Evaluation and data analysis of user experience for the Doctrina system: a skilled tutor

Danielli Araujo Lima ´[1], LICRO IFTM, danielli@iftm.edu.br, (0000-0003-0324-6690), Guilherme Pereira Santos Silva[1], guilherme.10.silva@hotmail.com, (0000-0002-5702-0366), Gustavo Pereira Santos Silva[1], gustavo7.silva@hotmail.com, (0000-0002-4961-1465), Henrique Yudy Franco[1], henriqueyfranco@gmail.com, (0000-0002-9888-7177), Igor Amalfi da Fonseca[1], igoramalfi23@gmail.com, (0000-0002-7988-9169), Rafael Machado de Oliveira[1], rafaelmo@hotmail.com, (0000-0003-2174-4466)

[1]Instituto Federal do Triangulo Mineiro (IFTM) – Campus Patrocínio
Laboratório de Inteligência Computacional e Robótica (LICRO),
Av. L´ıria Terezinha Lassi Capuano, 255 – 38.740-000 – Patroc´ınio – MG – Brasil

Abstract. Gamification is a growing concept and very popular mainly with young audiences. In the education context its application allows not only extra motivation by the student, but also a capacity for greater abstraction to more complex concepts. The purpose of this work is to present, through a software, developed specifically for mobile devices, a gamified environment for the teaching learning process. The software will work in a personalized way for each student through the intelligent tutor system, based on bio-inspired algorithms, aligned with reinforcement learning practices and hierarchy graph structures. A systematic analysis of system usability was performed through globally known heuristics. Finally, the user responses were analyzed with the objective of discovering patterns for system improvement.
Keywords: Gamification, Nielsen heuristics, intelligent tutoring system, user experience, bio-inspired computing, reinforcement learning.

Avaliação e análise de dados da experiência do usuário para o sistema Doctrina: um tutor inteligente

Resumo. A gamificação é um conceito em crescimento e muito popular principalmente entre o público jovem. No contexto educacional a sua aplicação permite não só uma motivação extra por parte do aluno, mas também uma capacidade de maior abstração de conceitos mais complexos. O objetivo deste trabalho é apresentar, através de um software, desenvolvido especificamente para dispositivos móveis, um ambiente gamificado para o processo de ensino-aprendizagem. O software funcionará de forma personalizada para cada aluno através do sistema tutor inteligente, baseado em algoritmos bioinspirados, alinhados a práticas de aprendizagem por reforço e estruturas de grafos hierárquicos. Uma análise sistemática do sistema de usabilidade do sistema foi realizada por meio de heurísticas conhecidas globalmente. Por fim, as respostas dos usuários foram analisadas com o objetivo de descobrir padrões de melhoria do sistema.
Palavras-chave: Gamificação, heurística de Nielsen, sistema tutor inteligente, experiência do usuário, computação bio-inspirada, aprendizagem por reforço.

1. Introduction

The growth and advancement of technological is an undeniable fact, and can be attributed to the innumerable factors, especially when technology is linked to innovation for providing evolution and revolution around the world. Along this growth came the video-games advances and so-called gamer culture. According to [Ogawa et al. 2016] this culture is defined as a specific type of culture that appropriates elements that are part of the video games universe and all its deployments.

In the educational context, technology has also been inserted in classrooms in an increasingly efficient way. When it comes to efficiency technology for learning, software has become not only pedagogical support tools [Alves and Lima 2018], but also a motivation device for students if well used [Lima et al. 2017b]. In order to understand the concept of gamification for educational systems we have several research such as [Gediel et al. 2016], [Lopes et al. 2015] and [Soflano et al. 2015], proving that games are an efficient way to abstract more complex concepts. The reinforcement is an important element in the learning process. [Ferneda 2006] reports that the student can rather be stimulated through games and challenges organized in difficulty levels.

Analyzing all these concepts and based mainly on the gamification and reinforce ment context for the teaching-learning process, this work aims to: (i) propose a teaching learning application environment, called Doctrina, where (ii) the inserted contents in the system by the teacher must be presented to the student through games using a hierarchical graph. In addition, (iii) an intelligent tutoring system was used to guide student learning through the self-reinforcement plan, modeled using bio-inspired computation algorithms. Finally, (iv) a systematic evaluation will be carried out to verify students' satisfaction when using the system developed here.

2. Theoretical reference

Nowadays people who have devices such as smartphones, tablets or computers have access to a large amount of information. Unfortunately, not all people are able to enjoy the knowledge that the Web can offer. The term M-Learning (Mobile Learning) has different meanings for different communities, that is considered a subset of E-learning (Electronic Learning) and can be understood as educational technologies for distance education (or hybrid education). M-Learning focuses mainly on learning with mobile devices. Unlike U-Learning (Ubiquitous learning), where the focus is learning at anytime, anywhere, independent of the technology used for this purpose [Pedro et al. 2018]. Aligned with this context, access to the Web is growing rapidly, especially in the mobile Internet, where it has seen the highest growth compared to other devices [Virtanen et al. 2018]. This technological advance provided new possibilities and opportunities for the new perspectives in active education methodologies and in the middle of these opportunities arises the concept of gamification that according to [Araujo 2018] consists in using the mechanics, strate- ´ gies and thoughts that are elements of a game, outside context of a game. The purpose of this strategy is to motivate individuals to help them solve some problem in order to teach, and consequently the student will be able to learn new concepts or reinforce existing ones [Silva et al. 2018].

The Doctrina fits into a context of application, because it seeks to promote students learning by inserting them into a competition scenario, with feedback system, rewards system, attempts, mistakes and fun that according to [Squire 2011] are some of the characteristics for gamified applications. But the gamification, as pointed out by [Ogawa et al. 2016], is not only about creating a game that takes a real-world problem and recreates it in a virtual environment, it aims to use strategies to solve a real problem that was addressed in a virtual environment. Contemplating the advancement of mobile applications and smartphones and with the context of gamification spreading in several areas, it has also been an element associated with education, exploring the competitiveness and interaction of students to solve various problems.

Another intelligent tutorial model feature is that it can be applied in a hybrid education context where the contents seen in the classroom can be reinforced by the student. In addition, the Doctrina has an intelligent tutor, called the virtual agent Dominus, who guides the student through a daily planning activities that must be redone, according to the technique of genetic algorithms to optimize the study plan. The Doctrina also has the technique of reinforced learning guided by optimization in colonies of ants, where less visited paths must be reinforced over time for an efficient learning. The domain model stores content that must be reinforced and learned. The student model stores the individual characteristics of each student, such as daily goals and the number of subjects that wish to revise, this model was implemented in this work. The interface model was also implemented and analyzed in this work, through the evaluation by data mining from collected responses of system users.

3. System development

The Doctrina was proposed as a prototype for a gamification learning environment in [Lima et al. 2017a], and in this work a system implementation is presented. To facilitate the Doctrina system construction, we seek to use the tools most used by developers in the methodology Xtreme Programming. For this, we made use of frameworks wich are a set of incomplete classes structures, when extended allow the production of different artifacts for a system. These classes provide a base or skeleton for the tool to be constructed and worked on. In the following subsections we will present the main frameworks that were used for the Doctrina construction - the student, domain and interfaces models.

3.1. Spring Boot

Spring Boot is a project that has made it easier to configure and publish applications. It is a framework that allows an application running in production quickly using the best design practices, and with already optimized configurations. It offers a series of templates, preconfigured tools. It was used to create the Doctrina project, as it has a project creation wizard that greatly simplifies the configuration work. In order to develop projects in the Spring Boot Framework, it is necessary to use the eighth version Java language. Thus, in this work we use the Java language, because it is fast and reliable.

3.2. Ionic Framework

Ionic is an open-source framework for multi-platform mobile application development. To do this, it enables the app implementation using technologies commonly employed in building Front-end web solutions: HTML, CSS and JavaScript. However, in contrast to the framework Apache Cordova, it adopts as a basis bringing features that simplify development even further. These differentials are related to the set of visual components that we can use to construct the front-end solution. Some other points of this platform that deserve prominence is the fact that it is Open Source, has a large community and there are several companies that uses it, so it was chosen for the Doctrina implementation. JavaScript is a high-level programming language becoming more versatile and complete over time. It is currently possible to use JavaScript for mobile applications, so it was used for the M-Learning Doctrina application creation.

3.3. Doctrina Model

The Doctrina interface model of the intelligent tutor can be seen in Figure 1. For the interface creation we aim for a lightweight, fast and functional application. The Doctrina

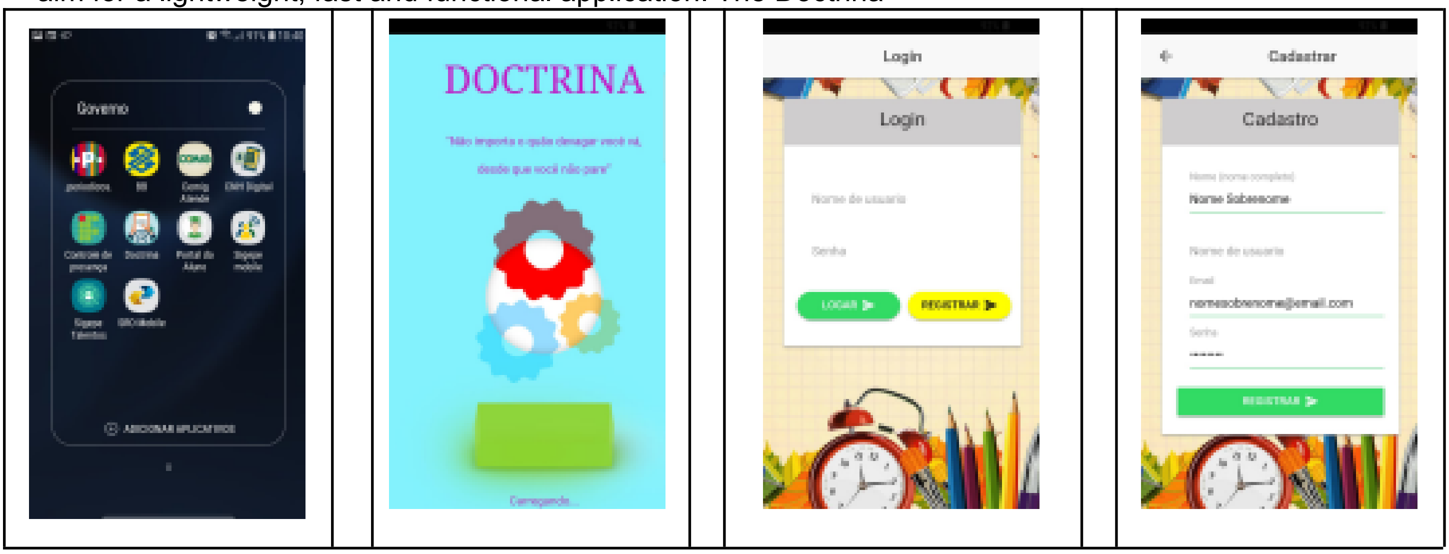


Figure 1. Examples of the Doctrina system interface model, where the login screens are displayed.

developed and its first prototype was presented in [Lima et al. 2017a] as a technological aid tool for education in the context of teaching-learning, where the material content can be understood through gamification reinforcement. The software also presents other main modules: the teacher's domain model, where it can register its disciplines, divide the discipline into modules and add questions and study materials to those modules; to the stu dent's module, which will visualize the subjects registered, can also download the study materials of each module, and answer the questions. Students have access to the ranking where they can foster competition between them; and finally, the tutor system part, that will have an artificial intelligence developed based on the ants' pheromone. This pheromone in the application will

DOI: https://doi.org/10.22456/1679-1916.122071

use as base - the students' average errors and correctness of each module and time in which this activity is not reinforced by users - drawing a customized study plan according to the difficulty of each student. In addition, a tutor who does the student planning through genetic algorithms, as described in [Lima et al. 2017a], through multi-objective optimization.

4. Methodology

To evaluate the Doctrina we will use the user experience through the set of 10 usability heuristics for interface evaluation, defined by [Nielsen 1994]. Usability on mobile devices applications is vital because it places the user's focus on the system. The analysis of heuristics consists in submitting the interface of a certain computational system to the specialists in usability evaluation, according to a predetermined set of "good usability principles". Based on these good practices of usability defined by Nielsen were elaborated two questionnaires for the Doctrina evaluation. The Level 1 questionnaire was applied to the project team, divided into 3 phases: (i) individual interface analysis, (ii) consolidation of the analysis with a group meeting, and (iii) final meeting to prepare a second questionnaire of Level 2. The Level 2 questionnaire presents some questions also according to the principles proposed by Nielsen for users evaluate the system usability system. Being that 74 answers were collected from the second questionnaire, between teachers and students.

5. Results analysis

The results were divided into three main steps that were analyzed according to the questionnaires applied. First, the Nielsen questionnaires analysis were evaluated at both levels. Subsequently, a visual and statistical data analysis collected was investigated. Finally, a data mining analysis with supervised and unsupervised learning was conducted in order to understand the data in a deeper way, trying to discover patterns that could be used to improve the Doctrina system.

5.1. Interface evaluation

As described above, for the Doctrina evaluation we use the Nielsen heuristics. These heuristics were applied in two levels, in which the first level represents the answers from Doctrina developers, and the second level represents the application of a questionnaire with questions created also based on the heuristics of Nielsen to IES (Higher Education Institution) students. The Nielsen heuristics results in the first level can be seen in the Table 1. Based on the Table 1, we conclude that for the first level of heuristics application

Table 1. Nielsen heuristics for the Doctrina application.

Number	Requirement	Grade	Relative
1	Visibility of the current state of the system	39	78%
2	Correspondence between the system/real world	44	88%
3	Freedom of user-friendly control	43	86%
4	Consistency and standards	48	96%
5	Error prevention	30	60%
6	Recognition instead of memorization	43	86%
7	Flexibility and efficiency of use	44	88%
8	Aesthetics and minimalist design	43	86%
9	Helps users recognize and recover from errors	35	70%
10	System help and documentation	35	70%

the lowest score obtained was 60% of the total and the highest score was 96%. The average presented a value close to ≈ 40.4 (80.8%), relatively high variance presented a value of 27.24, that some parameters should be improved. The standard deviation is 5.22 with a standard error of 1.74, which indicates that in general the interface presents good usability according to the developers, since the asymmetry coefficient is −0.60, indicating that the values density is concentrated to the right. According to the evaluation of Level 1, there are still some corrections that we must make in the interface in order to improve the system usability. The heuristics results for the second level can be seen in Table 2. Figure 2 presents the distribution of data collected in relation to the course and

DOI: https://doi.org/10.22456/1679-1916.122071

Table 2. Questionnaire applied in Google Forms for Doctrica users.

#	Questions	Answers	
		Class	
1	What is your main activity in the higher education institution?	0. System Analysis 1. Management 2. Electrical Eng. 3. Accounting	4. Informatics 5. Electronics 6. Business 7. Professor
2	If you are a student, you are enrolled at what level?	0. Undergraduate 1. Technical Education	2. Teacher
	Options	**Yes**	**No**
3	Does the system provides good communication for guidance on the use of same?	91.89%	08.11%
4	Does the system has an easy understanding: machine-human communication?	86.49%	13.51%
5	Does the user have freedom of choice regarding system options?	85.14%	14.86%
6	Does the system use a standard for feedback and user interactions?	58.11%	41.89%
7	As for deletion, editing and insertion does the system has good visibility of these features?	77.03%	22.97%
8	The user has freedom of choice regarding decision making when using the Doctrine system?	79.73%	20.27%
9	Is the software intuitive in navigation of screens?	74.33%	25.67%
10	Is the language used for user interaction easy to understand?	68.92%	31.08%
11	Is the interface used for the software is easy to understand and usability?	67.57%	32.43%
12	Does the system provide an easy recovery if it shows errors?	68.92%	31.08%
13	Does the system provide help to the user while using it?	89.19%	10.81%
14	Provides comments/final considerations.	Open and optional response.	

level of schooling of the users interviewed. The right part of figure represents the first question in Table 2, while the left side represents the second question of Table 2 indicating distribution by level of users. According to the figures, the highest number of respondents are from higher education (≈ 85% of the interviewees), of which more than half (≈ 54%) are Systems Analysis, which is a positive point, since they are students who have prior knowledge of human-machine interfaces.

Based on the answers of Table 2, we can affirm that most of the people interviewed

DOI: https://doi.org/10.22456/1679-1916.122071

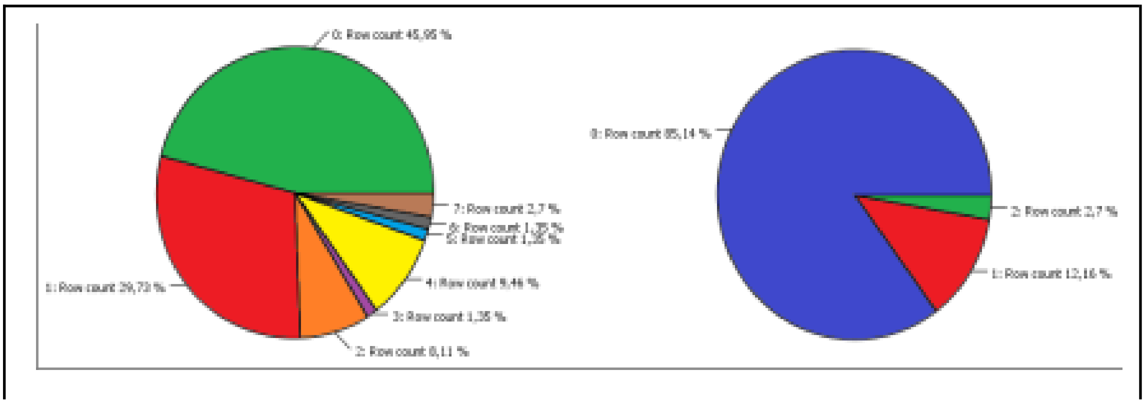


Figure 2. Pie chart that emphasizes the classifications of respondents by course, left and by level of schooling, on the left.

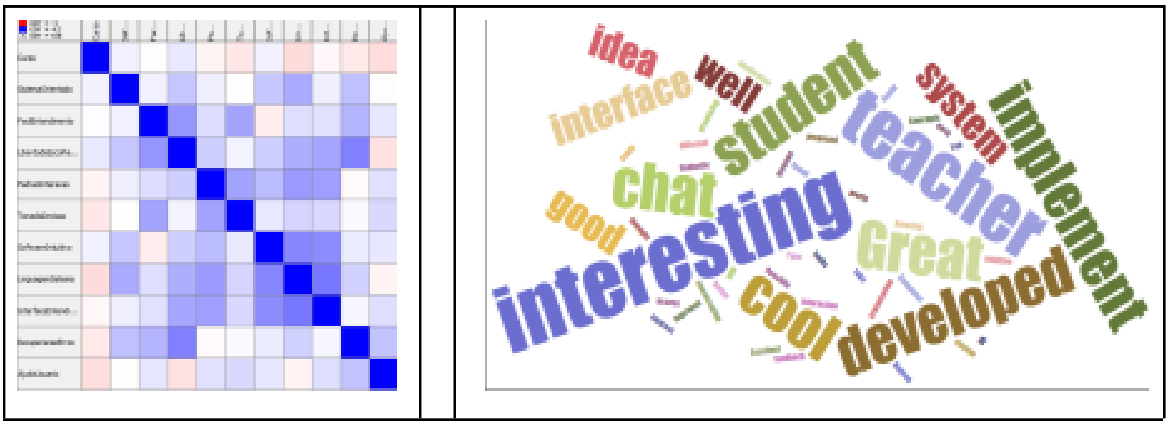
reported positive feedback regarding the Doctrina tool. There have also been many improvements that are very relevant to the application success in future. Figure 3, at left, presents a summary of the 14th question of Table 1, where the occurrences of the main suggestions were extracted in order to assemble a word cloud, where *n* occurrences passed through the transformation function log(*n*). The larger the word size, the greater its occurrence in the responses, for example, word interesting appears 5 times, while the word chat appears only 2 times. The main application characteristics according to the interviewees suggestions is that the Doctrina is very interesting, intuitive and the idea is great. The suggestions raised by the interviewees is that the interface could be improved and a chat system for teacher-student communication could be implemented.

5.2. Statistical data analysis

In order to understand the questionnaire elements in Table 1 some analyzes were made for the Doctrina improvement. Firstly, a correlation matrix was created to verify which questionnaire elements were most similar between each other. According to Figure 3, at right, the elements with the highest correlation are the questions directly referring to the system interface. That is, questions 4, 6, 9, 10 and 11 of Table 1 present a positive correlation level, and all refer to questions more related to the visual aspects interface system.

Parallel Coordinates, a representation of multi-dimensional information or data, in which multiple dimensions are allocated one-to-one to an equal number of parallel axes on-screen. An object in a data set is then mapped as a series of points, one per axis, with the position of each point on the axis being dependent on their value in the associated dimension. The points are then joined together by line segments from one axis to its immediate neighbor, forming a polyline across the set of axes. The selected points are those that presented the highest correlation seen in the matrix of Figure 3 at right. From the analysis of Figure 4 it is possible to notice that some interviewees answered that although the interface is intuitive, it does not present an easy understanding. Students of the technical courses that are not from information technology (IT) area answered that the interface does not present quality in the usability, while students of the undergraduate courses and technicians directed to linked to IT courses, answered that the interface presents good usability. A similar pattern was observed in the question about system communication with the user, where students of non-IT technical courses had difficulty in system usage.

[1]Some comments had grammatical errors, and these were corrected to accurately generate the word cloud.



DOI: https://doi.org/10.22456/1679-1916.122071

Figure 3. At left, cloud of words for each word occurrence in questionnaire suggestions. At right, correlation matrix between the questionnaire elements, ❙ positive correlation, unlike ❙ negative correlation.

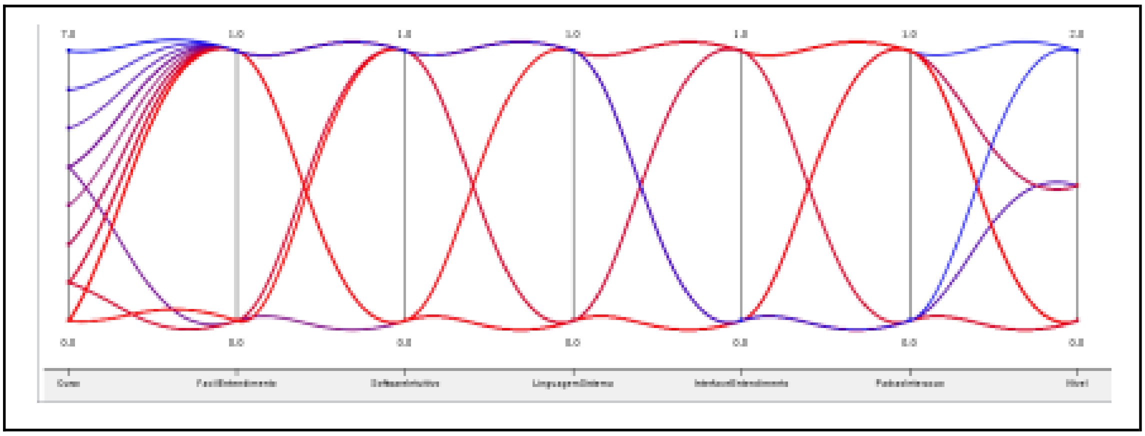


Figure 4. Parallel coordinates between the elements with the highest correlation matrix similarity.

5.3. Evaluation of results by machine learning

Machine learning is the process in which machines through artificial intelligence algorithms and data can learn. The database was divided into learning (70%) and test (30%) and the data were filtered according to the analysis of Figure 3 at right. At first, an unsupervised data analysis data was performed in a hierarchical way using the agglomerative technique to create data groups [Lima et al. 2021]. The Figure 5 shows the dendrogram which displays the whole cluster hierarchy, at the bottom are all data points. The closest data points are connected, where the height connection shows the distance between them. Thus, the coordinate displays the distance fusions and the hierarchy level. The x axis is nominal and displays the single data points with their row ID. The dendrogram shown in Figure 5 was created from 30% of the data since it was necessary to have a perspective clustered data. The distance plot displays the distances between the cluster for each number of clusters. This view can help to determine "good" number of clusters, since there will be sudden jumps in the level of similarity as dissimilar groups are fused. The distances show that three main clusters may be indicative for the input of the optimized non-supervised learning algorithm called k-medoids.

Then, the k-medoids algorithm (with k = 3) was applied with using the learning base, with 70% of total data, for visualizing the data and understanding how the algorithm behaves. Figure 6 shows data in two-dimensional and three-dimensional forms. Three

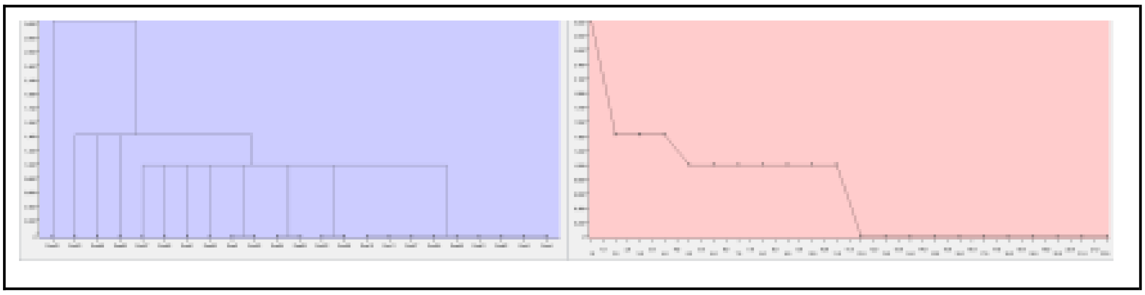


Figure 5. Dendrogram created by hierarchical agglomeration and distance graphic.

centroids were selected by the algorithm, the centroids are the lines {9, 24, 53}. The data were plotted from a reduction of their dimensionality from a high precision rendering. A total of 4 iterations were needed to group the data. The respective cost reductions were observed at each iteration {12.08, 6.47, 0.60, 0.00}. It usually starts with some large cost reduction (due to random initialization) but decreases as more iterations are run.

A supervised learning system is a program (inductor) capable of inducing a

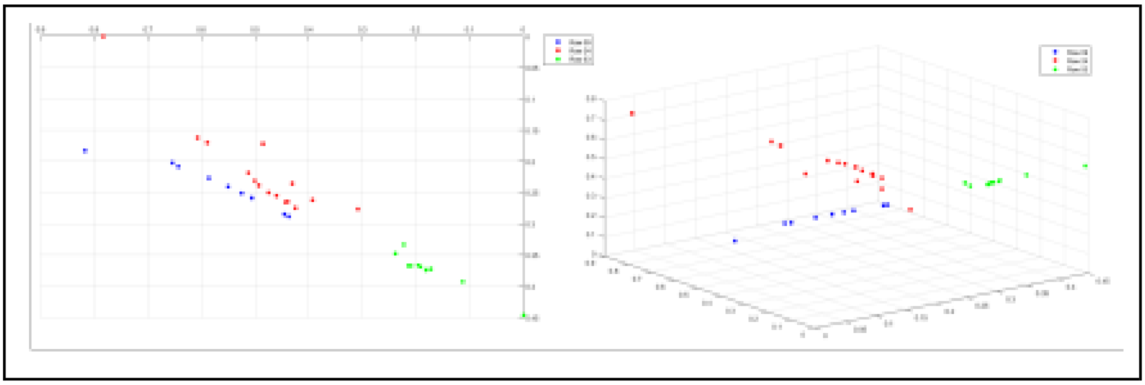


Figure 6. Two-dimensional and three-dimensional data visualization data grou ped by k-medoids.

description of a concept (classifier) from a set of known examples and previously labeled with their respective classes. From the previous clustering, a learning phase was performed with 70% of the data and 30% of the data was used for testing. From this, the classification algorithm called decision tree was used to verify a prediction rule that best classifies the data. Figure 7 shows the learning tree and the prediction tree returned by algorithm Decision Tree C4.5.

Thus, it is possible to verify a rule that best classifies the data from the filter returned by Figure 3 at right. So, it is possible that if users responded that the system language is adequate, it means that we can infer that the user is a teacher, and we still have 73% trust that that user is at the top level and only 37% chance that the individual is higher education. This means that the more experienced and older individuals showed better behavior during the Doctrina usage. On the other hand, the high school/technical students presented greater difficulty to use the system proposed in this work, which indicates that the interface can still be improved, since our intention is that the Doctrina system can be used by all the institution's students and teachers.

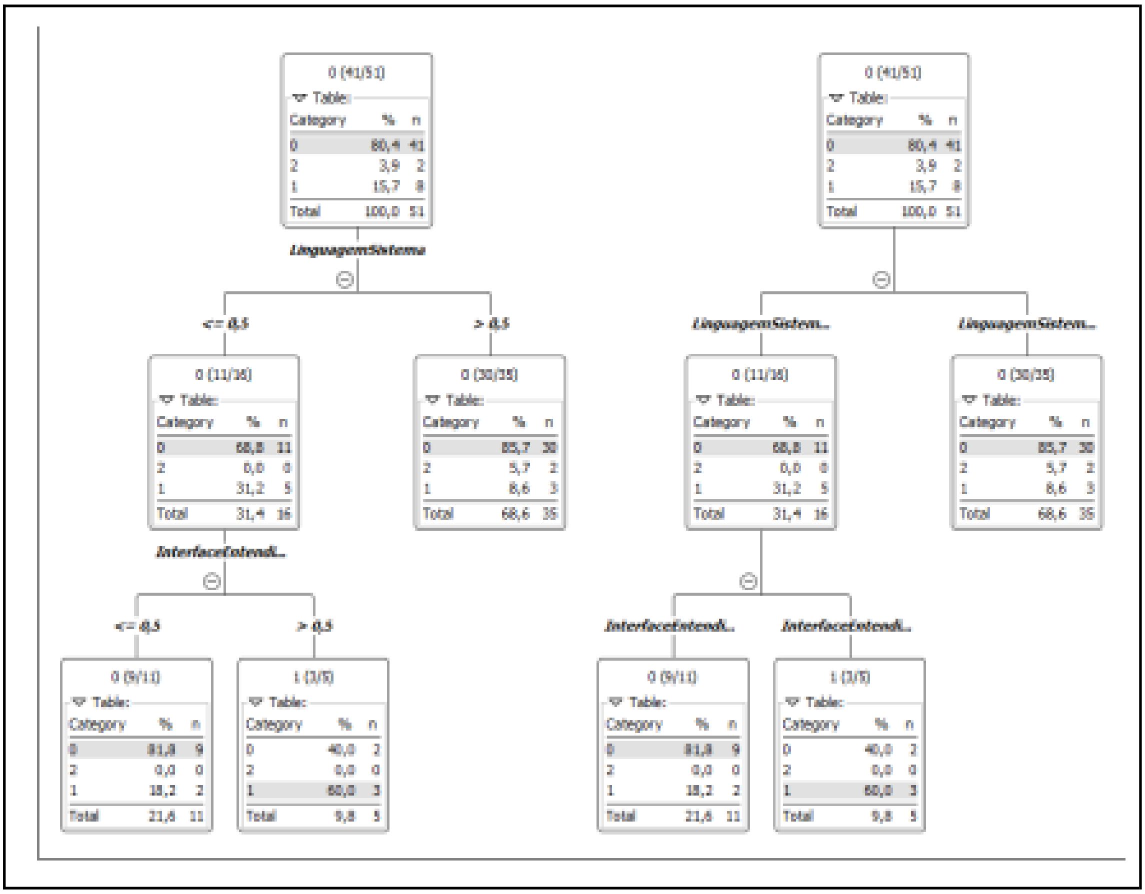


Figure 7. Decision learning tree for the prediction used for analysis of question naire answers.

DOI: https://doi.org/10.22456/1679-1916.122071

6. Conclusions and future work

Based on the data obtained through research using Nielsen heuristics questionnaires, we concluded that the Doctrina functionalities was implemented in accordance with what was idealized in [Lima et al. 2017a]. Analyzing the results, the software will undergo some changes and improvements in future versions. After data mining, it was possible verify that the high school students had greater difficulty in using the application, because of being a first contact with these technologies for educational support. In this way, a training should be given to these students with difficulty in using technologies. Through the feedbacks the authors intend to adjust, mainly in the interface, seeking greater standardization between the screens and the improvement of some application parts. For example, chat creation, Doctrina and ranking to Facebook API integration, feasibility of a student's digital signature on answers completion. The Doctrina application was awarded as the best idea (1st place) of the IFTM VII New Products, Processes and Services Fair (FNPPS) in 2018[2].

Referências

Alves, F. B. and Lima, D. A. (2018). Uso de la clasificación para el análisis y la minería de datos en la herramienta de enseñanza-aprendizaje google classroom. In XXIII *TISE*

Conferencia Internacional sobre Informática na Educação .

Araujo, M. S. (2018). Teaching-learning and digital technologies in pre-service english language teacher education. *Trabalhos em Linguıstica Aplicada*, 57(3):1590–1614.

Ferneda, E. (2006). Redes neurais e sua aplicação ̧ao em sistemas de recuperação de informação. Ciência *da Informação*, 35(1).

Gediel, A. L. B., Soares, C. P., and de Oliveira, C. L. R. (2016). O ambiente virtual como aliado no processo de ensino e aprendizagem da libras. *Revista (Con) textos Lingu ́ısticos*, 10(16):24–37.

Lima, D. A., Ferreira, M. E. A., and Silva, A. F. F. (2021). Machine learning and data visualization to evaluate a robotics and programming project targeted for women. *Journal of Intelligent & Robotic Systems*, 103(1):1–20.

Lima, D. A., Oliveira, C. C., Pestili, L. C., Silva, E. C., Bezerra, M. A., and Lima, H. A. (2017a). Uma proposta de sistema de aprendizagem com conteudo gamificado e com reforço guiado por algoritmos bio-inspirados. *Computer on the Beach*, page 140.

Lima, D. A., Zati, A. F., and Silva, E. C. (2017b). Análise de dados no google classroom para auxiliar na diminuição do distanciamento transacional nas disciplinas da área de informática. In XXII *TISE Conferência Internacional sobre Informática na Educação* .

Lopes, R. A., Toda, A. M., and Brancher, J. D. (2015). Um estudo preliminar sobre elementos extrínsecos e intrínsecos do processo de gamification. *Revista Brasileira de Informatica na Educacao ̃* , 23(3).

Nielsen, J. (1994). Enhancing the explanatory power of usability heuristics. In *Proc. of the SIGCHI Conf. Human Factors in Computing Systems*, pages 152–158. ACM.

Ogawa, A., Klock, A. C. T., and Gasparini, I. (2016). Avaliaçao da gamificação ao na área educacional: um mapeamento sistemático. In *Brazilian Symposium on Computers in Education-SBIE)*, volume 27, page 440.

Pedro, L. F. M. G., de Oliveira Barbosa, C. M. M., and das Neves Santos, C. M. (2018). A critical review of mobile learning integration in formal educational contexts. *International Journal of Educational Technology in Higher Education*, 15(1):10.

Silva, R. R., Fernandes, J., and Santos, R. (2018). Panorama da utilização de jogos digitais no ensino de programação no nível superior na última década: Uma revisão sistemática da literatura. In *29º Brazilian Symposium on Computers in Education*, page 535.

[2]DAL is grateful for the financial support granted by CAPES, FAPEMIG and CNPq and other authors are grateful for PROPI-IFTM for the execution of research project.

Soflano, M., Connolly, T. M., and Hainey, T. (2015). An application of adaptive games based learning based on learning style to teach sql. Computers & Education, 86:192.

Squire, K. (2011). Video games and learning. *Teaching and participatory culture in the digital age. New York, NY: Teachers College Print*.

Virtanen, M. A., Haavisto, E., Liikanen, E., and Kääriäinen, M. (2018). Ubiquitous learning environments in higher education: A scoping literature review. *Education and Information Technologies*, 23(2):985–998.