

232

ESTUDO DE ALGORITMOS PARA APRENDIZAGEM DE REDES BAYESIANAS. *Fausto Néri da Silva Vanin, Cristiane Koehler* (Campus Universitário de Vacaria, Curso de Tecnologia em Processamento de Dados, Universidade de Caxias do Sul-UCS).

O processo de aprendizagem de Redes Bayesianas é composto pelas etapas de aprendizagem da topologia e dos parâmetros associados a esta topologia. Atualmente, uma das pesquisas mais importantes na área de Inteligência Artificial é o desenvolvimento de técnicas de inferência eficientes para o uso em sistemas especialistas com domínio complexo. No entanto, o uso de tais técnicas pressupõe a disponibilidade de um modelo de conhecimento válido. A necessidade de extrair conhecimento de bases de dados vêm aumentando exponencialmente a cada ano. Cada vez mais, os volumes de informações excedem a capacidade de análise por métodos tradicionais que não conseguem analisá-los sob o enfoque do conhecimento. Para atender a essa necessidade vêm sendo pesquisadas novas técnicas e ferramentas, para extrair conhecimento a partir de bases de dados. Nesta pesquisa foi feito um estudo sobre os algoritmos de aprendizagem de redes bayesianas, onde foi possível analisá-los sob o enfoque do desempenho sobre a análise de grandes bases de dados. Para isto, foram implementadas rotinas utilizando a linguagem de programação CBuilder, versão 3.0 da Borland. Os resultados permitiram concluir que um dos grandes problemas encontrados nestes algoritmos é que os mesmos levam em consideração todas as variáveis disponíveis nas bases de dados gerando uma imensa explosão combinatorial de grande complexidade, sendo muitas vezes classificada como um problema de complexidade NP Completos. Em contrapartida, para se ter um modelo conhecimento válido para apoio à tomada de decisão, o mesmo deve ser composto apenas pelas variáveis mais relevantes. Para resolver este problema foi proposto uma modificação no algoritmo CDL incluindo uma rotina que utiliza o Teste de Correlação de Pearson para realizar uma análise da correlação entre as variáveis da base de dados com o objetivo de detectar quais destas variáveis possuem uma maior correlação entre si. A partir desta análise de correlação pode-se gerar a Rede Bayesiana como ferramenta para apoio à tomada de decisão (BIC-Fapergs).