Bachelor or Master Thesis

Mining Association Rules for Sepsis Treatments from Electronic Health Records

Semantic Computing Group

Philipp Cimiano
Basil Ell
Christoph Düsing

cimiano@techfak.uni-bielefeld.de
bell@techfak.uni-bielefeld.de
cduesing@techfak.uni-bielefeld.de

Association rule mining is a rule-based machine learning approach that aims to discover latent relations of the form $X => Y$ in large datasets. Their intuitive meaning is that most samples from the dataset that contain X tend to contain Y. The primary goal when mining these rules is to identify those rules with highest confidence, support, lift, and conviction.

Mining association rules is particularly useful when working with medical data for two reasons: (1) it provides inherently understandable rules that improve medical knowledge discovery, and (2) they can assist physicians during decision-making, e.g., when choosing the antibiotic for a given patient.

In order to reliably utilize them in a high-stakes domain such as healthcare, rules with high confidence and support are required. This is due to the fact that incorrect association rules may worsen a patient's medical condition further.

The offered thesis aims to derive high-confidence and high-support association rules from a real-world dataset that contains medical records of sepsis patients. The goal is to identify rules for the prescription of antibiotics. Subsequently, these rules should be evaluated in collaboration with physicians from the University Hospital OWL with respect to their validity and correctness.

The implementation of these approaches does not require prior knowledge of association rule mining. However, programming skills as well as basic data science skills are highly recommended.

Related literature

1. Srikant, Ramakrishnan, Quoc Vu, and Rakesh Agrawal. "Mining association rules with item constraints." Kdd. Vol. 97. 1997.

2. Lee, D. G., Ryu, K. S., Bashir, M., Bae, J. W., & Ryu, K. H. (2013). Discovering medical knowledge using association rule mining in young adults with acute myocardial infarction. Journal of medical systems, 37, 1-10.

3. Towfek, S. K. (2023). A Semantic Approach for Extracting the Medical Association Rules. Journal of Artificial Intelligence and Metaheuristics, 5(1), 46-6.

The Semantic Computing Group researches and develops methods that enable machines to acquire relevant knowledge as well as linguistic capabilities. Using methods from *natural language understanding* and *machine learning*, we are aiming at machines that are capable of knowledge acquisition by reading unstructured textual data. In particular, the group focuses on methods for information extraction, semantic parsing, ontology learning, entity linking, and question answering.

More information is available at: http://sc.cit-ec.uni-bielefeld.de.