International Journal of Engineering & Technology, 7 (2.28) (2018) 106-109

International Journal of Engineering & Technology

Website: www.sciencepubco.com/index.php/IJET

Research paper



Color code method design evaluation and data analysis

Laura Dzelzkaleja[1]*

[1]Riga Technical University Kalku street 1, Riga, Latvia
**Corresponding author E-mail: laura.dzelzkaleja@rtu.lv*

Abstract

Nowadays, with the growing attention to Educational data mining, it important to use the data analysis results to improve course assessment and improve the e-learning process. A new Color code method for understanding students' learning process has been introduced in the previous research. The method uses three color codes: red for "problem", yellow for "work in progress" and green for "job done". The method has proven to work in a classroom environment and recently implemented in the edX learning platform. In this paper the button design is tested in four student groups and the button press data analyzed taking into consideration students' gender and exam marks. Research show that 18% of the students used buttons without difference in gender and that highest button usage activity had the least course non-finishers and the highest number of students that got the highest mark. Survey showed that button design needs some improvement.

Keywords: *Data analysis; Educational data mining; Students behavior prediction*

1. Introduction

Data analysis in educational field is an important and promising way to build knowledge about student's learning process, developed course material success, student satisfaction and knowledge development. Nowadays, with the growing attention to ICT tools and solutions and their usage in education the explosive growth of educational data is available, and there is an opportunity to use this data to improve the quality of managerial decisions (Al-Twijri 2015). Especially true this is in e-learning and blended learning.
The quality of higher education institutions implies providing the services, which most likely meet the needs of students, academic staff, and other participants in the education system. Data Mining can help institutions of higher education to make more effective decisions as to improve the quality of instruction and services (Al-Twijri 2015).
Data Mining is very useful especially when examining students' learning behavior in online learning environment. This is due to the potential of data mining in analyzing and uncovering the hidden information of the data itself which is hard and very time consuming if to be done manually. In choosing the appropriate algorithms, researchers must first design the data and align it with the desired output (Khadijah 2013).
Educational Data Mining (EDM) is the application of Data Mining techniques on educational data. The objective of EDM is to analyze such data and to resolve educational research issues. EDM deals with developing new methods to explore the educational data and using Data Mining methods to better understand student learning environment. The EDM process converts raw data coming from educational systems into useful information that could potentially have a great impact on educational research and practice. Educational data mining uses many techniques such as Decision Trees, Neural Networks, Naive Bayes, K-nearest neighbor and many others. Prediction and analysis of student performance is an important milestone in educational environment (Kaur 2015). A major key application area of prediction in EDM is predicting student educational outcomes (Asif 2017).
In this paper there are experimental data and research data analyzed. The data are gathered from university students' behavior and exam results in a blended learning course. This research is a continuation of the previous work, where a new learning process assessment method was introduced.

2. Methodology

2.1. Experimental group description

As an experimental group was chosen Riga Technical University 1st Bachelor course students. The learning subject was Entrepreneurship. Registered students' number at the beginning was 191, 139 of them being males and 52 females. Seven (5 males, 2 females) of them dropped out of the course and didn't appear in the exam mark list, but two (males) form the previous years joined the course during the semester and was present in the exam mark sheet. In was decided to leave their data out of the data analysis, since they do not provide the full semester data which was decided to be an important characteristic of the experimental group, since it provides the possibility to compare students adequately. So, the total valid number of students is 182, 132 (73%) of them being males and 50 (27%) females.
The students were split into 4 groups, but the learning curricula was the same for all. The experiment took part from the end of September 2017 until the February 2018, but the database was active and started to record only from in the October 19. The last data were mined from January 23rd. The learning form was blended learning and for online learning purposes they used edX learning platform, which they were first introduced to in this course.

Copyright © 2018 Authors. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

2.2. Color code method description

A new approach, presented in the authors' previous papers (Dzelzkaleja, 2016 and Dzelzkaleja, 2017) for continuously assessing the learning process in a real time was presented. In this paper, the method is further analyzed with the data gained from the data mining and analysis.

The main principle of the method is as follows: there are three color codes which are used by a learner to show the teacher the progress in every moment of the learning process.

• "Red" is used to show that the task is not clear, or difficulties have appeared during the process, some assistance is needed (in the form of tutor or some extra learning materials);

• "Orange" is used whenever the task is being done and everything is clear – no need for assistance;

•"Green" is used when the learner has finished the task or isn't doing anything.

The colors have been chosen based on the traffic-lights color coding, since these colors are recognizable for almost every person and gives an opportunity to intuitively guess the meaning of the buttons – red as a something that slows down or stops, green is something that allows you to go to the next place and is connected to pleasurable associations, and orange (yellow in some cases) being something in the middle of both previous.

In the case of e-learning, the codes need to be installed so that the learner could click on the appropriate color on the screen conveniently in every moment of the learning process. It is important to note that this method is created to collect continuous rather than discrete data. For example, it can show that from time x till time y the student has been in the learning process, from time y till time w the student struggles with a problem and so on. It doesn't ask from the user to decide whether they like every single learning item, but just ask to record, when and how their learning situation changes.

The buttons module was developed in the edX platform in the summer of 2017. The edX learning platform was chosen due to its open code, possibility to add modules and growing popularity and prestige. The students' group were first to test this new module in the edX platform.

The use of the buttons is a voluntary choice. In the future it is planned to make user interface with data visualizations to motivate the students to use them, but in this experiment the students were only encouraged from the researcher and the course instructor to the buttons as a kind help for the research.

2.3. Data from buttons

Data analysis was done mostly using IBM SPSS Modeler, which is a tool that allows to process a big amount of raw data from a database and predict the future events with a help of models. For some of the data summaries and visualizations MS Excel was used. The information about button presses was further compared to course exam final marks as well as researched the influence of student's gender to the button usage.

As mentioned before, these type of data is potentially big-data, but in this particular case the total number of the data units were small due to the small number of students, button usage time – only one semester and the low motivation, since not feedback to the students is yet possible. The total number of valid presses is 129, 70% of them being the green button "done", 21% being the orange button "in progress" and 7% being the red button "problem".

2.4. Surveying

In the end of February 2018, after the semester ended, a survey about color code buttons was sent to all the students, and they filled it voluntarily. 14 valid surveys (8% of the students) were received, free-form questions were answered by less students. The student gender proportion in the survey the percentage was 71% to 28%, so a slight difference from the overall in the course, but within the error limits. The age of the students surveyed was 19-24, most (50%) being 19 years old and 20 (36%) years old.

The survey consisted of nine questions and student background information. The background information provided student's age, gender and group. The questions were about noticing color buttons in the learning platform, about the button design, about the button usage frequency and reasons and about situation where they used buttons, why and when they would use them.

3. Results and discussion

Students responded positively to most of the button design features, except the button respond reaction, that got more negative votes that other design elements. In the Figure 1. You can see the summary of the votes about the design section.

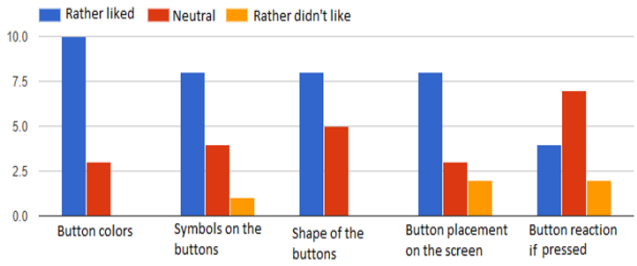


Fig. 1. Button design evaluation by students

The survey shows that some design changes as well as conceptual changes need to be implemented for the button visual reaction to a press. It seems that there is some confusion with the meaning and distinction between the colors and whether the pressing and response is recorded. A solution to this might be to make the buttons such as they would visually reflect the process more than discrete pressing events. The proposition it to remove the pop-up text box with the information about the registered press. Instead the buttons could light up light light-bulbs, showing in which process stage the student is at. Once a button is pressed, it lights up. Once another is press, the previous turns out and the newest pressed light up. The only difference would be with the green button – because finishing is an event and not a process. Hence, green button could turn out automatically after some seconds. In this way a room for a forth element appears: non-lit buttons, that means that a student takes a break from learning and is not actively involved in the learning platform even if he/she is not logged out from the platform. This idea differs slightly from the initial one, where only three different choices were available and with "green" could also be understood "taking a break".

There is also a room for improvement in the design of button meaning and the explanation of system meaning – a kind of a user manual. So far, this section has not been included in the design, but research show that it is very much needed. Even after the oral presentation of the system some of the students mentioned in the survey that they had forgotten the meaning of each button, from the survey it was also noted that there is a misunderstanding about the meaning of the buttons, since this system differs from the systems that are more commonly in use in students' everyday life, like voting systems and sentiment expression systems. Hence, the buttons need to be complemented with a thorough explanation about their meaning, and to particularly stress out that it is not a voting or opinion expressing system, but only the factual progress report – no need for decision making and extra cognitive load, just a pure non-emotional information about what the students is doing at each moment. The need for more information and a justification is also expressed in the survey with an opinion that "I don't understand, what the goal of using the buttons is, so I don't see the point of using them." There was another response in the section of design, that need to be taken into consideration and is probably connected to the implementation phase of the buttons – that the buttons where not visible at some moments in the learning platform. And that relates to the fact that some upgrades and maintenance

were made by the technicians to improve the work of the buttons. Probably, in the future it would be wise to notify the user that buttons are temporarily not available.

The students' activity in pressing the buttons was quite low – only 18% of the students used the color buttons at least once. When analyzing the data in terms of gender differences, no significant difference was found: both genders used the buttons the same – 18%. More preliminary results are presented in the previous research (Dzelzkaleja, 2018).

Student's motivation to use the buttons was only the kind request from the course instructor and the researcher, but this student test group didn't get any added value to pressing the buttons. Having that in mind, the results about the number of pressed buttons seems reasonably high. In the future the main motivation for a fully operative system will be a possibility to provide a fast feedback for the course instructor and be sure that the course instructor reacts to the provided information, as well as possibility to track one's own progress throughout the learning course and spot the learning process problems and risks easier. This is also what students say in the survey: students say that there is a point using these buttons, if the teacher pays a special attention to the learning materials that are problematic for the student as well as tell the teacher about the progress.

When asking, in what situation they would use the buttons, the answers were as follows: in the situations where there is a possibility to react faster; if I struggle or to show that I like the material (shows misunderstanding of the button meaning); Don't need them; to record the progress; Don't know; After reading materials; To report a context error in a text material; where multiple answers are needed; for positive and negative feedback. The answers show the lack of understanding and motivation as discussed previously.

Overall, the survey answers gave an impression that it was filled mostly by the students who had used color buttons at least once.

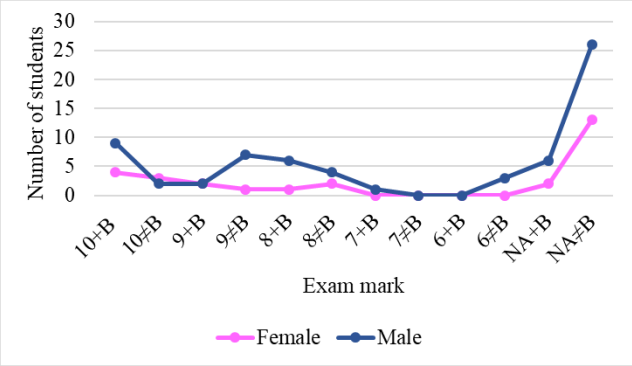


Fig.2. Exam mark correlation with gender and button presses

When looking at the exam marks, it was interesting to see, whether there is a correlation between the final course exam mark and students who used color buttons. There were many students, that hadn't finished the course after the end of the semester and the two additional weeks, so they hadn't received any mark – not attested for the finishing "NA". There were 47 NA students, 32 (68%) males and 15 (32%) females. This proportion differs by 5% from the overall course gender proportion – more female students hadn't finished the course. The research showed that most (83%) of these students didn't use the color buttons as well. So, it shows that the students that are less engaged in the course overall is less engaged learning tools as well.

In Latvia the lowest mark possible is 1 and the highest is 10. The lowest mark received in the exam during this research was 6 and three students got this mark, the second lowest was 7 and only one students got this mark and he also used the buttons. These students didn't use the buttons were as well. Situation is different with the one who got the best marks: 72% of the 18 students that got 10, 33% of 12 students that received 9 and 54% of the 13 students that received 8 used the buttons. In the Figure 2 correlation between gender, exam marks and buttons presses are presented.

Table 1: Students Performance Difference Among Study Groups

Group N°	Proportion of button users	Students with mark: NA	Students with mark: 6 or 7	Students with mark: 10
1	36%	14	0	5
2	50%	8	1	8
3	25%	14	1	5
4	26%	10	2	0

There were noticeable differences also among the four different student groups. In one group as much as 50% of the students had used the buttons at least once, in one group it was 36%, but in other two only about one quarter of the students. It can be seen in the Table 1.

4. Conclusion step before the final submission

Some conclusions and predictions from the button usage activity in different study groups can be made, since the 2nd group with the highest button usage activity had the least course non-finishers and the highest number of students that got the highest mark. The same tendency doesn't fully apply to other groups, but it can be seen though, that the second highest button activity correlates with the second best top mark student characteristic and no lowest grades, but the most course non-finishers as well. The tendency of the bottom two button using groups is that overall, they have more lower marks, less top marks and more non-finishers. But these suggestions need more data for verification.

Students' motivation to use the color codes without any personal gain is low, and the user interface with data visualization where students can follow their learning progress and to compare to the benchmark could provide this motivation both to the students and to the teacher, since it the engagement is also very dependent on the teacher's ability to justify the code usage and to remind of it time by time. Only when the analytical results are displayed in a user-friendly way are they effectively utilized by users. Reports, histograms, pie charts, regression curves, etc. are frequently used to visualize the results of data analysis. This leads to the topic of visualization being one of the main challenges in mining big data (Chen 2014).

When analyzing the data in terms of gender differences, no significant difference was found. Both genders used the buttons the same – 18% of the total. So, it seems that no significant button design differences should be implemented due to the gender differences. But the survey shows that some design changes as well as conceptual changes need to be implemented for the button visual reaction to a press. It seems that there is some confusion with the meaning and distinction between the colors and whether the pressing and response is recorded, so they should be solved in the future. There is also room for improvement in the design of button meaning and the system meaning explanation – a kind of a user manual. So far, this section has not been included in the design, but research show that it is very much needed.

Acknowledgement

I want to say big thanks to Jānis Kapenieks (jun.) for letting me do the research in his course and helping with the students' data. Thanks to Viktors Zagorskis for helping with data digging from the learning platform. This research has been funded by international European Project under the FLAG-ERA Joint Transnational Call (JTC) 2016 FuturICT 2.0.

References

[1] Al-Twijri MI, Noaman AY (2015), A New Data Mining Model Adopted for Higher Institutions. *Procedia Computer Science* 65, 836-844.

[2] Asif R, Merceron A, Najmi AAS, Haidera G. (2017), Analyzing undergraduate students' performance using educational data mining. *Computers & Education* 113, 177-194.
[3] Chen SM, Liu MY (2014), Big data: a survey. *Mobile Netw. Appl.* 19, no. 2, 171-209.
[4] Dzelzkaleja L, Kapenieks J (sen.) (2016), Real-time Color Codes for Assessing Learning Process. *International Conference; Meaning in Translation: Illusion of Precision, 11-13 May. Riga: Procedia - Social and Behavioral Sciences*, 263-269.
[5] Dzelzkaleja L, Timsans Z (2018) Colour Codes Method Digitalization in edX E-learning Platform. *Paper confirmed for publication. 10th International Conference on Computer Supported Education* 15-17 April, Portugal, Madeira.
[6] Dzelzkalēja L (2017), Real Time Color Codes in a Classroom. *Proceedings of 9th International Conference on Computer Supported Education*, 21-23 April, Portugal, Porto, 111-117.
[7] Kaur P, Singh M, Josan GS (2015), Classification and prediction based data mining algorithms to predict slow learners in education sector. *3rd International Conference on Recent Trends in Computing 2015 (ICRTC-2015). Procedia Computer Science* 57, 500 – 508.
[8] Khadijah S, Tasir ZM (2013), Educational Data Mining: A Review. *Procedia - Social and Behavioral Sciences* 97, 320-324.