MASTER2 INTERNSHIP

LEARNING CONTEXTUAL PREFERENCES

Supervisors: Karima Sedki, Rosy Tsopra, Jean Baptiste Lamy

Context: Many users are frequently confronting to make decisions like buying a house, choosing a movie to watch or a recipe to cook, etc. To satisfy users needs and help them in their decisions, integrating users preferences is a necessary step [1]. Preferences can be acquired by elicitation from users through a sequence of queries/answers or by learning them directly from data. However, as preference elicitation is time consuming, it is more appealing to learn preferences from data which is easy to observe and collect. The idea of preference learning [2,3] is to learn and construct a preference model from observed preference information. Once the preference model is learned, it can be used for making decision for example or predicting user's preferences on new products. We distinguish two principal types of preference models (quantitative and qualitative).

Objective: The objective of this internships is to propose and implement a method for learning contextual preferences from a database containing a set of instances described each one by a list of features and one label. The method consists in extending association rules algorithm [4]. Experimentations will be performed on two types of data: The first database containing the antibiotics, their properties, and their rank of recommendations (R) as defined in Clinical Practice Guidelines (the database was developed by medical experts). The aim is to learn preferences on the antibiotics properties for building a model that represents the closest possible experts reasoning and strategies to provide recommendations. An example of contextual preference is: *in the context of short duration protocol property, narrow spectrum property is preferred to low risk of side effects property.* This contextual preference means that given two antibiotics having each one short duration protocol, the antibiotic having narrow spectrum is preferred to the one having low risk of side effects. The second database concerns MovieLens dataset [5] which is one of the more popular ones. It contains movie rating data for research purposes.

The principal tasks of the internship are:
- understand the association rules algorithm (a priori algorithm) [4]. This algorithm is largely used in machine learning for particularly generating association rules on the basis of frequent itemsets,
- extend and implement the apriori algorithm in order to learn contextual preferences in our context.
- evaluate the implemented algorithm on the antibiotic database and MovieLens dataset.

Candidate: We are looking for a motivated Engineering School or Master's degree candidate in Computer Science who is motivated by artifcial intelligence and machine learning fields. Good programming abilities will be required.

Contact:

- karima.sedki@univ-paris13.fr (Assistant professor in Computer Science, LIMICS, Université Paris 13).
- rosytsopra@gmail.com (AHU, Médecin (MD, PhD), Centre de Recherche des Cordeliers, Université Paris-Descartes, Université Sorbonne Paris Cité, Department of Medical Informatics, Hôpital Européen Georges-Pompidou, AP-HP, Paris).
- jean-baptiste.lamy@univ-paris13.fr (Assistant professor in Computer Science, LIMICS, Université Paris 13).

Bibliographic references

[1] S. Benferhat, K. Sedki: Two alternatives for handling preferences in qualitative choice logic. Fuzzy Sets and Systems 159(15): 1889-1912 (2008)
[2] J. Furnkranz, E. Hullermeier, Preference learning: An introduction, 2010.
[3] Rosy Tsopra, Jean-Baptiste Lamy, Karima Sedki: Using preference learning for detecting inconsistencies in clinical practice guidelines: Methods and application to antibiotherapy. Artificial Intelligence in Medicine 89: 24-33 (2018)
[4] R. Agrawal and R. Srikant. Fast algorithms for mining association rules in large databases. In J. B. Bocca, M. Jarke, and C. Zaniolo, editors, VLDB, pages 487-499. Morgan Kaufmann, 1994.
[5] https://movielens.org