A LINGUISTIC COLLABORATIVE RECOMMENDER SYSTEM FOR ACADEMIC ORIENTATION*

E. J. CASTELLANO, L. MARTÍNEZ †

Department of Computer Sciences, University of Jaén, Campus Las Lagunillas s/n
Jaén, 23071, Spain

Students must face up to decision making situations along their academic journey, in order to keep on the chase of some professional competences. Several factors must be taken into account to achieve successful decisions namely: capacities, skills, social attitudes, etc. Many countries have created a figure so-called ***advisor***, whose role is to guide them in decisions regarding their academic future. The aim of our research is to help advisors in their task of guiding students by means of a linguistic Decision Support System (DSS) that uses students' marks and provides linguistic assessments about the choices that students can select. With such purpose we have built OrieB, a linguistic Web-based DSS based on Collaborative Filtering methods.

1. Introduction

Students must face up to decision making situations since early ages, in order to keep on the chase of some professional competences. However, the suitability of people in jobs or studies is not only based on their preferences. Other factors are involved: capacities, skills, social attitudes, etc., that must be taken into account to successfully decide [1].

Many countries have created a figure so-called ***advisor***, whose role is to guide them in decisions regarding their academic future. Advisors consider different criteria and indicators in their task being the key one the student's marks. Marks mean much more than a simple assessment: they indicate not only knowledge, but also skills, preferences about fields, attitudes, etc.

Advisors generally should guide and support several hundreds of students and hence analyze big amount of information. The aim of this contribution is to support them in their task of guiding students by means of a Decision Support System (DSS) that uses students' marks and provides linguistic recommendations about what subjects or modalities are better for each student

* This contribution is partially financed by the research projects TIN 2006-02121, P08-TIC-3548 and Feder Fonds.

† ejcastellano@gmail.com, martin@ujaen.es

in order to achieve a successful performance. We found out that the analysis of huge amount of data to make recommendations such as advisors follows a similar scheme to personalized marketing in Internet accomplished by Recommender Systems (RS) [2, 3]. These systems offer recommendations to users according to their preference profiles, guiding them through search spaces to find out the most suitable items for their needs.

Due to the complementarities between the necessities of academic orientation and the facilities of RS, we have built OrieB, a Web-DSS based on a Collaborative Recommender System (CRS) that supports advisors in their student guidance task. RS usually compute numerical degrees to rank the best items to be recommended. But in academic orientation we consider more appropriate the use of linguistic values, based on the fuzzy linguistic approach, [4, 5] for supporting advisor's tasks in order to manage the vagueness and uncertainty inherent to the problem, rather than precise numerical values which are just misleading approximations.

This contribution is structured as follows: Section 2 reviews in short some necessary concepts to understand our proposal, Section 3 presents our proposal of a linguistic DSS for academic orientation and finally some conclusions are pointed out in Section 4.

2. Preliminaries

In order to understand our proposal in this section we review in short the framework in which our DSS is applied as well as concepts about recommender systems and linguistic information.

2.1. *Academic orientation*

Academic orientation is related to supporting students to make decisions about their curriculum in order to be successful in their aim of obtaining a degree.

Most educational systems allow students to choose among different *specialization* branches building a personalized curriculum so-called *Academic Profile*. Each branch consists of a set of subjects (core subjects and elective subjects), and can group subjects in ***profiles*** or ***modules*** which try to specialize students in an area by means of *module subjects*. The objective is that every student reaches an adequate level of specialization; such level is easier achieved if students have adequate skills or feel affinity to the area of specialization, i.e., the more accurate they choose the better the development of their potential.

On the other hand, all academic institutions and educational systems have in common that they *evaluate* their students by means of different tools (tests, essays, tasks, etc.). The final result of this process is a *mark* that reflects not only the students' knowledge but also their skills, preferences about the subjects, etc. By means of CRS we can use these marks to build recommendations.

2.2. *Collaborative Filtering in Recommender Systems*

CRS gather ratings for items in a given domain and group customers with similar needs, preferences, etc. [6]. In a CRS, customers share each other their judgments and opinions about items they have already experienced, such that, the system can support them in order to make right and better decisions about items not experienced. The CRS provide customized recommendations for interesting items by using CF algorithms which try to predict user's satisfaction regarding an unrated item based on users with similar profile to the target user.

We have used in our proposal due to its success, the k-NN scheme in CF [6, 7] that carries out the following process to compute recommendations:

1. Analyzing and selecting data sets in order to define further processes, performing a *user-item* matrix of ratings [6].
2. Grouping users by a k-NN algorithm according to their similarity by means of a measure of similarity. k-NN algorithm simply chooses then the *k* most similar users to the target user [8].
3. Predicting items not rated yet by the target user, in order to choose which one/s will be recommended, by aggregating ratings of the selected neighbors [8].

The use of the k-NN scheme as memory-based approaches may present a drawback so-called *scalability problem*: the more data the less computing performance of the system. To avoid this, model-based algorithms generate a model offline from the dataset for computing the predictions. There exist different approaches [8], and we have implemented in our DSS the latter approach.

2.3. *Fuzzy linguistic approach*

In academic orientation are involved subjective and vague factors that imply uncertainty. In order to manage such uncertainty we propose the use of linguistic information to express the recommendations obtained by the DSS.

The fuzzy linguistic approach [4, 5] represents qualitative aspects as linguistic values by means of linguistic variables. We have to choose the appropriate linguistic descriptors for the term set and their semantics. Because the linguistic assessments are just approximate ones given by the individuals, we can consider that linear trapezoidal membership functions are good enough to capture the vagueness of those linguistic assessments, since it may be impossible or unnecessary to obtain more accurate values [9].

The universe of the discourse over which the term set is defined can be arbitrary, usually linguistic term sets are defined in the interval between 0 and 1, but in our case due to our framework (student marks) will be the marks range, [0,10]. In Figure 1 is shown the semantics of the linguistic term set $S = \{VL, L, M, H, VH\}$, that will use our system to support advisors.

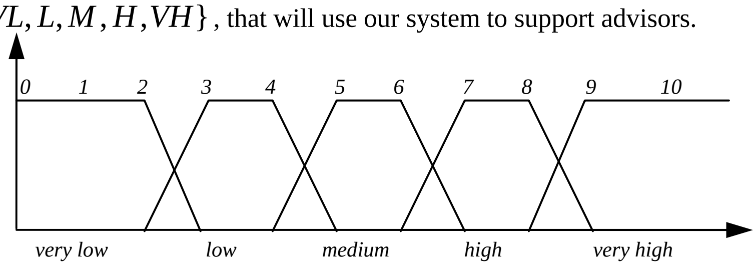

Figure 1. Semantic terms and membership functions.

3. OrieB, a linguistic Web-based DSS for academic orientation

OrieB is a DSS that deals with a dataset collected from several Spanish high schools in order to perform recommendations about which modules/subjects should students choose in Spanish Bachelor to achieve successful results in their academic journey. The proposal is to support advisors helping students in this task that becomes harder every day, due to the number of students they manage.

We have implemented OrieB by using our dataset of students, subjects and marks, and the results and algorithms obtained from detailed survey and optimization process are shown in [10]. The system will aid advisors to know (i) which module or modules will fit better a student, (ii) which subjects in each module and elective subjects will be better for her success and finally (iii) which core subjects can manifest extra difficulty to that student.

3.1. *Linguistic Recommendations*

As it was pointed out, it seems more natural the use of linguistic terms to make and explain recommendations than precise numerical values that can mislead the students in their decisions. So, OrieB will provide the recommendations using linguistic labels belonging to the term set showed in Figure 1.

OrieB uses a CF engine to compute a numerical prediction for subjects, *r*, and basing on it automatically assigns a linguistic label in *S*, according to equation (1). An example of this assignment can be viewed in Figure 2.

$$\tau : [0,10] \rightarrow S$$

$$s_i, i = \max_i (\gamma_i), \gamma_i = \mu_{s_i}(s_{ij}^N) = \begin{cases} 0 & if \ s_{ij}^N \notin Support \ (\mu_{s_i}(s_{ij}^N)) \\ \dfrac{s_{ij}^N - a_i}{b_i - c_i} & if \ a_i < s_{ij}^N < b_i \\ 1 & if \ c_i < s_{ij}^N < d_i \\ \dfrac{c_i - s_{ij}^N}{c_i - d_i} & if \ d_i < s_{ij}^N < c_i \end{cases} \quad (1)$$

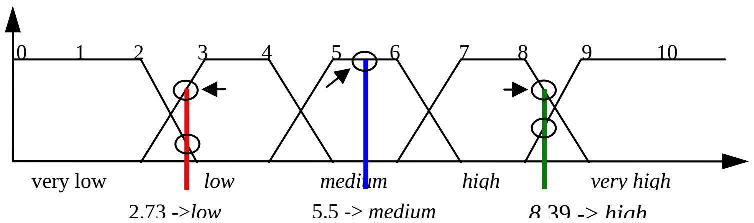

Figure 2: Example of linguistic labels assignment

3.2. *Supporting decisions*

OrieB offers three different types of support:

Vocational Program Recommendation

Trust	Interest	Program
57%	Very High	Arts
60%	High	Humanities and Social sciences
64.22%	Medium	Natural sciences and health
54.5%	Very Low	Technology

Figure 3: Module Recommendation

- **Module recommendation**: In order to aid advisors guiding students about the Module that better suits her according to her marks, OrieB computes a list of Modules ordered by relevance (Figure 3).
- **Subject recommendation**: Once students have chosen what module they prefer, they need to complete their curriculum with module and elective subjects. OrieB offers separate recommendations for each group (Figure 4).

Recommendation	Elective subject
Very high	Mass Media
High	Psichology
High	Computer science
Medium	French (2nd Language)

Figure 4. Subject recommendation in OrieB.

- **Warning difficulties in core subjects**: Students may need advises about what core subjects could be hard for them. In this sense, the system offers a list of core subjects with low predictions; it will warn the advisor which core subjects could cause difficulties to the student.

4. Conclusions

The use of collaborative filtering techniques applied to academic orientation can provide good results in order to support advisors in their task of guiding students in their academic decisions. Such techniques have been used to develop a linguistic DSS that can provide easy understanding recommendations for the students and which is being used to evaluate CF behavior in this domain.

Acknowledgements

This paper has been partially supported by the research projects TIN2006-2121, P08-TIC-3548 and Feder Fonds.

References

1. Salgado, J.F., *Personality and job competences: A comment on the Robertson & Kinder (1993) study.* Journal of Occupational and Organizational Psychology, 1996. **Vol 69**
2. Resnick, P. and H.R. Varian, *Recommender systems*. Vol. 40. 1997.
3. Adomavicius, G. and A. Tuzhilin, *Toward the next generation of recommender systems: A survey of the state-of-the-art and possible extensions.* IEEE Transactions on Knowledge and Data Engineering, 2005. **17**.
4. Martínez, L., L.G. Pérez, and M. Barranco, *A Multi-granular Linguistic Content-Based Recommendation Model.* International Journal of Intelligent Systems, 2007. **22**.
5. Zadeh, L.A., *The concept of a linguistic variable and its application to aproximate reasoning.* Information Science, 1975. **8 and 9**.
6. Herlocker, J.L., et al. *An algorithmic framework for performing collaborative filtering*. in *Sigir'99: Proceedings of 22nd International Conference on Research and Development in Information Retrieval*. 1999.
7. Sarwar, B., et al., *Analysis of recommendation algorithms for e-commerce*. Proceedings of the 2nd ACM conference on Electronic commerce. 2000.
8. Breese, J.S., D. Hecherman, and C. Kadie. *Empirical Analysis of Predictive Algorithms for Collaborative Filtering*. in *Uncertainty in Artificial Intelligence. Proceedings of the Fourteenth Conference.* 1998.
9. Delgado, M., et al., *Combining numerical and linguistic information in group decision making.* Journal of Information Sciences, 1998. **107**.
10. Castellano, E.J., et al. *Recomendación de Perfiles Académicos Mediante Algoritmos Colaborativos Basados en el Expediente*. in *Conferencia IADIS Ibero-Americana WWW/Internet 2007*. 2007.