

# Supporting Enrollment in Higher Education Through a Visual Recommendation System

Julià MINGUILLÓN<sup>a,1</sup>, Noe RIVAS<sup>a</sup> and Jonathan CHACÓN<sup>b</sup>

<sup>a</sup> *Universitat Oberta de Catalunya, Spain*

<sup>b</sup> *Elisava, Spain*

**Abstract.** Nowadays, most universities offer programmes and subjects online, specially in the case of fully online open/distance universities. Students have a higher degree of flexibility, which allows them to choose among an endless list of subjects for advancing within their degree. Although this can be seen as a positive result of enrollment flexibility policies, it may be also the source of one of the most well-known problems in open/distance education: high dropout rates, partly caused by inadequate enrollment. In this paper we propose a recommendation system that helps students to navigate through the list of available subjects using a visual metaphor, taking into account students' preferences and previous enrollment data. Our system is based on a two-dimensional map (2D) where subjects that can be taken together appear close to each other, as neighboring regions.

**Keywords.** recommendation system, visualization, enrollment, higher education

## 1. Introduction

Traditionally, distance and open universities provide students with greater flexibility during enrollment than brick-and-mortar universities. Students have almost no requirements about the number or kind of subjects they can take during one academic semester, although different recommendations are provided, in form of static information (from official degree documentation in institutional web pages) and dynamic information (from official mentors or other peers). Nevertheless, this flexibility can be misunderstood by some students, leading them to take wrong decisions regarding enrollment (too many subjects or inappropriate combinations of subjects). Previous work in this topic [1] showed that most students make decisions about enrollment taking into account their available time, semester organization and subject characteristics. In this paper we propose a recommendation system that helps students to navigate through the list of available subjects using a visual metaphor, namely a map. Subjects become regions in a 2D map, organized according to two complementary premises: subjects that are not supposed to be simultaneously taken should appear far from each other (i.e. they should not be neighboring regions), while subjects that are safe to be simultaneously taken may appear close to each other. In order to do so, we propose to combine different criteria related to user needs in a

---

<sup>1</sup>Corresponding Author: Universitat Oberta de Catalunya, Rambla Poblenou 156, 08018 Barcelona, Spain; E-mail: jminguillon@uoc.edu

subjects' distance matrix. According to their preferences, students can specify the partial weight of each distance criterion (based on historical data, including subject difficulty, satisfaction, semestral organization, etc. [1]) and visually determine the most appropriate subjects to enroll according to their previous achievements and goals.

## **2. Recommendation systems in educational scenarios**

In a recent review on academic advising systems [2] the authors found that recommending subjects is a primary research objective. Recommendation systems are classified in five different categories, according to their nature: Content-based, Collaborative filtering-based, Knowledge-based, Hybrid and Computational intelligence-based. Unlike other popular recommendation systems used in Amazon, Spotify or Netflix, providing students with appropriate subject recommendations needs to take into account not only students' preferences and background but also other constraints related to university policies and semestral organization, among others, that is, contextual information becomes important to provide users with good recommendations [3]. Actually, most recommendation systems can be considered hybrid in practice, as they combine aspects from the different categories. In the case of enrollment:

- Content-based: some subjects are part of the same learning path, develop the same competencies or can be prerequisites for other subjects.
- Collaborative-based: previous enrollment data and academic performance of students with similar enrollment patterns.
- Context-based: for instance, some subjects are offered only in one semester.

In [4] the authors showed that learning dashboards and the use of data visualization in educational scenarios is still limited. More recently, in [5] the authors describe a system that takes historical academic data and allows students to select subjects and predict their performance, using classical tables and line charts as a visualization. Instead, we would like to represent "time" (i.e. sequence of subjects) in a more organic way, making the student understand the concept only by taking a quick look. In order to do so, we propose a different approach, using a visual metaphor (namely, a map) that helps students to decide which subjects they want to enroll into. In summary, subjects that can be taken at the same time should appear close to each other, while subjects that should not be taken at the same time should appear far from each other. Unlike [5], we do not want to force absolute positions of subjects, but make them be part of a 2D map where each subject becomes a region, surrounded by neighboring subjects.

## **3. System design**

Students should be able to see a personalized 2D map according to their interests and previous enrollment data and academic results. As we pursue interpretability over accuracy [6], we focus on "typical" students and "reasonable" enrollment patterns only. In order to do so, we follow Shneiderman's mantra: Overview first → Zoom and filter → Details on demand, by means of an interactive learning dashboard [4].

### 3.1. Data

We used enrollment data from a Computer Engineering degree extracted from the institutional learning repository store [7]. Between the academic years 2010-2011 and 2019-2020, a total of 10,957 students enrolled 42,889 times, generating 10,409 different enrollment patterns, choosing among 695 different subjects (from all the programmes offered by the university). Actually, 7,418 enrollment patterns were taken only once, showing the underlying long tail distribution. As we want to build an usable recommendation system, we discarded data from those students enrolling into 7 or more subjects (less than 1.1% of students), and the long tail of subjects with less than 100 enrollments, so only the 54 most popular subjects were taken into account.

### 3.2. Distance criteria

The following criteria  $C$  (described in [1]) were used to compute the  $54 \times 54$  partial distance matrices  $D_C$  between each pair of subjects  $i$  and  $j$ :

1.  $D_S$ : Semestral organization ( $S_{i,j}$ ), subjects that are not supposed to be taken together according to the degree semestral organization should appear far from each other (i.e. separated).
2.  $D_P$ : Popularity ( $P_{i,j}$ ), subjects that are not usually taken together by students should appear separated. This factor is opposed to the previous one.
3.  $D_D$ : Difficulty ( $D_{i,j}$ ), subjects most likely to fail when taken together should appear separated.
4.  $D_R$ : Requisites ( $R_{i,j}$ ), subjects that need to be taken in a particular order (i.e. prerequisites) should appear separated.
5.  $D_O$ : Overlap ( $O_{i,j}$ ), subjects that have a high overlap between their assessment activity calendars should appear separated.

The resulting distance matrix  $D$  is as a linear combination of partial distance matrices  $D_C$ , where  $w_C \in [0, 1]$  is the weight assigned by the student to criterion  $C$  from the previous list and  $\varepsilon$  is a small random amount used to avoid zero distances that would generate subject overlaps:

$$D = \sum_C w_C D_C + \varepsilon$$

Then, using Sammon's non-metric multidimensional scaling algorithm [8] we obtain points in a 2D space, which are rotated to ensure that the degree capstone project is in the right part of the map. Finally, a Voronoi diagram is computed as the final 2D map.

### 3.3. Prototype

Currently now, an interactive proof-of-concept developed in R and Shiny is available<sup>2</sup> to demonstrate the possibilities of this approach, as shown in Figure 1. The map shows student's achievements (in shades of gray) and the rest of subjects are shown in blue, according to the weights given by the student to each criterion  $C$ . Subjects in blue closest to subjects in gray are candidates for recommendation and can be easily identified.

<sup>2</sup><http://personal.uoc.edu:8080/VE/>

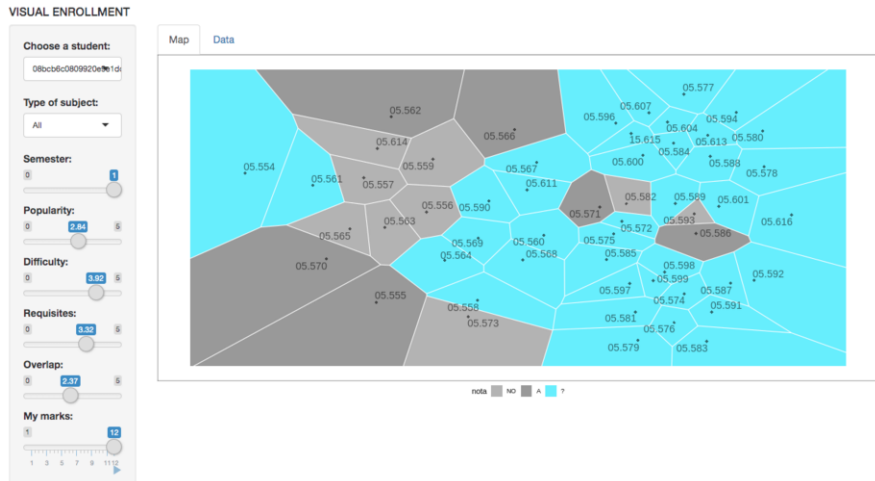


Figure 1. Current prototype showing data from a real student.

## 4. Conclusions

As an ongoing research project, several questions still need to be formally posed. For instance, how can we evaluate the proposed recommendation system? By means of student's satisfaction when using the tool? By comparing to student's real enrollment data after using our system? By measuring academic performance at the end of the semester? On the other hand, shall we use all available enrollment data (10 years, i.e. 20 semesters) or only data from the last  $N$  semesters? Is there an optimal  $w_C$  configuration?

## References

- [1] Rivas N, Minguillón J, Chacón J. ENROLLING HABITS IN HIGHER EDUCATION. WHAT SOURCES OF INFORMATION DO STUDENTS HAVE AND WHAT ARE MISSING? In: INTED2021 Proceedings. 15th Int. Technology, Education and Development Conf. IATED; 2021. p. 4980–4988.
- [2] Iatrellis O, Kameas A, Fitsilis P. Academic advising systems: A systematic literature review of empirical evidence. *Education Sciences*. 2017;7(4):90.
- [3] Kulkarni S, Rodd SF. Context Aware Recommendation Systems: A review of the state of the art techniques. *Computer Science Review*. 2020 Aug;37:100255.
- [4] Schwendimann BA, Rodríguez-Triana MJ, Vozniuk A, Prieto LP, Boroujeni MS, Holzer A, et al. Perceiving learning at a glance: A systematic literature review of learning dashboard research. *IEEE Transactions on Learning Technologies*. 2016;10(1):30–41.
- [5] Castells J, Mohammad PD, Galárraga L, Méndez G, Ortiz-Rojas M, Jiménez A. A Student-oriented Tool to Support Course Selection in Academic Counseling Sessions. In: Proceedings of the Workshop on Adoption, Adaptation and Pilots of Learning Analytics in Under-represented Regions (co-located with the 15th European Conference on Technology Enhanced Learning 2020); 2020. p. 48–57.
- [6] McNea SM, Riedl J, Konstan JA. Being accurate is not enough: how accuracy metrics have hurt recommender systems. In: CHI 06 Extended Abstracts on Human Factors in Computing Systems. CHI EA 06. Association for Computing Machinery; 2006. p. 10971101.
- [7] Minguillón J, Conesa J, Rodríguez ME, Santanach F. 8. In: Spector JM, Kumar V, Essa A, Huang YM, Koper R, Tortorella RAW, et al., editors. *Learning Analytics in Practice: Providing E-Learning Researchers and Practitioners with Activity Data*. Springer; 2018. p. 145167.
- [8] Sammon JW. A Nonlinear Mapping for Data Structure Analysis. *IEEE Transactions on Computers*. 1969 May;C18(5):401409.