

# Using Cortical Learning Algorithm to Arrange Sporadic Online Conversation Groups According to Personality Traits

Roberto Agustín García-Vélez<sup>1</sup>, Martín López-Nores<sup>2</sup>(✉),  
Yolanda Blanco-Fernández<sup>2</sup>, and José Juan Pazos-Arias<sup>2</sup>

<sup>1</sup> Área de Ciencias Exactas, Universidad Politécnica Salesiana,  
Calle Vieja 12-30 y Elia Liut, Cuenca, Ecuador  
rgarciav@ups.edu.ec

<sup>2</sup> Departamento de Ingeniería Telemática, AtlantTIC Research Center  
for Information and Communication Technologies, EE Telecomunicación,  
Universidade de Vigo, Campus Universitario s/n, 36310 Vigo, Spain  
{mlnores,yolanda,jose}@det.uvigo.es

**Abstract.** Online conversation spaces are becoming an increasingly popular tool for language learning. Various portals on the Internet offer technological platform where groups of students can meet native teachers and arrange conversation sessions in convenient dates and times. The students' satisfaction is very important in this context because potential new users pay much attention to the comments and ratings provided by others in the past. Therefore, the portals implement simple mechanisms for the students to rate their experiences, and use the feedback so gathered to promote the teachers who get the best evaluations. Notwithstanding, the current online conversation portals do not implement any means to proactively supervise the formation of the sporadic groups of students, aiming to ensure that the conversations will be balanced, engaging and pleasant to everyone. In this paper we look at the question of whether social data mining and machine learning technologies can be used to maximise the chances that the people put into the same group will get on well together. Specifically, we present one approach based on mining personality traits and using *Cortical Learning Algorithm* to aid in the planning of the sessions.

**Keywords:** Online language learning · Sporadic social networks · Personality traits · Cortical Learning Algorithm

## 1 Introduction

The concept of *sporadic* (aka *spontaneous*) *social networks* refers to groups of people (acquaintances or strangers) who happen to share a common space (physical or virtual) and have similar interests in it for a certain period of time [1–3]. Online conversation spaces are one context in which such sporadic groups form, specifically for the purposes of language learning. Various portals on the Internet

(e.g. Verbling and HowDoYouDo) offer a meeting point for groups of students to meet native teachers and arrange conversation sessions (henceforth, classes) in convenient dates and times. The teachers' personality and capabilities are key aspects to achieve positive group dynamics, and for that reason those portals ask the students to provide one rating for their labor after the classes have finished. The students often consider the aggregate value of those ratings when choosing the next classes in which they would like to participate.

Previous studies [4, 7, 10] have shown that the personality traits are one of the most important aspects to take into account in the formation of collaborative learning groups, because positive dynamics can potentiate the participation of all individuals. Likewise, other studies [8] have shown that the teacher ratings are an important feature for online learning portals, as they encourage the teachers to always do a good job and the students' satisfaction improves. However, despite the importance of ensuring positive dynamics and engaging topics in online conversation spaces, nowadays there are no solutions in place to proactively supervise the formation of the sporadic groups of students, and neither to assist the teachers in the preparation of appealing material.

In this paper we look at the question of whether social data mining and machine learning technologies can be used to maximise the chances that the people put into the same group will get on well together. The goal is challenging because the sporadic groups are most often independent of one another and many of the participants do not know each other beforehand. Therefore, unlike in [4, 7, 10] and similar works that focused on long-standing learning communities, it is not possible to make a first distribution of students into groups, and then refine progressively after successive rounds of observation and feedback. In order to face the more dynamic context on online learning spaces, we have developed an approach based on (i) mining personality traits of students and teacher from social networks, and (ii) using the advanced machine learning features of the *Cortical Learning Algorithm* (CLA) [9] to discover which combinations work well and which ones do not. This approach has been implemented in a real portal and some early results are now available.

The paper is organized as follows. Following this introduction, Sect. 2 describes the main modules and information flows of our approach. Then, Sect. 3 summarizes the findings obtained after the first few months of the approach running. Finally, Sect. 4 indicates the contributions we expect to make out of this project in relation to the research questions we are addressing, and also explains how we plan to continue with this work in the short and medium terms.

## 2 The Proposed Approach

We are developing the doctoral work in collaboration with an online conversation portal that specialises in English teaching for Spanish-speaking people. Rather than merely displaying the available classes and letting any students book their places in them, we want to advertise the new classes selectively and proactively, so that they are first known to the students who would most likely get

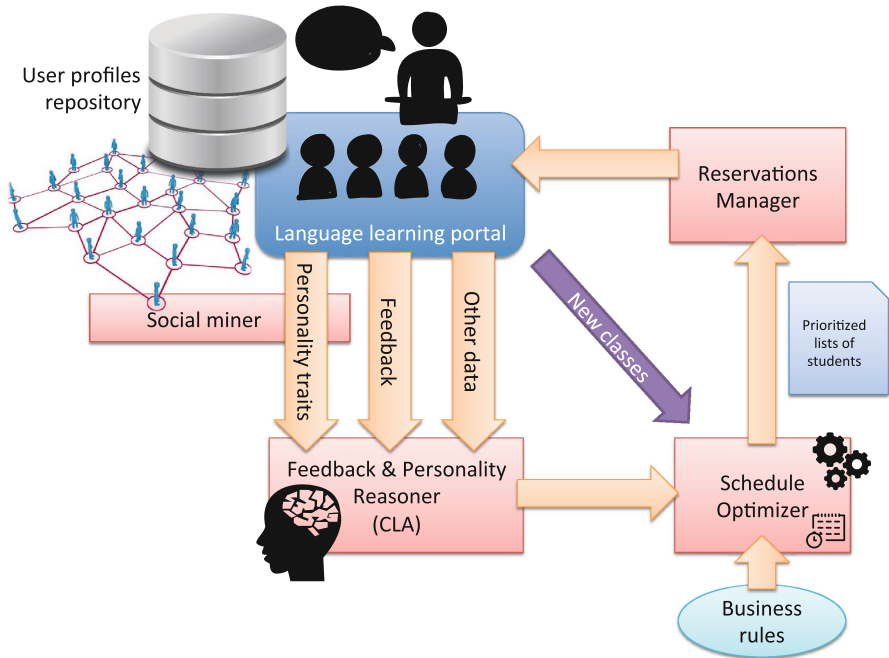
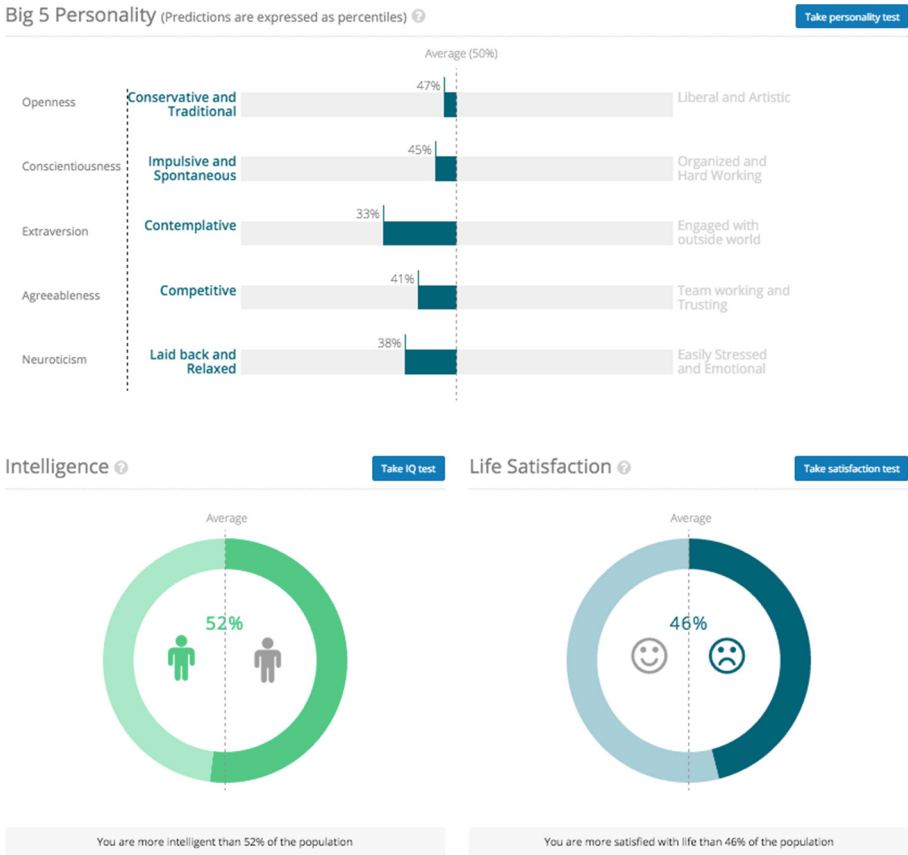


Fig. 1. Overall design of the proposed solution.

on well together. Of course, personality is assessed along with the data about the language level and the daily/hourly availability of teachers and students, bearing in mind a set of business rules about the number of classes each student is expected/allowed to attend during a period of time (depending on his/her type of subscription), the minimum average number of students in the classes to ensure that the portal makes profit, etc. The ultimate business goal is to achieve the greatest occupancy in the classes, so as to deliver service to more students with fewer classes.

The architectural scheme of our proposed solution is depicted in Fig. 1, comprising four main modules: the “*Social miner*”, the “*Feedback & personality reasoner*”, the “*Schedule optimizer*” and the “*Reservations manager*”. All the planning depends on a repository of student and teacher profiles, storing their personal data, learning needs and availability as usual, plus newly-added fields about personality traits. These are inferred from the users’ activities in online social networks by the “*Social miner*” module. Currently, this module works with Facebook only, and gets data directly from the *Apply Magic Sauce Prediction API* (<http://applymagicsauce.com/>), which computes psycho-demographic profiles that include estimations of the classical *Big 5* personality traits [6] (one of the most popular frameworks used by psychologists to describe the human psyche along the dimensions of openness, conscientiousness, extraversion, agreeableness and neuroticism [6]) as well as estimations of intelligence, life satisfaction,

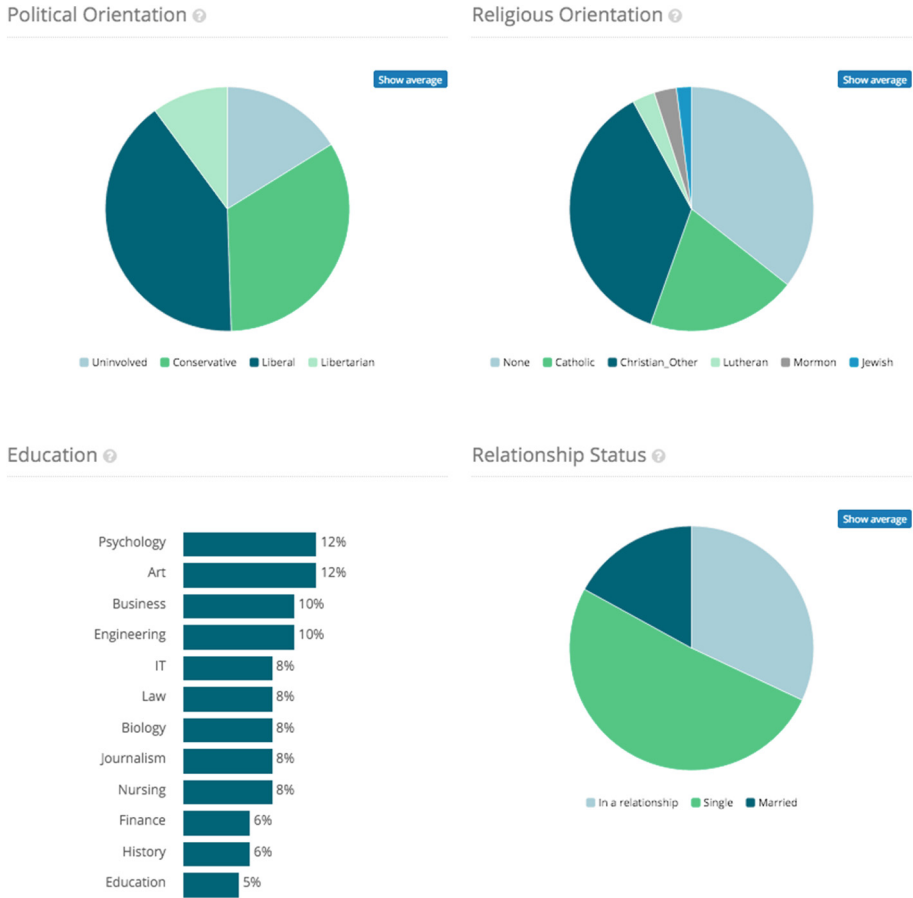


**Fig. 2.** Estimations of the *Big 5* personality traits, intelligence and life satisfaction.

sexual preference, political and religious orientations, education and relationship status. Figures 2 and 3 show graphical representations of the results obtained for one student.

The data obtained by the “*Social miner*” for all the profiles is fed into the “*Feedback & personality reasoner*” along with the feedback provided by students and teachers after the classes, as well as traces of the number and duration of everyone’s interventions in the class, their ages, genders, locations and timezones. As regards the feedback, currently, we ask students to rate their satisfaction with the class, the teacher’s labor and their interactions with the other students on a 5-point Likert scale (“*very negative*”, “*negative*”, “*neutral*”, “*positive*”, “*very positive*”). In turn, we ask the teacher to rate each student’s attitude and performance with regard to the class objectives.

The “*Feedback & personality reasoner*” is expected to learn progressively which combinations of traits and other data ensure positive feedback and which ones do not, what are the proper balances, etc. The *Cortical Learning Algorithm*



**Fig. 3.** Estimations of political and religious orientations, education, and relationship status.

models structural and algorithmic properties of the human brain’s neocortex to discover patterns in more sophisticated ways than classical artifacts like Bayesian networks and neural networks do [9]. On the one hand, we use CLA in *learning mode* to assimilate new bundles of information coming after every class, which serves to continuously evolve a complex model of the aspects that may influence the satisfaction of students and teachers. On the other hand, we use it in *inference mode* to make predictions and aid in the arrangement of forthcoming classes, in a loop with the “*Schedule optimizer*”.

The “*Schedule optimizer*” is based on the code of the open-source project OptaPlanner (<http://www.optaplanner.org/>), which provides a lightweight, embeddable constraint satisfaction engine that optimizes planning problems (typically, problems which are probably NP-complete or harder). Our module tries to harmonize the composition of the conversation groups with the

availability and interests of every teacher and student, as well as with the business rules of the online conversation portal. At the output, we get a prioritized list of potential students for each new class, which is used by the “*Reservations manager*” selectively deliver warnings about the available classes, to filter the lists of classes offered on the web site to each student, and to inform teachers about opportunities to offer new classes. The list of potential students is revised as the places in the classes are booked, until they are full.

### 3 Preliminary Evaluation

Our approach is being tested and some early results are now available, following some training of the “*Feedback & personality reasoner*” module. The data gathered thus far suggests that the average satisfaction does improve with regard to statistics from previous months, when there was no artificial intelligence aiding in the planning of the classes. However, ANOVA tests [5] indicate that the amount of data is not yet sufficient to fully confirm this hypothesis. Besides, we have found that the reported levels of students’ ratings of other participants improve as more of them come from the lists computed by the “*Schedule optimizer*”.

### 4 Conclusions and Future Work

The approach presented in this paper will serve to address the following research questions:

1. Is it possible to predict the levels of satisfaction of students and teachers in a class before it takes place, considering their personality traits and topics of interest?
2. Is it possible to get consistently more positive feedback by arranging the conversation groups according to personality traits, mined from the students’ and teachers’ activities in online social networks?
3. What is the quickest, least cumbersome way of gathering feedback after a class, aiming to get relevant information about each participant’s impressions of everyone else?
4. Does an increase in levels of satisfaction correlate with better learning outcomes, or do the students accept a trade-off between learning and socializing?

In replying to these questions, we expect to come up with a working solution to improve the experience in the collaborating portal, thereby gaining valuable insight (from the socioeconomic point of view) into the interrelations and trade-offs among the many factors involved: learning needs and availability, personality traits, social networking activities, business rules, learning outcomes, etc. Hopefully, some of the findings will be relevant not only to online conversation portals, but also to other areas of application of the concept of sporadic social networks.

As for the continuation of this work, we are seeking to complete the preliminary analysis of results with further evidence, gathered during at least one year

of our approach running. Besides, we want to evaluate alternative solutions for the social mining and machine learning tasks. First, we want to compare the learning capabilities of HTM against those of *Big Data* tools like Weka (<http://www.cs.waikato.ac.nz/ml/weka/>). Second, we are working to optimize the conversation groups not only on the grounds of personality traits, but also topics of interest mined from social networks and, thereby, make CLA reason about broad categories of topics (e.g. “Culture”, “Sports”, “Health”, etc.). Finally, in the medium term, we want to implement solutions of our own to get additional information about personality, by monitoring the students’ and teachers’ participation in the classes, using sound processing and face recognition tools to appraise shyness and mood, to recognise smiles and laughs, etc. This would be a significant improvement of the technological platform of an online conversation portal.

**Acknowledgment.** This work is being supported by the European Regional Development Fund (ERDF) and the Galician Regional Government under agreement for funding the Atlantic Research Center for Information and Communication Technologies (AtlantTIC), and by the Ministerio de Educación y Ciencia (Gobierno de España) research project TIN2013-42774-R (partly financed with FEDER funds).

## References

1. Ben Nejma, G., Roose, P., Dalmau, M., Gensel, J.: Service discovery for spontaneous communities in pervasive environments. In: Wang, J., et al. (eds.) WISE 2015. LNCS, vol. 9419, pp. 337–347. Springer, Heidelberg (2015). doi:[10.1007/978-3-319-26187-4\\_32](https://doi.org/10.1007/978-3-319-26187-4_32)
2. Boutet, A., Frenot, S., Laforest, F., Launay, P., Le Sommer, N., Maheo, Y., Reimert, D.: C3PO: a network and application framework for spontaneous and ephemeral social networks. In: Wang, J., et al. (eds.) WISE 2015. LNCS, vol. 9419, pp. 348–358. Springer, Heidelberg (2015). doi:[10.1007/978-3-319-26187-4\\_33](https://doi.org/10.1007/978-3-319-26187-4_33)
3. Bravo-Torres, J.F., López-Nores, M., Blanco-Fernández, Y., Pazos-Arias, J.J., Ordióñez-Morales, E.F.: Leveraging ad-hoc networking and mobile cloud computing to exploit short-lived relationships among users on the move. In: al Saidi, A., Fleischer, R., Maamar, Z., Rana, O.F. (eds.) ICC 2014. LNCS, vol. 8993, pp. 84–102. Springer, Heidelberg (2015)
4. Dascalu, M., Bodea, C.N., Lytras, M., Ordez de Pablos, P., Burlacu, A.: Improving e-learning communities through optimal composition of multidisciplinary learning groups. *Comput. Hum. Behav.* **30**, 362–371 (2014)
5. Gamst, G., Meyers, L., Guarino, A.: *Analysis of Variance Designs: A Conceptual and Computational Approach with SPSS and SAS*. Cambridge University Press, New York (2008)
6. McCrae, R., Costa, P.: The five-factor theory of personality. In: John, O.P., Robins, R.W., Pervin, L.A. (eds.) *Handbook of Personality: Theory and Research*, 3rd edn, pp. 159–181. Guilford Press, New York (2008)
7. Mudaliar, D., Modi, N.: Contemplating crossover operators of genetic algorithm for student group formation problem. *Int. J. Emerg. Technol. Adv. Eng.* **2**, 192–197 (2012)

8. Neubaum, G., Wichmann, A., Eimler, S., Kramer, N.: Investigating incentives for students to provide peer feedback in a semi-open online course: an experimental study. In: Proceedings of 10th International Symposium on Open Collaboration (OpenSym), Berlin, Germany, August 2014
9. Numenta, Inc.: HTM Cortical Learning Algorithms. White paper (2011). [http://numenta.org/resources/HTM\\_CorticalLearningAlgorithms.pdf](http://numenta.org/resources/HTM_CorticalLearningAlgorithms.pdf)
10. Sinha, T.: Together we stand, together we fall, together we win: dynamic team formation in massive open online courses. In: Proceedings of 5th International Conference on Applications of Digital Information and Web Technologies (ICADIWT), Bangalore, India, February 2014