Predicting students' outcome by interaction monitoring

Samara Ruiz
Department of Languages and Computer Systems
University of the Basque Country, UPV/EHU
San Sebastian, Spain
samara.ruiz@ehu.es

Maite Urretavizcaya
Department of Languages and Computer Systems
University of the Basque Country, UPV/EHU
San Sebastian, Spain
maite.urretavizcaya@ehu.es

Isabel Fernández-Castro
Department of Languages and Computer Systems
University of the Basque Country, UPV/EHU
San Sebastian, Spain
isabel.fernandez@ehu.es

ABSTRACT

In this paper we propose to predict the students' outcome by analyzing the interactions that happen in class during the course. PresenceClick lets teachers and students register their interactions during learning sessions in an agile way to give feedback in return about the students' learning progress by means of visualizations. Some of the registered interactions are the students who are attending class and a subset of the students' emotions felt during learning sessions. We have found correlations among attendance, emotions and performance in the final exam. This paper presents the study carried out to build a prediction model for the students' mark in the final exam based on these interactions. The purpose is to advice teachers about students in risk to fail.

Keywords

F2F interactions, mark prediction, linear regression, decision tree

1. INTRODUCTION

Drop out or failure is a common issue related to university students. Many studies have been carried out to detect students' problems, or even to predict the students' outcome, by applying data mining techniques to their interactions with intelligent tutoring systems [1] or course management systems [2] [3]. Other works include a wide range of potential predictors –i.e. personality factors, intelligence and aptitude tests, academic achievement, previous college achievements, or demographic data− to predict drop out and students' performance in high school [4] [5]. However, these works leave aside all the information that can be collected from the interactions that happen in face-to-face learning, the most extended way of education.

During traditional learning courses there is no way to detect problems or to know the performance of students in the final exam, except applying the teacher's intuition on the in-class students' interactions. This is even more difficult as the number of students in class grows, which is a current common issue at university worldwide. In this line, this papers aims to answer the next research questions: *Is it possible to predict the students' outcome by analyzing the interactions that happen in class? And, can we detect any interaction that especially influences the mark?*

2. PRESENCECLICK

PresenceClick is a distributed and modular environment that captures the interactions in learning sessions in an agile way. On the one hand, the *AttendanceModule* automatically captures the list of attendees to class. On the other hand, the *EmotionsModule* lets teachers capture the emotional state of the classroom related to whatever specific activity of the course. Students quantifies their emotions (six positive –*enjoyment, hope, pride, excitement, confidence* and *interest*– and six negative –*anxiety, anger, shame, hopelessness, boredom* and *frustration*–) in a 6-likert scale questionnaire based on the models described in [6] and [7]. The analyzed data belong to two subjects of Computer Science:Modular and Object Oriented Programming, (MOOP) and Basic Programming (BP). In MOOP 97 students were enrolled whereas 81 students participated in BP. The data were collected asking students to fill different event questionnaires. The MOOP students were asked three times to fill events where 41, 20 and 41 students responded respectively. The BP students were asked six times and 56, 36, 57, 48, 29 and 13 students participated (last event participation was low due to a server problem).

3. PREDICTING OUTCOME

Building a predicting model for students' outcome in the final exam was aimed to let teachers foresee those students that could be in risk to fail in the subject or even drop out.

In MOOP 44 students out of 97 enrolled attended the exam and 50 responded at least one emotion event, while 59 students attended the exam from 81 students enrolled in BP and 68 responded at least one emotion event. As the students dropping out the subject precisely are an important sample set to study, and as a considerable number of students did not attend the exam in both subjects, three different cases were studied: (Case1 - NA=F) Students non attending the exam were not considered; (Case2 - NA=T; mean=F): Students non attending the exam were assigned 0 as mark; (Case3 - NA=T; mean=T): Students non attending the exam were assigned the mean of the fails as mark, where fails are all the students with mark<5.

The three phase experiment that follows was carried out.

3.1 Phase 1: Correlation analysis

Pearson-correlation analysis was conducted between mark-attendance and mark-emotions. All the positive/negative emotions were gathered together, and the mean from all the events where each student participated was calculated in order to normalize the data. Table 1 shows the correlations for the three cases between mark-attendance, mark-positive emotions and mark-negative emotions. In both subjects *attendance and students' negative emotions influence the mark in the final exam* (except when non attendees to exam were not considered in BP) according to literature (p>|0.3|) [8]. Student's positive emotions influence the mark only in MOOP. This could be due to the fact that being aware of the negative emotions is usually easier than being aware of the positive ones. In addition, we could also suppose that students expressing negative emotions in questionnaires are not lying, whereas students could increase the value of their positive emotions in order to be closer to the group feelings.

Table 1. Correlations with the mark in MOOP

	Case	Attendance	Pos emo	Neg emo
MOOP	**NA=F**	0.45 (p=0.0048)	0.45 (p=0.0056)	-0.46 (p=0.0034)
	NA=T, mean=F	0.6 (p=4.02e-06)	0.46 (p=0.0008)	-0.65 (p=3.78e-07)
	NA=T, mean=T	0.54 (p=4.7e-05)	0.46 (p=0.0008)	-0.59 (p=5.45e-06)
BP	**NA=F**	0.25 (p= 0.071)	0.13 (p= 0.35)	-0.29 (p=0.034)
	NA=T, mean=F	0.48 (p=0.0004)	0.28 (p=0.019)	-0.34 (p=0.0042)
	NA=T, mean=T	0.39 (p=0.0009)	0.23 (p=0.054)	-0.33 (p=0.006)

3.2 Phase 2: Multiple linear regression

In this stage of the experiment we looked for a model with a multiple linear regression analysis to predict the numeric mark of the student. For both subjects, 2/3 of the population was taken for training while the remaining was taken for validation. The three variables together were tested as dependent in order to predict the mark (w+ x* attend.+ y* posEmotions+ z* negEmotions). However, for all cases the standard deviation of the model prediction error rounded two points, which implies a margin too big (in a scale grade from 0 to 10, where fails are above 5). All the emotions were also studied individually to check if any of them could explain the mark, but the error rounded the two points.

3.3 Phase 3: Classification tree

Finally, we ran a decision tree to predict whether a student drops out, fails or passes the exam. Data from both subjects were normalized and gathered in a unique dataset, and different models were tested taking into account different variables in order to find the one that better predicted the students' performance. Figure 1 presents the decision tree for the training set that best predicted the students' performance taking into account the *attendance* and the *students' negative emotions.*

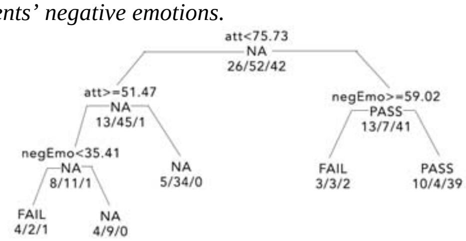

Figure 1. Training set's classification tree

As we can see in table 2 failed students are not well predicted with a 30% precision and 50% recall (F_1=37,5%), but dropping out students (F_1=86,36%) and passing students are quite well predicted (F_1=81,63%). The low correction of the fails could be due to the fact that few students are in this category and more data is required to refine the model. However, we consider that the most important measure is the recall for drop out and fail, in order to discover the students in risk and make the teacher aware. Taking into account that only 16% of failed students and 8,4% of drop out students have been predicted with PASS, we can conclude that the model is quite good, although a major sample is needed in order to adjust it for a better prediction.

Table 2. Predictions table

		Real				
		FAIL	**NA**	**PASS**	**Precis.**	**Recall**
Class	**FAIL**	3	3	4	30%	50%
	NA	2	19	2	82,61%	90,48%
	PASS	1	2	20	86,96%	76,92%

4. CONCLUSIONS

This paper has presented the preliminary study developed to propose a predicting model for the students' outcome in the final exam based on the interactions captured by the PresenceClick system. Those interactions data give teachers and students the possibility to avoid failure and drop out. So far, we have tested the *attendance to class* and *the students' emotions* as model predictors. The study was divided in three phases: correlation analysis, multiple linear regression and decision trees. We founded that *attendance* as well as *student's emotions* influence the mark. In particular, the *negative emotions* together with the *attendance* seem to be the interactions with bigger influence on the mark, although the multiple linear regression did not provide an accurate model. However, the decision tree brought us the possibility to foresee the students' performance in the final exam according to these factors, although a major sample is needed in order to refine the model.

5. ACKNOWLEDGMENTS

This work has been supported by the Govern of the Basque Country (IT722-13), the EHU/UPV university (PPV12/09, UFI11/45) and Gipuzkoako Foru Aldundia (FA-208/2014-B).

6. REFERENCES

[1] Baker, R., Corbett, A., Koedinger, K., 2004. Detecting Student Misuse of ITS, in: Lester, J., Vicari, R., Paraguaçu, F. (Eds.), ITS, Lecture Notes in CS, pp. 531–540.

[2] Romero, C., Ventura, S., Espejo, P.G., Hervás, C., n.d. Data mining algorithms to classify students, in: In Proc. of the 1st Int. Conf. on Educational Data Mining (EDM'08), P. 187191, 2008. 49 Data Mining 2009.

[3] Calvo-flores, M.D., Galindo, E.G., Jiménez, M.C.P., Pérez, O., n.d. 586 Current Developments in Technology-Assisted Education (2006) Predicting students ' marks from Moodle logs using neural network models.

[4] Kabakchieva Dorina, 2013. Predicting Student Performance by Using Data Mining Methods for Classification. cait 13,61.

[5] Pal, A.K., Pal, S., 2013. Data Mining Techniques in EDM for Predicting the Performance of Students. International Journal of Computer and Information 11/2013; 2(6):1110-1116.

[6] Pekrun, R., Goetz, T., Frenzel, A.C., Barchfeld, P., Perry, R.P., 2011. Measuring emotions in students' learning and performance: The AEQ. Contem. Educat. Psych. 36, 36–48.

[7] Arroyo, I., Cooper, D.G., Burleson, W., Woolf, B.P., Muldner, K., Christopherson, R., 2009. Emotion Sensors Go To School, in: Proc. 14th Conference on Artificial Intelligence in Education, pp. 17–24.

[8] Gray, G., McGuinness, C., Owende, P., 2013. An Investigation of Psychometric Measures for Modelling Academic Performance in Tertiary Education, in: Sixth International Conference on Educational Data Mining.