Journal of Engineering and Applied Sciences 12 (22): 5944-5948, 2017
ISSN: 1816-949X

Use of data mining to identify trends between variables to improve implementation of an immersive environment

[1,2]Ronald Zamora-Musa and [3]Jeimy Velez

[1] Department of Computer Science and Electronica, Universidad de la Costa, 130012 Barranquilla, Colombia

[2] Universidad Pontificia Bolivariana, 050004 Medellin, Colombia

[3] Corporación para la ciencia, la investigación, la innovación y el emprendimiento, CORCIEM, Monteria, Colombia

Abstract: Globally the implementation of immersive environments for learning activities have been in constant growth which indicates that their development must improve daily. For this reason, this study identifies trends (co-occurrences) and relationships between variables associated with an immersive environment to improve its implementation. Results were found which show that a good design of information guides, organization of menus and useful instructions generates that the users enjoy using the immersive environment for learning and foments recommendations of use to other users.

Key words: Immersive environment, e-Learning, data mining, educational data mining, association rules mining

INTRODUCTION

An immersive environment is a three-dimensional space where users represented by avatars perform education, work and entertainment activities as if they were present in that virtual place. These environments are a trend in the educational sector, because they allow the interaction of users with materials and tools in a 3D space for an immersive experience, (Comas-Gonzalez et al., 2017; Kovács et al., 2015).

Immersion is a concept that generates differentiation in learning, contributing with positive effects for education through 3D environments (Zamora et al., 2016; Cho et al., 2015). Therefore, immersion is a recent trend in e-learning, fostering the construction of knowledge in an innovative way (Pollock & Biles, 2016; Peng et al., 2015).

It is important to mention that for the implementation of immersive environments based on Information and Communication Technologies (ICT) reaches a positive impact, must be necessary, planning and design according to processes related to the user, and not only technological processes (Khalifah et al., 2017; Freire et al., 2016; Tawil et al., 2012), i.e, the use of ICT in education without planning generates e-learning resources without the necessary elements that allow the users to reach the desired objectives (Zamora-Musa et al., 2017; Arantes et al., 2016, Zamora-Musa & Villa, 2013).

In the same way, education through immersive environments in addition to being related to ICT is also associated with the pace of learning, (Long and Qing-Hong, 2014) and training needs of users (Paez et al., 2017), Therefore, it is necessary to use data mining to identify trends between variables to improve implementation of an immersive environment.

Data mining has been widely used in information systems, engineering, marketing, among others (Maqsood, 2017; Poorani et al., 2014) , but according to Angeli et al., (2017), in the last 10 years, it is beginning to be used in the field of education, to improve the implementation of virtual learning platforms (e-Learning) as immersive environments. The use of data mining allows to find valuable information from organized data, information that is important for discovering trends and relationships (Medvedev et al., 2017; Mendoza et al., 2015; Marengo et al., 2013).

The results obtained when applying data mining are important to improve learning processes, for example, how to design or redesign a learning environment (Lovkesh, 2016; Romero & Ventura, 2012). In related works, data mining in the field of e-learning has been used to give feedback on course structure (Merceron & Yacef, 2010), user performance prediction (Ahmed & Elaraby, 2014), design suggestions to improve environment (Udupi et al., 2016; Romero & Ventura, 2012), among others.

Corresponding Author: Ronald Zamora-Musa, Universidad de la Costa, 130012 Barranquilla, Colombia

In this article, we apply the rules of association, which is a technique of data mining, its objective is to determine rules of the form "if-then", where a set of values has associations with another set of values, generating a prediction or behavior of those set of values (Kumar & Pathak, 2016; Baker & Siemens, 2014; Lin et al., 2002). There are other data mining techniques such as: clustering (Gunasekara et al., 2014), sequential pattern mining (Mohajer et al., 2016) and regression (Buja & Lee, 2001).

MATERIALS AND METHODS

The objective of this research is to improve the implementation of an immersive environment through data mining, for which we must discover trends (co-occurrences) and relationships among the variables associated to a survey made to users. For this purpose the following methodology is developed: data selection, pre-processing and data transformation, using Weka software.

In the research, users are represented by avatars and perform learning activities, a preliminary study is implemented where users in subjects corresponding to "digital circuits" interact with the immersive environment and then proceed to answer the survey found in the link: http://sumi.uxp.ie/en/

The survey seeks to know how the different variables associated with each of the 50 questions are related, the survey was answered by all users of the mentioned subject, which are 24 students.

The 50 questions in the survey are divided into 10 question packs, where the first 10 questions correspond to the efficiency of the environment, the second 10 questions to the influence of the environment, questions 21 to 30 correspond to the utility of the environment, Questions from 31 to 40 correspond to the control that is had on the environment and the questions from 41 to 50 correspond to the ease of learning in the environment.

Data selection: From the immersive environment survey we have the following information, for data selection:

@attribute Efficiency_1{Agree, Undecided, Disagree}, @attribute Efficiency _2{Agree, Undecided, Disagree }… @attribute Efficiency_10{Agree, Undecided, Disagree}

@attribute Influence_1{Agree, Undecided, Disagree}, @attribute Influence_2{Agree, Undecided, Disagree}… @attribute Influence_10{Agree, Undecided, Disagree}

@attribute Utility_1{Agree, Undecided, Disagree}, @attribute Utility_2{Agree, Undecided, Disagree}… @attribute Utility_10{Agree, Undecided, Disagree}

@attribute Control_1{Agree, Undecided, Disagree}, @attribute Control_2{Agree, Undecided, Disagree}… @attribute Control_10{Agree, Undecided, Disagree}

@attribute Learning_1{Agree, Undecided, Disagree}, @attribute Learning_2{Agree, Undecided, Disagree}… @attribute Learning_10{Agree, Undecided, Disagree}

Pre-processing and data transformation: A description of the data is made with some histograms of the 50 questions, for example to the first question: "This software Responds too slowly to inputs", it has the following histogram shown in Figure 1.

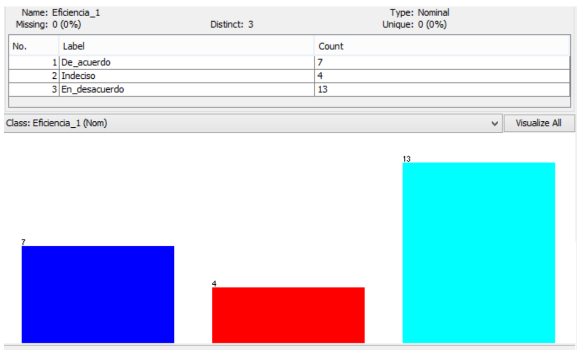

Figure 1. Histogram for question 1

In Figure 1, it can be seen that 7 users, i.e. 29.1% say that the software responds very slowly to the inputs and that 13 users, i.e. 54.1% say they disagree with the previous statement.

For question 12 "Working with this software is satisfying" it has the following histogram shown in Figure 2.

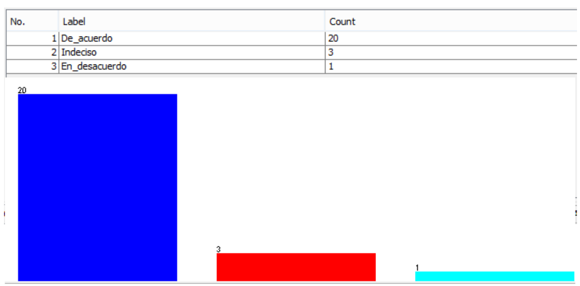

Figure 2. Histogram for question 12

In Figure 2, it can be seen that 20 users, i.e. 83.33% say that working with this software is satisfying.

For question 27 "Using this software is frustrating" it has the following histogram shown in Figure 3.

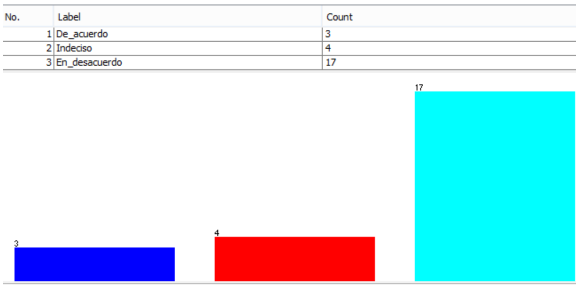

Figure 3. Histogram for question 27

In Figure 3, it can be seen that 17 users, i.e. 70.83% say they disagree that using this software is frustrating.

This research does not need to perform data transformation in any of the variables because they are all categorical, a condition that must be met when applying the unsupervised technique "association rules", which is used for this descriptive analysis Seeks to know how the different variables associated with each of the 50 questions of the immersive environment.

Finally, the "A priori" method is applied to determine how the different variables associated with each of the 50 questions of the immersive environment survey are related.

By applying the A priori method, the rules or associations are obtained; it is specified that no class attribute is configured in the Weka software, an evaluation measure of 90% and a number of rules of 40, taking into account number of variables

RESULTS AND DISCUSSION

When applying the unmonitored technique "association rules" to discover trends (co-occurrences) and relationships between variables associated with the immersive environment survey, the following rules are observed (for evaluation and interpretation purposes, 6 rules were chosen)

Rule 1: When approximately 80% of students (users) agree that "the information given by the software" can be understood and guided (question 23 - Utility_3) that percentage also agrees that "I would recommend this software To my colleagues "(question 2 - Efficiency_2); Then it can be said that a good information guide generated by the immersive environment ensures that students recommend their colleagues to work (study and learn) with this.

Rule 2: When approximately 80% of students (users) agree that "The organization of the menus seems quite logical" (question 33 - Control_3) also that percentage agrees that "I would recommend this software to my classmates" (Question 2 - Efficiency_2); Then it can be said that a logical organization of menus in the immersive environment ensures that students recommend their peers to work (study and learn) with this.

Rule 3: When approximately 80% of students (users) agree that "Instructions and aids are useful" (question 3 - Eficiency_3) and at the same time another 80% agree that "I enjoy working with this Software "(question 7 - Efficiency_7) also that percentage agrees that" I would recommend this software to my colleagues "(question 2 - Efficiency_2); Then it can be said that the existence of useful instructions and help, and enjoy working with the immersive environment ensures that students recommend their colleagues to work (study and learn) with this.

Rule 4: When 75% of students (users) disagree that "software is inconsistent" (question 21 - Utility_1) that percentage also agrees that "I would recommend this software to my classmates" (question 2 - Efficiency_2); Then it can be said that the immersive environment is consistent for the student's perception ensures that students recommend their peers to work (study and learn) with this.

Rule 5: When 75% of students (users) agree that "Working with this software is satisfactory" (question 12 - Influence_2) and at the same time another 75% agrees that "I can understand and guide me through Information given by the software "(question 23 - Utility_3), that percentage also agrees that" The way in which the system presents the information is clear and understandable "(question 13 - Influence_3).

Rule 6: When approximately 80% of students (users) agree that "I can understand and be guided by the information given by the software" (question 23 - Utility_3), 75% of students (users) That "I enjoy when I work with this software" (question 7 - Efficiency_7), then it can be said that a good guide in the information given by the immersive environment ensures that students enjoy studying and learning with the environment.

CONCLUSION

With the descriptive analysis made to the survey associated to the immersive environments to know how the different variables associated to each of the 50 questions, through the unsupervised technique:

"rules of association" it can be concluded that this technique offers valuable information to re-design or strengthen aspects of the immersive environment taking into account the results of the preliminary study.

For example it can be observed that a good design of information guides, organization of menus and useful instructions generates that the users enjoy using the immersive environment for the study and learning and that this one is recommended to its companions.

As future work can be mentioned that for the technique of association rules due to the large number of variables it is necessary to perform analysis with more rules to take into account important aspects for the final design of the immersive environment, in addition to performing a descriptive analysis of all rules.

ACKNOWLEDGMENT

The authors thank the support and financing of this project to the research group in Electronica GIECUC and CORCIEM.

REFERENCES

Ahmed, A. B. E. D., & Elaraby, I. S. (2014). Data mining: A prediction for student's performance using classification method. World Journal of Computer Application and Technology, 2(2), 43-47

Angeli, C., Howard, S., Ma, J., Yang, J., & Kirschner, P. (2017). Data mining in educational technology classroom research: Can it make a contribution?. Computers & Education, 113, 226-242.

Arantes, E., Stadler, A., Del Corso, J., & Catapan, A. (2016). Contribuições da educação profissional na modalidade a distância para a gestão e valorização da diversidade. Espacios, 37(22), E-1.

Baker, R. S., & Siemens, G. (2014). Educational data mining and learning analytics. In K. Sawyer (Ed.), Cambridge handbook of the learning sciences (2nd ed., pp. 253e274). NY: Cambridge University Press

Buja, A., & Lee, Y. S. (2001, August). Data mining criteria for tree-based regression and classification. In Proceedings of the seventh ACM SIGKDD international conference on Knowledge discovery and data mining (pp. 27e36). ACM

Chen, L. & Yang, Q. (2014). A group division method based on collaborative learning elements. In The 26th Chinese Control and Decision Conference (pp. 1701-1705). Changsha.

Cho, Y. H., Yim, S. Y., & Paik, S. (2015). Physical and social presence in 3D virtual role-play for pre-service teachers. The Internet and Higher Education, 25, 70–77

Comas-Gonzalez, Z., Echeverri-Ocampo, I., Zamora-Musa, R., Velez, J., Sarmiento, R. and Orellana, M. (2017). Tendencias recientes de la Educación Virtual y su fuerte conexión con los Entornos Inmersivos. Espacios, 38(15), p.4.

Freire, P., Dandolini, G., De Souza, J., Trierweiller, A., Da Silva, S., & Sell, D. et al. (2016). Universidade Corporativa em Rede: Considerações Iniciais para um Novo Modelo de Educação Corporativa. Espacios, 37(5), E-5.

Gunasekara, R., Wijegunasekara, M., & Dias, N. (2014). A Study on How to Improve the Perfomance of k-mean Data Mining Algorithm in a Parallel Environment. Journal Of Engineering And Applied Sciences, 9(10), 441 - 446.

Kovács, P., Murray, N., Rozinaj, G., Sulema, Y., & Rybárová, R. (2015). Application of immersive technologies for education: State of the art. In 2015 International Conference on Interactive Mobile Communication Technologies and Learning (IMCL) (pp. 283 - 288). Thessaloniki.

Kumar Ameta, G., & Pathak, V. (2016). A Survey on Improved Association Rule Mining for market based analysis. International Journal Of Advances In Computer Science And Technology, 5(12), 173-175.

Lin, W., Alvarez, S. A., & Ruiz, C. (2002). Efficient adaptive-support association rule mining for recommender systems. Data Mining and Knowledge Discovery, 6(1), 83-105.

Lovkesh. (2016). Enhancing E-Learning Through Data Mining in the Context of Education Data. International Conference On Computing For Sustainable Global Development (Indiacom) - IEEE, 109 - 113.

Marengo, A., Pagano, A., & Barbone, A. (2013). Data mining methods to assess student behavior in adaptive e-learning processes. Fourth International Conference On E-Learning "Best Practices In Management, Design And Development Of E-Courses: Standards Of Excellence And Creativity" - IEEE, 303 - 309.

Maqsood, A. (2017). Study of Big Data: An Industrial Revolution Review of applications and challenges. International Journal Of Advanced Trends In Computer Science And Engineering, 6(3), 31-34.

Medvedev, V., Kurasova, O., Bernatavičienė, J., Treigys, P., Marcinkevičius, V., & Dzemyda, G. (2017). A new web-based solution for modelling data mining processes. Simulation Modelling Practice And Theory, 76, 34-46.

Mendoza, F., De la Hoz, A., De la Hoz, E., & Ariza, P. (2015). Feature selection, learning metrics and dimension reduction in training and classification processes in intrusion detection systems. Journal Of Theoretical And Applied Information Technology, 82(2), 291 - 298.

Merceron, A., & Yacef, K. (2010). Measuring correlation of strong symmetric assocation rules in educational data. In C. Romero, S. Ventura, M. Pechenizkiy, & R. S. J. D. Baker (Eds.), Handbook of educational data mining (pp. 245e255). Boca Raton: Taylor & Francis Group

Mohajer, A., Somarin, A., Yaghoobzadeh, M., & Gudakahriz, S. (2016). A Method Based on Data Mining for Detection of Intrusion in Distributed Databases. Journal Of Engineering And Applied Sciences, 11(7), 1493 - 1501.

Mustami, M., Suryadin and Suardi Wekke, I. (2017). Learning Model Combined with Mind Maps and Cooperative Strategies for Junior High School Student. Journal of Engineering and Applied Sciences, 12(7), pp.1681 - 1686.

Paez, H., Zabala, V. and Zamora, R. (2017). Análisis y actualización del programa de la asignatura Automatización Industrial en la formación profesional de ingenieros electrónicos. Educación en Ingeniería, 11(21), pp.39 - 44.

Peng, J., Tan, W., & Liu, G. (2015). Virtual Experiment in Distance Education: Based on 3D Virtual Learning Environment. In 2015 International Conference of Educational Innovation through Technology (EITT) (pp. 81-84). Wuhan.

Pollock, C. & Biles, J. (2016). Discovering the Lived Experience of Students Learning in Immersive Simulation. Clinical Simulation in Nursing, 12(8), 313-319.

Poorani, M., Nithya, P., & Umamaheshwari, B. (2014). A Method for Mining Infrequent Causal Associations with Swarm Intelligence Optimization for Finding Adverse Drug Reaction. International Journal Of Computing, Communications And Networking, 3(1), 25-32.

Romero, C., & Ventura, S. (2012). Data mining in education. Wiley Interdisciplinary Reviews: Data Mining And Knowledge Discovery, 3(1), 12-27.

Tawil, N., Zaharim, A., Shaari, I., Ismail, N. and Embi, M. (2012). The Acceptance of E-Learning in Engineering Mathematics in Enhancing Engineering Education. Journal of Engineering and Applied Sciences, 7(3), pp.279-284.

Udupi, P., Sharma, N., & Jha, S. (2016). Educational Data Mining and Big Data Framework for e-Learning Environment. 5Th International Conference On Reliability, Infocom Technologies And Optimization (ICRITO) (Trends And Future Directions) - IEEE, 258 - 261.

Zamora-Musa, R. and Villa, J. (2013). Estudio de la alternativa de ambientes virtuales colaborativos como herramienta de apoyo a laboratorios tele-operados en ingeniería. WEEF – World Engineering Education Forum.

Zamora, R., Velez, J. and Villa, J. (2016). Contributions of Collaborative and Immersive Environments in Development a Remote Access Laboratory: From Point of View of Effectiveness in Learning. In: F. Mendes Neto, R. de Souza and A. Sandro Gomes, ed., Handbook of Research on 3-D Virtual Environments and Hypermedia for Ubiquitous Learning, 1st ed. Pennsylvania: IGI-Global, pp.1-28.

Zamora-Musa, R., Velez, J., Paez-Logreira, H., Coba, J., Cano-Cano, C. and Martinez, O. (2017). Implementación de un recurso educativo abierto a través del modelo del diseño universal para el aprendizaje teniendo en cuenta evaluación de competencias y las necesidades individuales de los estudiantes. Espacios, 38(5), p.3.