hierárquica, é associado a múltiplas classes descrevendo uma estrutura hierárquica com diversos níveis específicos e regras bem definidas que conduzem as relações entre as diferentes classes presentes na hierarquia. A escolha da hierarquia está diretamente relacionada ao problema no qual o modelo de classificação será aplicado. **Objetivo Geral:** Apresentar experimentos computacionais que obtiveram melhoria de sua previsão e precisão em conjuntos de dados hierárquicos em relação aos já presentes na literatura, indicando os classificadores adotados e o uso da seleção de atributos responsáveis pelo resultado. **Questão de Pesquisa:** Técnicas de seleção de atributos são exploradas em conjunto com classificadores hierárquicos de diferentes categorias objetivando melhorar suas respectivas performances. **Método de Pesquisa:** Caracterizar medidas associadas ao objeto que se deseja extrair, de forma que as medidas sejam similares para objetos similares e diferentes para objetos distintos usando-se a extração de atributos. Segundo Plastino, métodos baseados em diferentes abordagens, considerando relações hierárquicas em diferentes formas, devem ser mensurados na tentativa de conseguir um melhor desempenho preditivo.

Palavras-chave: Seleção de atributos. Mineração de dados. Classificação. Classificação Hierárquica.