A Serious Game as a Tool for Teaching Outlier and Fraud Detection: A Case Study

Jean Avila Rangel, Maria Claudia Figueiredo Pereira Emer, Adolfo Gustavo Serra Seca Neto
Academic Department of Informatics
Federal University of Technology - Paraná
Curitiba, Brazil
jeanrangel@alunos.utfpr.edu.br, mcemer@utfpr.edu.br, adolfo@utfpr.edu.br

***Abstract*—Outlier detection is an important subarea of data mining that can be used for fraud detection but, as far as we know, there are no games developed to teach this subject. To address this gap, in this paper we present Encontre a Fraude, a serious game for teaching outlier detection that uses data visualization techniques and three-dimensional graphics. Our goal is to evaluate if players learn outlier detection concepts when playing a serious game intended to teach this subject. We evaluated the game in a case study with three different groups of undergraduate students and public auditing professionals. After playing the game, each player of each group answered a questionnaire. In the first two evaluations, the players answered that playing the game was indifferent to learning oulier detection (neutral answers in 2 of 3 questions). After these two evaluations, the comments and considerations of the players were used to implement an improved version of the game. At the last evaluation, the game got positive responses (neutral answers in 3 of 3 questions) as to its effectiveness in teaching outlier detection. Our results are preliminary but they suggest that a serious game can be used as a tool for teaching outlier detection.**

***Keywords*-serious games education; data mining teaching; outlier detection; fraud detection;**

I. Introduction

The outlier detection subarea is established in the data mining field, where computing is used to generate information from data collected, in most part, with automated procedures to create clusters, classifiers, neural networks, etc [1]. A small amount of data does may be meaningless but, in administrative and business fields, large amounts of data can be relevant to the success of companies.

Outlier detection is derived from statistics and lately it is present in computing studies, usually associated with data mining [1]. A fraud is a kind of outlier. The definition of fraud is: the misuse made by one or a group of people from a sector to increase their benefits. This fraud can be done in an intentional or unintentional way [2], [3]. The same definition is also applied in public sectors of government, where people can commit illicit acts to gain advantage [4].

Whatever the context, outlier detection is an area that will probably never find a definitive solution to every problem. With this challenge, the cost of finding a fraud has to be less than the fraud itself.

The literature about data mining dedicated to outliers and fraud detection presents supervised and unsupervised techniques. Some authors show the possibility of semi-supervised techniques. In the case of semi-supervised, the technique will contain some characteristics of a supervised and unsupervised techniques.

In data mining, visualization of data facilitates for a human to understand the information of files and presents them in a more intelligible manner in comparison to plain text and tables [5]. Data visualization is one of the most indicated tools for outlier detection in non-supervised contexts [5] because, in this environment, the algorithms that need some attribute of classification will not have a good result [6].

There are few studies in literature about works describing data visualization applied to outlier and fraud detection [5]. In digital games, object of study of this work, the studies are related to how data visualization can be used to observe the route followed by the player in a map. This observation is made to analyze their behaviour for commercial purposes (e.g. to offer products in games that the player probably will buy) [7].

Serious games can teach disciplines of interest to the player. Game characteristics can facilitate the knowledge absorption [8], [9]. Teaching computing disciplines to students can also be done with serious games [10]. Therefore, it is possible to teach outlier and fraud detection through serious games to students.

Our main goal in this work is to evaluate if users can learn outlier and fraud detection, sub-fields of data mining, after playing a serious game.

We also want to understand and evaluate the learning perception of undergraduate students and public sector auditors about the data mining theme through the game and, according to their opinions, whether the game motivates them to learn about data mining. Finally, we try to understand whether the game encourages the study of data mining and the development of citizenship.

This paper is organized as follows: in this part, the Introduction, we discussed our goals and the motivational factors to conduct this study. In the next section, Background, we discuss data mining and serious games. In the third section, Method, we explain our method of study, case study, and

explain it. In the fourth section, Analysis and Results, the results and discussions are presented and commented. Also, in this section, our serious game developed is presented and discussed. Finally, at the Conclusion, we talk about lessons learned and future works.

II. Background

In this section we describe concepts of data mining and serious games, showing some works related with this study.

A. Data Mining

Mostly data mining algorithms work with data similar to a database table, which every instance is represented by a row and the attributes are indicated in columns. In supervised data, a column (generally, the last one) have one classifier information, and it can be a variable binary or nominal.

With focus on outlier detection, Ahmed et al. [11] worked only with unsupervised data, which are data without a classifier attribute, because the majority of data in the world do not have a classifier attribute. The work indicate that techniques of clusterization (and the most popular algorithm of the area, *K-Means*) are common to detect anomalies in this kind of data.

In contrast with unsupervised data, supervised data are information that researchers have control of which elements are frauds. An example of supervised data is a table with information of spends on a credit card, which a bank specify previously what data is fraudulent [12], [13]. To supervised data, algorithms are trained using machine learning to classify new values unlabeled.

Therefore, meanwhile cluster methods (unsupervised) try to identify patterns on data because they do not have a label, supervised algorithms train a classifier knowing which instance is true or false. The problem of supervised data is the difficulty to find new anomalies non similar to previous ones [11].

Focusing in unsupervised methods, Ahmed et al. [11] identified in their literature review on anomaly detection the three most common types. The groups are similar to the groups presented by Bansal et al. [14] and indicate: punctual anomaly, where the element is located far from the others; contextual anomaly, where the element is located far from the others due a contextual circumstance; and collective anomaly, where a group of elements are located far from the rest.

Finally, a task in cluster algorithms (unsupervised methods) is to indicate previously the number of groups that data has. On *K-means*, this number is indicated (apriori) with the letter K. Therefore, the computational complexity in this technique is low, occurring in order O(n), where N represent the number of elements in the data set [11].

B. Related Works in Serious Games

As basic premise, serious games must teach new and useful knowledge to players [15]. In this way, game's topics should be discussed in an educational manner. A video game that do not just have entertainment can be considered a serious game [10]. A serious game must give to player information source in educational areas, health or public policy.

All works found that relate serious game with data mining are aimed to analyze actions performed by the player in gameplay, for example, observe their punctuation and routes by the maps. Educational Data Mining (EDM) is an area that analyze the game results by computational algorithms [16]. Also were found works in Learning Analytics (LA), an area derived from statistic destined to visualize, in an analytical way, the informations of a student in their learning period [17].

The comparison between EDM and LA is frequent and a study tried to identify the relation in those areas. The main difference is the human and computational approach. Meanwhile EDM utilizes automated computers to analyze and show results to users, LA aim to encompass specialized humans to verify learning data generated by students [17].

The vast majority of works realized analysis on player scores to evaluate the learning perception of students. In commercial games, the analysis of results have the objective to offer to players services, products or a differential treatment. In these games is also possible identify players profiles to give them special characteristics [18].

Searching by serious game with corruption combat on Google Play[1] we have found two free games with the goal to teach basics aspects about politics to the player. The *Juega Limpio*[2] was developed by the Ministerio de Transparencia Institucional y Lucha Contra la Corrupción from Bolivia and the *Certified Fraud Examiners 2017*[3] was developed by the homonymous organ. The games presented an educational purpose and gives to player activities to discover and learn about laws and legislation.

III. Method

The method used for the development of this work is presented in Figure 1. We can observe the following steps in this figure: 1) Perform research in the literature on data mining, fraud detection and outlier detection, serious games and application of data mining in games. 2) Development of the first version of the serious game to be played and receive the first criticism. 3) Adaptation of the MEEGA+ questionnaire to include research questions that reflect the serious game developed. 4) Elaboration of the first case study, with undergraduate students in computing. 5) Analysis

[1] https://play.google.com/store/
[2] https://play.google.com/store/apps/details?id=com.tic.seven/
[3] https://play.google.com/store/apps/details?id=com.studyguidecorp.cfe

of the answers provided by the students. 6) Implementation of the second version of the serious game, considering the answers and criticisms obtained in the first questionnaire. 7) Second application of the serious game and questionnaire in public auditors. 8) Analyzing the answers provided in the last questionnaire, two experts also made criticisms and suggestions for the game. The opinions of the specialists were not considered for this study and are commented in the section of future works, at the end of this work. 9) Relevant comments were considered for the implementation of the last version of the game for this case study. 10) The new version was tested by undergraduate students who had knowledge in data mining. 11) Finally, the results of the last questionnaires answered in the previous stage were analyzed and this case study work was finalized.

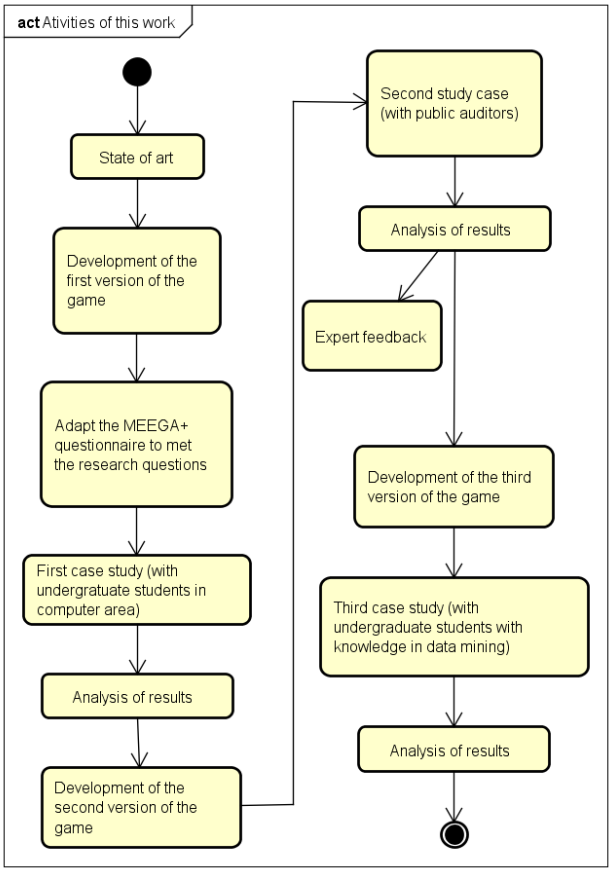

Figure 1. Activities realized to conclude this work.

A. Questionnaire

In the context of our work, an area that is without a boundary definition is serious game as a form of entertainment and, at the same time, as a tool to aid in the teaching of outlier detection. To guide the development, we used a questionnaire to conduct a case study. In addition to the interpretation that the questionnaire can provide, the case study accepts subjective factors that can be obtained by observing the study environment.

In the literature, there are two tools for evaluating assistance of serious games or educational games. The first tool developed was EGameFlow [19]. Afterwards, the Model for Evaluating Educational Games (MEEGA) was created following the technique Goal/Question/Metric (GQM) [20]. Both tools were systematically developed and tested following a case study. The two tools are concerned with evaluating issues such as user learning and user experience. However, the MEEGA model is most commonly used in studies by most authors [21] and [20].

In order to improve the MEEGA model, MEEGA+ was developed after several tests and analyzes in studies that used and reported the MEEGA model. It was verified that the MEEGA tool generated some errors of comprehension in the evaluated users and its reliability and validity had limitations [22]. We can say, then, that the MEEGA+ model is an updated and improved version of the MEEGA model.

The new MEEGA+ model is an evolution from the previous one (MEEGA) and has been validated by a group of specialists in the areas of statistics and computation. In the new model, issues such as usability, confidence, challenge, satisfaction, social interaction, fun, focused attention and relevance are asked for users to evaluate after playing the proposed game [22].

For our work the MEEGA+ questionnaire was more suitable, since the authors provide a version in Brazilian Portuguese, the official language of the country where the case study was applied (Brazil). Another decisive factor for the choice of the MEEGA+ questionnaire was the higher frequency and reports of tests present in the literature. In most tests, the tool proved to be efficient. A characteristic found in the most part of the studies was the lack of explanation of how to proceed and analyze the data collection of the questionnaires. Tool-related articles, produced by the authors themselves, focused on reporting how the tool was constructed and evaluated, but not how to evaluate the results provided by them. To better represent our context, some questions of the MEEGA+ questionnaires were adapted to improve the volunteers' understanding of the case study.

B. Tools for the serious game development

We've developed a mobile app designed to teach outlier detection concepts through a serious game. The game was developed using an game engine because it facilitates the coding of three-dimensional objects and elements of compatibility and gameplay of devices. The tool used is the *Unreal Engine*, provided without costs by the company *Epic Games*. The coding was done using the *visual scripting* tool called *blueprint*, which allows the creation of codes in the C++ language using visual elements with nodes and edges.

C. Case study

After the completion of the first version of the serious game Encontre a Fraude, 12 undergraduate students of the course of Computer Engineering and Information Systems of the The Federal University of Technology - Paraná (UTFPR), campus Curitiba, played and opined about the game. The volunteer group was part of the Software Engineering course. In order to carry out the case study, we try to influence the environment and the case studies as little as possible.

In the second case study, five public servants related to auditing in the state of Paraná (Tribunal de Contas do Estado do Paraná) interacted with the game and also provided their considerations.

In the third and last case study carried out by this work, we applied the game to 30 undergraduate students of the Information Management course of the Federal University of Paraná (UFPR) of the Product and Information Service discipline. The choice of class in question was due to the students' likely contact with the data mining discipline. The students were in the seventh semester of the course and probably had contact with the discipline in the previous semester. This desired profile characteristic was due to the fact that we obtained neutral answers in questions related to the objectives of our work. When interviewing students who had recent interest in data mining, we expected to have a better understanding of the subject and match the interest of the game.

After completing the game coding, the MEEGA + questionnaire (adapted for this study) was applied to the three different groups that played the game and afterwards, answered the questionnaire.

IV. The Serious Game

To better exemplify and design the software, we developed diagrams to guide the project. In Figure 2, we observe the use case of the proposed game and its main functionalities, which are: a player can enter the game, exit the game, learn to play and play with synthetic data (data created by us for a fictional context).

We can visualize in the diagram of deployment, represented in Figure 3, that the player will use his device (preferably a smartphone) to access the online store to get his game for free through an internet connection. The entire application of the game will remain installed on the user's device, requiring the internet connection only for updates or obtaining the game for the first time.

A. Version of the game for the case study

The production of the software generated a portable application for the *Android* platforms, in the *Android Package* (APK) and *Microsoft Windows* format (in the executable and HTML5 web version format). Controls and interactions were predominantly designed for mobile devices running the *Android* operating system. However, as we will see in the following sections, most users have played the game on desktop computers with the *Microsoft Windows* operating system.

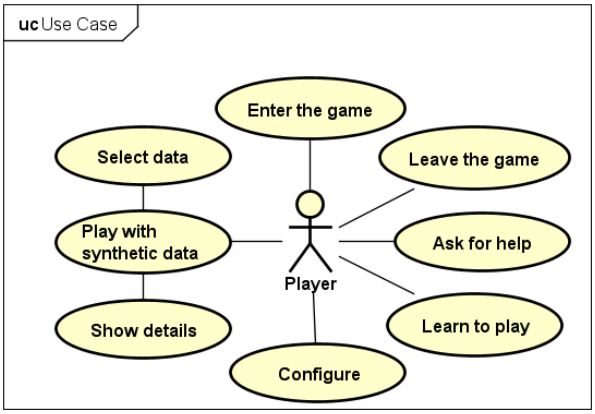

Figure 2. Use case diagram of the serious game.

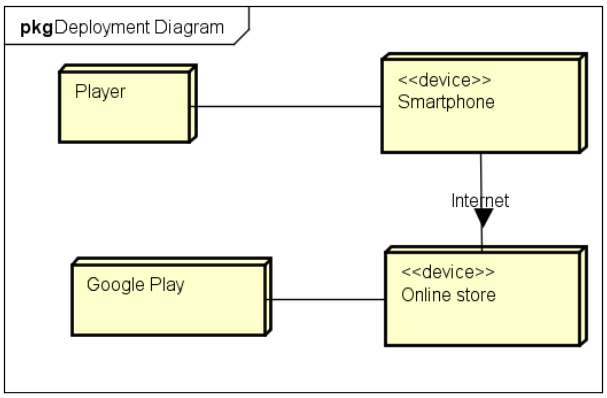

Figure 3. Deployment diagram of the serious game.

When the game starts, the player receive three-dimensional maps that represent graphs with information and values. In the example shown in Figure 4, we observe a more separate element from the others. In the case of this step, this element means an outlier and must be selected, as shown in Figure 5. When proceeding to the results, by choosing the option "Next", the player gets the interface with his score and the possibility of advance to next level, which can be visualized in Figure 6. If it obtains the minimum score to advance the step, a message is displayed on the screen providing some formal explanation of the content treated by the step, as shown in Figure 7.

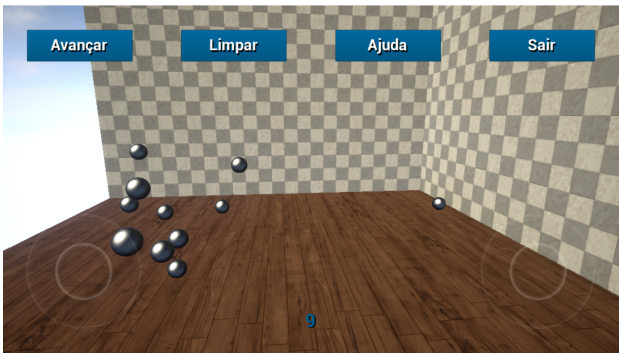

Figure 4. User interface of a game step.

Figure 5. User interface of a game step after select an item.

Figure 6. User interface to present the score of each step.

V. Analysis and Results

A. First context of the case study: undergraduate students in computing

Analyzing the demographic information of the students, we find that the majority is male (92%), aged 18-28 years (92%). In addition, it has been identified that most students play digital games at least once a week (67%). With this information, we can infer that the case study group possessed mastery and knowledge about digital game controls and movement in two-dimensional and three-dimensional environments. It has also been found that most (83%) of students

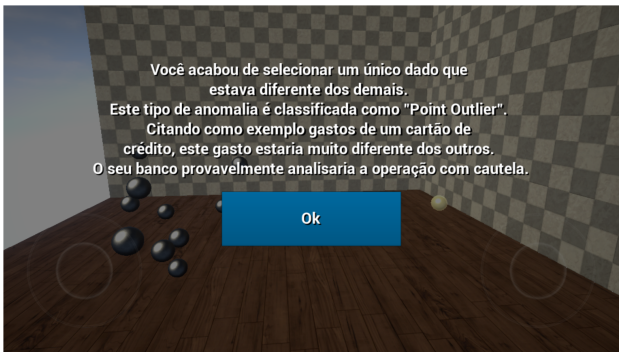

Figure 7. User interface with the explanation of the content discussed in the step.

rarely play non-digital games (cards or board games, etc.).

Analyzing the results, we noticed that some items did not obtain any positive response. However, this result was expected in some cases, such as questions regarding interaction between participants. This is due to the fact that the students were not instructed to communicate and interact with each other during the execution of the game and the questionnaire, in addition to being a singleplayer style game. In Figure 8 we can see three questions related with perceived learning and the answers show neutral opinion in two of them.

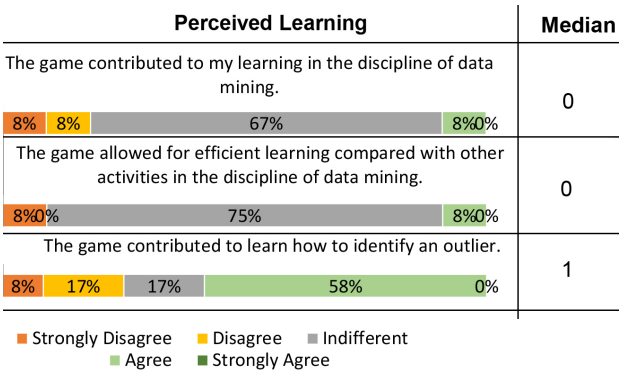

Figure 8. Perceived learning of our first case study.

Another question that did not receive a positive answer was the question to evaluate if the game made possible the personalization of visual appearance and texts. As the purpose of the work is not to study or propose interfaces with the user, modifying the visual or elements of the game is not in the scope.

Focusing on questions related to the object of study of this work, we identify different results than expected. The MEEGA+ questionnaire is indicated and intended to be applied to students who are taking the same course that the game proposes to tackle. In our case study, the game was applied to students of the discipline of Software Engineering, not the discipline of Data Mining. Therefore, questions about

the relationship of the game were adapted to: the relationship of the game with the discipline of data mining. This may have generated some confusion for the students. By this factor, the median of students' response to questions related to the discipline was as "indifferent".

Most of the criticisms that users have suggested was focused on the controls of the game. As some went awry with dizziness problems and others had trouble with the move around the scene, more solutions were thought out and implemented for better user experience playability. This negative factor was already being predicted, since the controls had not been pre-tested on the devices used for the first case study version of the game.

B. Second context of the case study: public auditors

Implementing some of the suggestions that the previous case study had suggested, a new version of the serious game Encontre a Fraude was developed to begin a new case study stage with public auditors. Most of the previous comments was related to the game's design and gameplay, as students encountered difficulty navigating the graphics and found the game to be very repetitive. Considering the comments, the new version of the game got full re-adaption in the controls for movement and the colors of some scenarios were changed. The values present in the data remained the same.

At the time of the case study, not all participants played the game entirely because the game was presented on a projector to everyone in the room simultaneously. The participants also talked to each other about their views on the game. In the end, each one responded individually to a questionnaire.

The demographic characteristics of the case study group were similar between the participants and the sample was small (five people), compared to the other case study (12 and 30 people). We found that all individuals in the case study were male and the vast majority were over 40 years of age. When questioned about how often they play digital games, respondents indicated different values. One respondent reported playing digital games daily. Another respondent commented on playing at least once a week and one participant said he never plays digital games. Two respondents said they play analog (card or board) games and none reported ever playing. When asked about the frequency of playing analog (non-digital) games, the participants replied that they play rarely (60%) or monthly (40%).

Analyzing the answers in comparison with the results obtained in the first case study, we obtained a similar result in questions related to the discipline of data mining, since the median response of the public auditors was "indifferent". Other values have also been shown to be the same as the previous case study, such as the player experience with design aspects.

In the area of the questionnaire designed to assess the player's perception of learning after interacting with the game, we again got the median "indifferent" in most questions. These specific questions are considered important for the case study of this paper as they are strongly related to our general and specific objectives. We can see the percentage of the answers in Figure 9 and we receive the median 0 in Likert scale in two of three questions.

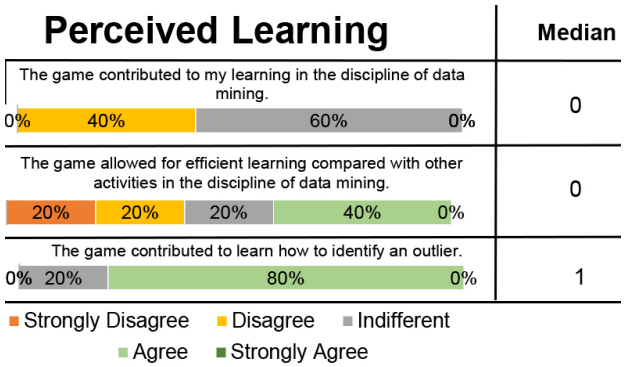

Figure 9. Perceived learning of our second case study.

C. Third context of the case study: undergraduate students in Information Management who had contact with data mining

Considering the comments of the last case study of the game, we added explanations about the contents covered in each stage. The purpose was to provide a more detailed explanation to the user about the concepts dealt with in the game and to better explain the context in which the themes are presented and worked out.

In the case study conducted, 59% of the class was male and 41% female. Most interviewees were between 18 and 28 years old. Also, 14% were between 29 and 39 years old and one respondent was over 50 years old.

In this group, 34% of respondents play digital games rarely, 24% weekly, 21% daily, 14% never play digital games and 7% play at least once per month. Furthermore, 64% of respondents play analog games rarely, one person plays weekly, 7% daily, 7% never plays digital games, and 21% plays at least once a month.

As one of the requirements for class selection for questionnaire preparation, familiarity with the subject of data mining was considered relevant for this stage of the case study. Therefore, we included in the questionnaire a question regarding the student's familiarity with the data mining discipline. In the new section of the questionnaire, we obtained the answers from which 19 participants studied and were approved in the discipline. Four participants were disapproved and six did not study it, but they had knowledge about what is covered and the content covered. It is important to emphasize that the inconsistency of the number of respondents is correct. There was one respondent who did

not complete all the answers, but his questionnaire was still considered.

In the part of the questionnaire regarding free text comments, we obtained returns from some students. Briefly, we can see students' satisfaction with the game's theme, as many have reported finding the proposal interesting to teach some course content through a serious game. One of the students wrote in the strengths section the word "free" because the game is made available at no cost. We consider this an important factor, since the goal of the game is to promote the social good, and the attribution of costs would be an impediment to this. The strengths listed by the participants are compiled as follows:

- The game have examples for the content;
- Three-dimensional map;
- Present the data in charts help to understand the information;
- Free;
- The goal of the game is to teach the player something;
- The game is related with the data mining discipline.

Although some of the strengths indicated by the participants were related to the game's visual and design, some participants criticized these aspects. Some participants suggested changes in the game interface. For this work, it was not objective or purpose to study the visual elements. The design was planed only to display information on the screen in a way that the player could interact with the steps. Some comments from participants regarding improvements to the game are as follows:

- Improve the design;
- More steps and examples;
- Improve the movement and controllers;
- Create a ranking between players;
- The game must be violent to be funny;
- Improve the language of the game, with short, simple and patronized texts;
- Great teaching method;
- Creative and fun.

In additional comments, we received a great deal of praise on the subject of work, the quality of the software presented and the positive perspective of using digital games to teach undergraduate subjects.

When we apply the case study questionnaire about the serious game Encontre a Fraude in undergraduate students who have been in contact with the discipline of data mining, we have received a positive feedback regarding the usefulness of the game for teaching. Many comments were on the visual of the game, which was not the object of direct study of this work.

The answers of the important questions to this work can be seen in Figure 10. In contrast with the first and second case study, in our last case study we found positive answers in questions related with the effectiveness of the game in teach data mining and outlier detection. In three questions, we can see the median 1 (using Likert scale) in all of them.

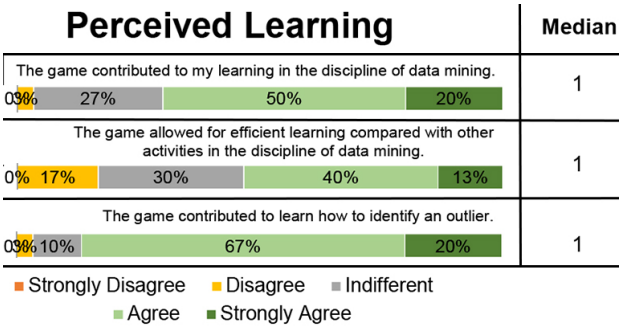

Figure 10. Perceived learning of our third case study.

VI. Conclusions

In the present work we carry out a case study to verify if it is possible to exercise concepts of outlier detection with a serious game. In order to evaluate the game and our hypothesis, a questionnaire for the evaluation of serious games was applied in the case study.

The first version of the game was tested by 12 undergraduates studying Software Engineering. In answers, some students said they could not play the game because the controls were confusing, causing dizziness. Visual and design factors were criticized in all case studies, however, it was not the purpose of this work to develop an ideal visual identity for this type of game. Another feature discovered in the first case study was that participants considered their opinion about the game for teaching data mining, the main purpose of the work.

After fixing problems with controls, moving and cosmetically alter some visual aspects of the game, a new version was tested by five public servants of the audit area in the state of Paraná. Participants indicated that the game did not provide the proper contextualization of what content was being addressed in the game, how the subjects is related to the phases and what users were learning. Overall, the problems pointed out by the auditors were the same as those reported by the students in the previous case study. The results obtained in the questionnaire were also similar, including the neutrality in the questions regarding to evaluate if the game serves the purpose of teaching data mining concepts.

In the third version of the game, produced to run the third case study, 30 students of the Product and Information Service discipline volunteered to play it. The undergraduate students were invited because most had already had contact with the subject of data mining. In the questionnaire presented to them, the term data mining was replaced by outlier detection, because it better portrays the content addressed by the game. As results, we obtained differences in comparison

to previous questionnaires. The students reported that the game is efficient in teaching outlier detection.

The questions that presented similar values to the previous questionnaires were related to the incentive of individual or collective citizenship. As a specific objective, the work sought to identify if the serious game encourages the personal and collective citizenship of the users. All the questionnaires presented answers that abstained from positive or negative comments regarding this topic.

Therefore, according to participants, the game is not the best tool if the goal is to encourage personal or collective citizenship, however, it may be a good choice for teaching outliers. In the first case study, students who had never had contact with the discipline of data mining did not consider gaming an efficient alternative for teaching this content. The students in the last case study, who had had contact with the discipline, reported that the game is a good way to teach outliers detection. By the contrast of the answers, we have two possibilities: (1) - the students of the first case study remained indifferent to this question because they did not know what is data mining; (2) - the game serves as a complement to the discipline of data mining, and not to introduce the concept from the beginning.

For this work, we consider the two valid possibilities, but, as a conclusion, we emphasize the second one: the serious game developed serves as a complement to the teaching of the data mining discipline, more specifically, contents related to outlier detection. However, it should be noted that some users felt physically ill due to dizziness. With this factor, we conclude that the application of a serious game with the same characteristics of the game developed by this work (three-dimensional and first-person) should not be the only tool for all those interested in learning some discipline.

VII. Contributions

The software produced by this work consists of a serious game in three-dimensional, singleplayer and first-person style called Encontre a Fraude. The three versions developed for each case study presented by this work can be individually get in[4]. The development source code is considered free software and can be obtained, modified and distributed freely by the online repository[5]. The latest version of the game, developed for mobile devices with the *Android* operating system, can be obtained for free through Google Play[6]. The entire development was made in the Unreal Engine[7] game engine, distributed free of charge by the company Epic Games[8].

[4]https://goo.gl/dgwjm8
[5]https://gitlab.com/jeanrangel/jogo-encontre-a-fraude
[6]https://play.google.com/store/apps/details?id=com.rangel.encontreafraude
[7]https://www.unrealengine.com
[8]https://www.epicgames.com

VIII. Future Works

As future work, we can approach design and visual characteristics, subjects always discussed among users when answering the questionnaires. Techniques of how to display the data in the graphical user interface and how to organize the texts and contents of the game are still little explored in the literature.

A factor to be considered in future works was the median value of zero in the question to evaluate if the game is attractive (interface, graphics, etc). This is a factor that we consider important for the acceptance of the game by a greater part of users, but was not object of study in this work.

As some users reported problems with dizziness, new ways and mechanics to manipulate the game may also be considered objects of study for future studies. In addition, most of the answers indicated the possibility of more phases and greater challenges throughout the game. The question of gamification and fun can also be explored in this respect.

All data and values available to the player were created syntactically only to exemplify outlier detection concepts. Due to the time and scope of the project, data from actual bases were not inserted into the game. As future work, it is suggested to create three-dimensional maps with real data or data known by the state of the art. With the creation of a map that has data already classified or grouped by some algorithm, we can make a comparison of which data were selected as outliers by human players. With this, we could test otherwise whether the serious game teaches outlier detection concepts and whether it is possible to use the human ability to manipulate a serious game to classify elements.

Acknowledgment

Our acknowledgment goes to Tribunal de Contas do Estado do Paraná (TCE-PR), Europe - Brazil Collaboration of Big Data Scientific Research Through Cloud-Centric Applications (EuBra-Bigsea), blind reviewers and staff of SBGames. We would also like to thank the Federal University of Technology (UTFPR) and every person who helped us in our work.

References

[1] T. Fawcett and F. Provost, "Adaptive Fraud Detection," *Data Mining and Knowledge Discovery*, vol. 1, no. 3, pp. 291–316, 1997.

[2] A. Abdallah, M. A. Maarof, and A. Zainal, "Fraud detection system: A survey," *Journal of Network and Computer Applications*, vol. 68, pp. 90–113, 2016.

[3] T. Allan and J. Zhan, "Towards fraud detection methodologies," in *2010 5th International Conference on Future Information Technology*, May 2010, pp. 1–6.

[4] Brasil, "Referencial de combate à fraude e à corrupção," *Tribunal de Contas da União (TCU)*, p. 164, 2016.

[5] W. N. Dilla and R. L. Raschke, "Data visualization for fraud detection: Practice implications and a call for future research," *International Journal of Accounting Information Systems*, vol. 16, pp. 1 – 22, 2015. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S1467089515000020

[6] V. Chandola, A. Banerjee, and V. Kumar, "Anomaly detection: A survey," *ACM Comput. Surv.*, vol. 41, no. 3, pp. 15:1–15:58, Jul. 2009. [Online]. Available: http://doi.acm.org/10.1145/1541880.1541882

[7] G. Wallner and S. Kriglstein, "Visualization-based analysis of gameplay data - a review of literature," *Entertainment Computing*, vol. 4, no. 3, pp. 143 – 155, 2013. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S1875952113000049

[8] K. Millis, C. Forsyth, H. Butler, P. Wallace, A. Graesser, and D. Halpern, *Operation ARIES!: A Serious Game for Teaching Scientific Inquiry*. London: Springer London, 2011, pp. 169–195. [Online]. Available: http://dx.doi.org/10.1007/978-1-4471-2161-910

[9] W. Westera, "How people learn while playing serious games: A computational modelling approach," *Journal of Computational Science*, vol. 18, pp. 32 – 45, 2017. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S1877750316304483

[10] M. Muratet, P. Torguet, J.-P. Jessel, and F. Viallet, "Towards a serious game to help students learn computer programming," *Int. J. Comput. Games Technol.*, vol. 2009, pp. 3:1–3:12, Jan. 2009. [Online]. Available: http://dx.doi.org/10.1155/2009/470590

[11] M. Ahmed, A. N. Mahmood, and M. R. Islam, "A survey of anomaly detection techniques in financial domain," *Future Generation Computer Systems*, vol. 55, pp. 278–288, 2015.

[12] L. Akoglu, H. Tong, and D. Koutra, "Graph based anomaly detection and description: a survey," *Data Mining and Knowledge Discovery*, vol. 29, no. 3, pp. 626–688, 2015. [Online]. Available: http://dx.doi.org/10.1007/s10618-014-0365-y

[13] P. Branco, L. Torgo, and R. P. Ribeiro, "A survey of predictive modeling on imbalanced domains," *ACM Comput. Surv.*, vol. 49, no. 2, pp. 31:1–31:50, Aug. 2016. [Online]. Available: http://doi.acm.org/10.1145/2907070

[14] R. Bansal, N. Gaur, and S. N. Singh, "Outlier detection: Applications and techniques in data mining," in *2016 6th International Conference - Cloud System and Big Data Engineering (Confluence)*, Jan 2016, pp. 373–377.

[15] M. Zyda, "From visual simulation to virtual reality to games," *Computer*, vol. 38, no. 9, pp. 25–32, Sep. 2005. [Online]. Available: http://dx.doi.org/10.1109/MC.2005.297

[16] D. Hooshyar, C. Lee, and H. Lim, "A survey on data-driven approaches in educational games," in *2016 2nd International Conference on Science in Information Technology (ICSITech)*, Oct 2016, pp. 291–295.

[17] G. Siemens and R. S. J. d. Baker, "Learning analytics and educational data mining: Towards communication and collaboration," in *Proceedings of the 2Nd International Conference on Learning Analytics and Knowledge*, ser. LAK '12. New York, NY, USA: ACM, 2012, pp. 252–254. [Online]. Available: http://doi.acm.org/10.1145/2330601.2330661

[18] S. Benmakrelouf, N. Mezghani, and N. Kara, "Towards the identification of players' profiles using game's data analysis based on regression model and clustering," in *Proceedings of the 2015 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining 2015*, ser. ASONAM '15. New York, NY, USA: ACM, 2015, pp. 1403–1410. [Online]. Available: http://doi.acm.org/10.1145/2808797.2809429

[19] F.-L. Fu, R.-C. Su, and S.-C. Yu, "Egameflow: A scale to measure learners enjoyment of e-learning games," *Computers and Education*, vol. 52, no. 1, pp. 101 – 112, 2009. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0360131508001024

[20] G. Petri and C. G. von Wangenheim, "How games for computing education are evaluated? a systematic literature review," *Computers and Education*, vol. 107, no. Supplement C, pp. 68 – 90, 2017. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0360131517300040

[21] A. Calderon and M. Ruiz, "A systematic literature review on serious games evaluation: An application to software project management," *Computers and Education*, vol. 87, no. Supplement C, pp. 396 – 422, 2015. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0360131515300166

[22] G. Petri, C. G. von Wangenheim, and A. F. Borgatto, "Meega+: an evolution of a model for the evaluation of educational games," *INCoD/GQS*, vol. 3, 2016.