1st International Conference on Higher Education Advances, HEAd'15
Universitat Politècnica de València, València, 2015
DOI: http://dx.doi.org/10.4995/HEAd15.2015.333

Data mining teaching throughout cards game competition

J. Antonanzas-Torres*, R. Urraca*, E. Sodupe-Ortega*, F.J. Martinez-de-Pison*, A. Pernía-Espinoza*

* EDMANS Research Group, Department of Mechanical Engineering, University of La Rioja, C/ San Jose de Calasanz 33, Logroño 26004, Spain.

Abstract: Data-mining techniques and statistical metrics learning can be complicated because of the complexity and overwhelming nature of this field. In this paper a class competition to improve learning of designing Decision Support Systems (DSS) by playing a classic cards game named "Copo" is proposed. The fact that this game is based on a probabilistic problem and that different solutions can be obtained represents a very typical kind of problem in the field of engineering and computer science. During the last four years in which this methodology has been applied to students of Computer Engineering Grade at University of La Rioja, both professors and students h satisfied with the results achieved. The programing of this project was performed in WEKA free software, a data mining suite widely spread in universities and companies. Therefore, not only students learn how to solve stochastic and probabilistic problems, but also the software tools and functions to develop DSS solutions.

Keywords: data mining; learning by playing; decision support systems; WEKA, soft-computing.

Introduction

Educational models that focus on student's autonomous workload and Project-based learning (PBL) (Buck Institute for Education, 2002; Boss & Krauss, 2007) are promoted by the European Space of Higher Education or EHEA (EEES in Spanish, 2015). These methods allow students to better internalise concepts and learn how to apply them to specific situations. Traditional teaching methods have shown the lack of promoting interactions among students, who in some cases becomes a passive character that has difficulties in assimilating the theoretical content of the subject and in thinking of possible real applications of what he is studying. For these reasons, it is necessary to introduce some changes in the traditional teaching methods, enhancing the student's personal work, their interpersonal relationships and communication skills (Reynolds & Miller, 2013). Some alternatives to traditional teaching have risen, such as PBL, Problem-based Learning (Merril, 2007) or Collaborative-based Learning (Slavin, 1980).

In this paper, a Competition-based Learning (CnBL) (Gredler, 2004; Burguillo, 2010; Andreu-Andrés & García-Casas, 2011) scheme was used to enhance student's motivation, which normally leads to a better understanding of the subject (Bergin & Reilly, 2005). CnBL promotes learning through a class competition, but without the scores affecting the student's learning results. It can be combined with other teaching techniques, listed above.

In this experience, a class competition, with students divided in teams, have to work and collaborate to win the quiz. Thus, students take active part in the course and get enthused to be in the top of the classification.

The methodology herein presented has been implemented during the last four academic courses in the subject 'Knowledge Engineering' of Computer Engineering Grade at University of La Rioja. Along the course, the methodology has lead substantial improvements in the teaching methods, as motivation for the subject among the students has increased.

The main objective of the methodology proposed is to make the students assimilate the concepts shown during the course and to boost their motivation in the learning of the

This work is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License (CC BY-NC-ND 4.0).

subject in which the methodology is applied. This is driven through putting student's acquired theoretical knowledge into practice by designing Decision Support Systems (DSS) for supporting decisions in real problems. Students, organized in groups, have to construct a useful DSS for a card game, developing a model as precise as possible based on a database of historical results of a game of cards and deciding how to use it to maximize the benefit. In addition to applying the knowledge to a real case and interpreting the results obtained, students also interact and discuss with each other to put ideas in common and reach the solution, which enables to develop communication skills.

Methodology

The methodology is based on the competition between groups of students, who must develop an expert system to play a Spanish cards game known as *'copo'*. Students have to apply the knowledge discovery databases (KDD) process in order to make explicit the knowledge on a large database of different *"copo"* plays. In general, a KDD process consists on extracting 'hidden' knowledge from large datasets, which is obtained after processing the data, applying data mining techniques and interpreting the results. Therefore, in this experience, after analysing the different plays (dataset), an expert system composed by a classifier with some additional rules was obtained.

The game is played with the Spanish pack of cards, which is comprised by 4 different suits (coins, cups, clubs and swords) and 10 cards per suit (Figure 1). For this game, the cards from the same suit are sorted as follows: 2, 4, 5, 6, 7, 10 (jack), 11 (knight), 12 (king), 3, 1 (ace); according to the traditional numbering of the Spanish deck. However, to keep it simple, hereafter, cards will be referred from 1 to 10 based on the previously described sorting position.

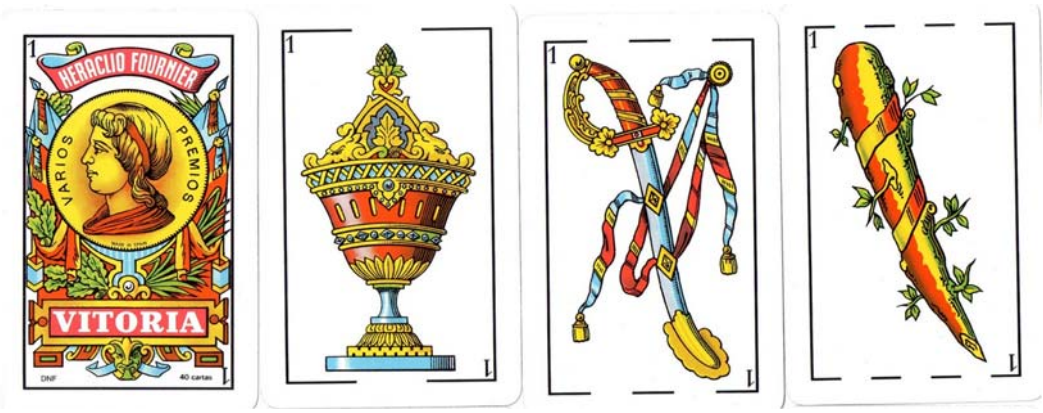


Figure 1 Suits of the Spanish deck: coins, cups, swords and cubs from left to right.

The game can be played by a variable number of players and it is comprised by the following stages.

1. Initially, each player puts an initial amount of money in the centre of the table in order to comprise the pot.
2. Four cards are dealt to each player.
3. Each player chooses whether he wants to bet or not and the amount of money bet. The bet cannot be superior to the pot. The player will have to pay or receive from the pot the amount of money bet depending on if he wins or loses.
4. Once each player has made his bet, the top card from the deck is turned over.
5. The player wins if he owns a higher card of the same suit of the uncovered card.

This work is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License (CC BY-NC-ND 4.0).

In brief, the logic behind the game is quite simple as the probability of winning or losing can be straightforwardly obtained by dividing the number of cards which defeat the 4 cards owned by the player by the total amount of cards in the deck. Nevertheless, a KDD methodology was proposed in order to teach students a more general procedure that could be extended to model more complex situations.

In order to ease the work, it was considered that only one player was playing at each time. The development of the expert system was divided into the following steps:

1. The students received a training set composed by 1000 plays, with the cards owned by the player and the uncovered card.
2. Based on the training set, each group of students had to select the best classification technique. They could choose between: decision trees, decision rules, Bayes-based methods, logistic regression, neural networks, support vector machines or instance-based learners, and besides, they had to select the proper metric to evaluate each classifier: *Kappa statistic*, *Precision*, *TP-Rate*, *Recall* or *F-Measure* among others (Witten & Frank, 2005). This process was carried out in the WEKA suite (Hall et al., 2009).
3. Once each group had selected the best classifier, students received a testing set composed by 500 additional plays, where they had to decide how to use the model to support the decision to bet or not. Then, for each play, the classifier predicted if the student was going to lose or win and its probability. Afterwards, for the wining plays, each group had to set a threshold to decide if the probability of wining was high enough to bet. Finally, if the decision of betting was taken, the system had to decide the amount of money to bet. Students fine-tuned this expert system based on a trial-and-error procedure until they maximized their gains.
4. Once every group had completed the training stage, each group uploaded a *.doc* file with all results and steps followed in the process, the selected model, and the threshold and formula which defined the DSS.
5. Lastly, the professor organized a competition in class with a new testing set composed by other 500 additional plays unseen by the expert systems of each group. The grade obtained by each group was related to the amount of money gained in this competition. The group with more gains was graded with a 10/10, while the remaining groups were graded proportionally.

The experience was carried out along two session, of 2 hours each, within practical classes of the subject. During the first session steps 1 and 2 were accomplished. Steps 3, 4 and 5 (final class competition) were developed during second session.

Results and Discussion

The afore-mentioned methodology was applied during four different courses of "Knowledge Engineering" subject in the Computer Engineering Grade at University of La Rioja, from academic years 2010/2011 to 2014/2015. Results presented herein describe results obtained from academic year 2013/2014, with 25 students.

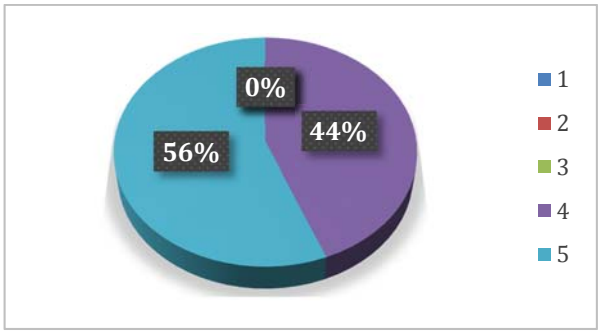
This cards game is a stochastic and probabilistic problem based on determining the probability of finding a win-case (hereinafter referred to as WIN). The database was not balanced in order to keep probabilities of appearance as they were. If balanced, the model could derived a wrong probability distribution and therefore, provide good results in cross-validation but then fail when testing with new data.

This work is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License (CC BY-NC-ND 4.0).

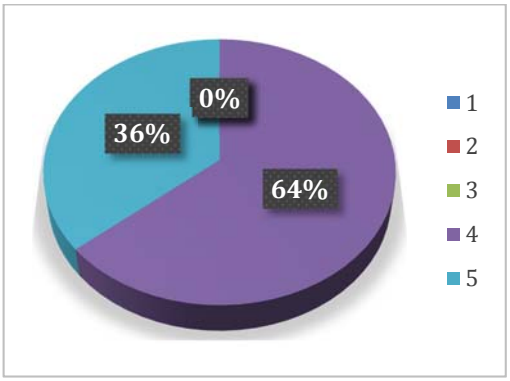
Students had to deal with statistical metrics such as *TP-Rate* or *Sensitivity*, which indicates the percentage of winning cases the model detects from all WIN's in the database; *Precision*, which shows the percentage of cases guessed rightly from all WIN's predicted by the model; *F-Measure* which is the harmonic average of previous TP-Rate and Precision; and others. As a result, different strategies came up, such as very conservator models with very high *Precision* and low *TP-Rate*, or models with higher risk with higher *TP-Rate* and lower *Precision*.

At the end of the course, a survey was conducted to quantify the level of satisfaction of the students with the methodology proposed, whose results are shown below.

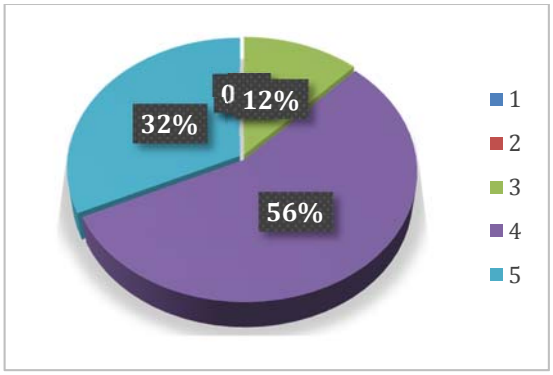
1. From 1 to 5 (where 1 represents 'very low' and 5 'very high'), how interesting did you find the card game competition based-learning compared to traditional teaching model?



2. From 1 to 5 (where 1 represents 'very low' and 5 'very high'), how do you grade the students' involvement in the card game competition based-learning?



3. From 1 to 5 (where 1 represents 'very low' and 5 'very high'), do you think the number of practical cases is enough for a correct understanding of the subject?



This work is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License (CC BY-NC-ND 4.0).

The following positive experiences were obtained:

- High level of attention and participation of students during the lessons corresponding to soft-computing and DSS. All students considered involvement to be "high" or "very high" in class, as well as the degree of interest, which resulted to be very high. This is very important for students as previously to this methodology they were over-whelmed by the great number of techniques taught in a relatively short period of time.
- Class competition to optimize results can be considered a success, since students are motivated because they consider learning as a game indeed.
  Students combine theoretic and practical knowledge in soft-computing and DSS during the course to solve other real life projects, which can be approached in a similar way. 32% of students found that this experience was complete enough so as to proportionate a good understanding of the subject (while other 56% considered the experience almost complete enough).

Conclusions

This innovative teaching-learning project deals with the teaching of data mining and decision support systems in "Knowledge Engineering" course for students of Computer Engineering Grade at University of La Rioja, throughout the application of soft-computing techniques to conventional problems. A typical-within students' cards game ("copo") was chosen. Students had to learn and apply different soft-computing techniques and statistical metrics, and also understand possible strategies (conservator or higher risk player) to achieve best results. By making this task a class competition, students kept a high level of attention during the whole class and a high level of satisfaction in comparison to traditional teaching-learning method.

To conclude with, not only this methodology can be applied to 'Knowledge Engineering', but it could be also applied to other technical courses, where a great number of techniques and metrics are taught in a reduced period of time.

Acknowledgements

The authors would like to acknowledge the University of La Rioja for the financial support received through the call of proposals '*Proyectos de Innovación Docente 2014/2015*', as well as for the fellowships *FPI-UR* to train researchers (J. Antonanzas and R. Urraca). The authors also want to express their gratitude to the *Instituto de Estudios Riojanos* (IER), to the *Banco Santander* for the project PROFAI 13/06 and to the *Agencia de Desarrollo Económico de La Rioja* for the project ADER-2012-I-IDD-00126 (CONOBUILD). Finally we would like to thank to the special collaboration of all student involved in this experience.

References

Andreu-Andrés, M. A. & García-Casas, M. (2011). Perceptions of Gaming as Experiential Learning by Engineering Students. International Journal of Engineering Education Vol. 27(4), 795–804.

Bergin, S., & Reilly, R. (2005). The influence of motivation and comfort-level on learning to program. Proceedings of the 17th workshop on psychology of programming-

This work is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License (CC BY-NC-ND 4.0).

PPIG'05, 293-304.

Boss, S., & Krauss, J. (2007). Reinventing project-based learning: your field guide to real-world projects in the digital age. Eugene, OR: International Society for Technology in Education.

Buck Institute for Education, An overview of Project Based Learning (2002). http://www.bie.org/pbl

Burguillo, J. C. (2010). Using game theory and Competition-based Learning to stimulate student motivation and performance. Computers & Education 55, 566-575.

European Space of Higher Education. http://www.eees.es/

Gredler, M. E. (2004). Games and Simulations and their Relationships to Learning. Handbook of Research on Educational Communications and Technology, 2nd ed., Lawerence Erlbaum Associates, Mahwah, NJ, 571–582.

Hall, M., Frank, E., Holmes, G., Pfahringer, B., Reutemann, P., & Witten, I.H. (2009). The WEKA data mining software: an update. SIGKK Explorations, 11(1).

Merrill, M. D. (2007). A task-centered instructional strategy. Journal of Research on Technology in Education, 40(1), 33-50.

Reynolds, W.M. & Miller, G.E. (2013). Educational psychology: Contemporary perspectives. Handbook of Psychology, Educational Psychology, 7 Wiley, Hoboken, NJ, 1–22.

Slavin, R. E. (1980). Cooperative learning. Review of Educational Research, 50(2), 315-342.

Witten, I.H. & Frank, E. (2005). Data mining: practical machine learning tools and techniques. San Francisco, CA. Morgan Kaufmann Publishers Inc.

This work is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License (CC BY-NC-ND 4.0).