

24. Developing a knowledge base for PID-expert system

Crina Samarghitean¹, Jouni Väliäho¹, Mauno Vihinen^{1,2}

¹Institute of Medical Technology, University of Tampere, Finland, and

²Research Unit, Tampere University Hospital, Finland

PID-expert is a medical expert system, which tries to assist physicians in diagnosis of primary immunodeficiencies, especially antibody deficiencies.

The knowledge base includes data and facts from IDR (Immunodeficiency Resource), literature and medical experts. For developing the system we employ knowledge engineering methods (manual and automatic) and data mining techniques such as decision tree induction (DTI).

Expert systems (ES) are computer-based applications developed to contain and use knowledge from human experts, to solve problems in a narrow domain. They improve the abilities of experts and increase the consistency and quality of problem-solving activities. In medicine, expert systems have been used in a variety of fields (internal medicine, infectious disease, paediatrics, neurology, otology, oncology, geriatrics, dentistry) and they fulfil a large number of tasks starting from assisting in clinical diagnosis and ending with guidance for a treatment protocol.

PID-expert is a medical expert system, which tries to assist physicians in diagnosis of primary immunodeficiencies, especially antibody deficiencies. In this area will be the first ones, who apply methodologies from data mining and machine learning.

The knowledge base of the system includes data and facts from IDR (Immunodeficiency Resource), an online (<http://bioinf.uta.fi/idr/>) database we have been developed since 1998, literature and from medical experts.

We try to apply data mining techniques to discover diagnostic knowledge, automatically from the patient data. We employ several knowledge engineering methods (manual and automatic) to acquire the knowledge and special attention is paid to pre-processing of the data. Inference mechanism is based on decision tree algorithm, which is well suited for medical domain due to its symbolic knowledge representation and explanation of decisions it makes.