Knowledge-Based Recommender System in a Gamified Environment for Computational Thinking Development

Klinge Villalba-Condori[1], Julio Vera-Sancho[1], Cesar Vera-Vasquez[2], Jorge Mamani-Calcina[3] and Hector Cardona-Reyes[4,*]

[1]*Universidad Católica de Santa María, Arequipa, Perú*
[2]*Universidad Continental, Arequipa, Perú*
[3]*Universidad Tecnológica del Perú*
[4]*CONACYT-CIMAT, Zacatecas, Mexico*

Abstract

There is currently a great digital divide in Peru, only 48.7% of the population aged six and over have access to the Internet, preventing them from accessing resources for the development of computational thinking. Gamification in the educational context helps the development of computational thinking by understanding concepts of its application, but many of them do not have a data processing approach from interactions, in this research project we propose the development of a knowledge-based recommender system, in a supervised learning model, which is implemented in a video game, helping the student learning process, under the STEM framework, in this proposal students obtained improvements in the learning process, going from obtaining a 5.658 out of 14 to 10.150, and noting the interest of students in this type of technology..

Keywords

Gamification, STEM, recommender system

1. Introduction

Currently, we have a large number of students, who receive basic training that is not following the vanguard of the educational system, such as computational thinking. Part of our reality is the still existing digital gap that exists in our country Peru since only 48.7% of the population of 6 years and older have access to the Internet and only 58.2% of the urban area has a connection, while only 15.4% [1] of the rural area has a connection. This digital gap makes it impossible to access resources and technologies to train teachers and the development of computational thinking in students; also in Peru, at least 52.1% of people have internet access in Lima and 17.5% in the rest of the country [1]. This leads us to the idea of proposing the use of mobile devices, as they are becoming more accessible, and can support students in their learning about the operation of algorithms through video games that give them basic knowledge and thus increase their knowledge. and thus increase the interest in pursuing a technological career and that has a great need in the global market.

The use of educational games as learning tools is promising because of their ability to teach and reinforce not only knowledge but also important skills such as problem-solving, collaboration, and communication, as well as a motivational effect, because they use several mechanisms to encourage people to engage with them, often without any reward, just for the joy of playing and with the

HCI 2022: VIII Iberoamerican Conference of Human Computer Interaction, October 13-15, 2022, La Habana, Cuba
EMAIL: kvillalba@ucsm.edu.pe (K. Villalba-Condori); jveras@ucsm.edu.pe (J. Vera-Sancho); cverav@continental.edu.pe (C. Vera-Vasquez); e16187@utp.edu.pe (J. Mamani-Calcina); hector.cardona@cimat.mx (H. Cardona-Reyes)
ORCID: 0000-0002-8621-7942 (K. Villalba-Condori); 0000-0001-5526-5223 (J. Vera-Sancho); 0000-0003-4168-5117 (C. Vera-Vasquez); 0000-0001-6633-2102 (J. Mamani-Calcina); 0000-0002-9626-6254 (H. Cardona-Reyes)

possibility of winning. However, creating a highly engaging and comprehensive instructional game is difficult, time-consuming, and costly [2, 3, 4].

Gamification has been used in a variety of application domains to promote behavior change, and while proposing a recommendation system for personalized gamification is not yet used to perfection, recent research shows that personalized gamification approaches can achieve better results than generic approaches. However, a general framework for creating personalized game applications is still lacking, so framing within Human-Computer Interaction (HCI), and can take data: items, users, transactions, and contextual information. Achieving results that the framework contributes to understanding and building effective persuasive and playful applications by describing the different building blocks of a recommender system [5].

A recommendation system in the entertainment domain, specifically in the domain of movie schedules as RecomMetz accompanied by a context-based mobile recommendation system using Semantic Web technologies, develops a domain ontology that primarily serves a semantic similarity metric tuned to the concept of "individual item bundles", which allows leveraging schedule correlation identification and inferences about a user's preferences, which demonstrates that user profiles are not modeled using ontologies [6].

A model of educational content recommendation is based on the context of a user, and by the context, it incorporates the role, tasks, and programming exercises and that recommendations are made based on the estimation of the difference that exists between the level of knowledge of a user versus the skills required by the user. The experiments developed in the student context show that when using a probabilistic reasoning model using Bayesian Networks, this model helps to obtain better recommendations of educational content, according to the skills that the student requires, but also allows to capture the student's interactions through a web platform, allows to improve the development of computational thinking [7, 8, 9].

After analyzing the context and a review of proposals that are framed in the development of a recommendation system to improve the learning process, the objective of this research is to develop a tool for a mobile video game that allows classifying student learning according to their interaction with the video game and its results, using supervised automatic learning techniques such as Support Vector Machine (SVM), which will be analyzed and it will be seen if the student learned or not and it will be recommended what actions should be followed to make up for the lack of information on a specific topic through a repository of computational thinking topics within the framework of STEM.

2. Background

2.1. Gamification

It is the use of techniques, elements, and dynamics typical of games and leisure in non- recreational activities to enhance motivation, as well as to reinforce behavior to solve a problem, improve productivity, obtain an objective, activate learning, and evaluate specific individuals [10].

In this context, it is important to provide the differentiation between two concepts: gamification and gamification; gamification may or may not involve a playful environment, while a game environment necessarily includes a playful component, and that gamified environments have a positive correlation with the motivation of learners with the level of participation in learning activities [11].

2.2. Recommender systems

Recommender systems provide modified suggestions for different things such as hotels, restaurants, movies, books, resources, etc. Such recommendations are based on the information available in the system, and the system suggests items similar to favorable items mentioned in a query by the users [12]. Recommender systems study the characteristics of each user and by processing the data, find a subset of items that may be of interest to the user [12].

The recommender system can predict whether a particular user prefers an item or not based on the user's profile. Recommender systems are beneficial to both service providers and users. Recommender systems have also been shown to improve the decision-making process [13].

2.3. Recommender systems

2.3.1. Supervised Learning Systems

Supervised learning, which relies on prior knowledge about a data set, is concerned with predicting the value of a response variable, or label (either a categorical or continuous value), based on the input variables/features. Supervised learning accomplishes this feat by using a training set of labeled data examples [14].

2.3.2. Unsupervised Learning Systems

Unsupervised learning consists of finding patterns in data that are initially considered pure unstructured noise. Two very simple classical examples of unsupervised learning are clustering and dimensionality reduction. The machine simply receives inputs (x1, x2, ...) but does not obtain objective results or any guiding data. The goal of this learning is to build representations of the input that can be used for decision making, predict future inputs, efficiently communicate inputs to another machine, etc. [14].

2.4. STEM Methodology

STEM is a curriculum based on the idea of educating students by integrating Science, Technology, Engineering, and Mathematics, intending to increase scientific and engineering power where it is applied [15].

U.S. Bureau of Labor Statistics 2018 analysis, most STEM careers, which are composed [15]:

- Computing - 71%
- Traditional Engineering - 16%
- CPhysical Sciences: 7%
- Life Sciences: 4%
- Mathematics - 2%

3. Materials and methods

3.1. Anaconda Navigator

This tool will be used because it is an open-source suite that encompasses a series of applications, libraries, and concepts designed for the development of data science with Python. In general terms, Anaconda Distribution is a Python distribution that works as an environmental manager, and a package manager and has a collection of more than 720 open-source packages.

Anaconda Distribution is grouped into 4 sectors or technological solutions, Anaconda Naviga- tor, Anaconda Project, Data Science Libraries, and Conda. All of these are installed automatically and in a very simple procedure. One of its main features is that it is open source, with detailed documentation and a large community, helps to develop data science projects using various IDEs such as Jupyter, JupyterLab, Spyder, and RStudio, and has tools such as Dask, NumPy, pandas, and Numba to analyze data and allows data visualization with Bokeh, Datashader, Holoviews or Matplotlib [16].

3.2. Unity3D

It is a multiplatform video game engine created by Unity Technologies. Unity is available as a development platform for Microsoft Windows, OS X, and Linux. The development platform has compiler support with different types of platforms. As of its version 5.4.0, it no longer supports browser content development through its web plugin, instead, WebGL is used. Unity has two versions: Unity Professional (pro) and Unity Personal.

The graphics engine uses OpenGL (on Windows, Mac, and Linux), Direct3D (Windows only), OpenGL ES (on Android and iOS), and proprietary interfaces (Wii). It has support for bump mapping, reflection mapping, parallax mapping, ambient occlusion in screen space, dynamic shadows using shadow maps, render to texture and full-screen post-processing effects which will help us a lot for the levels we will develop for the video game [17].

In the following Figure 1 the game developed with Unity is shown below. The game consists of two mini-games, one by which you have to move a block from an initial side to a final side, and according to the time it takes, and the number of movements calculated the understanding of the field of Pattern Recognition, Abstraction, and the use of Algorithms that had to follow to meet the completed level.

In the following Figure 2 is shown below the game towers of Hanoi, in which the student had to move the blocks of the tower from one side to another, initially only one tower was used, but with the maximum difficulty the three characteristic towers of the game are used, in which the student has to demonstrate his level of Decomposition to understand that the level will only be satisfied when a smaller block is on top of a larger one and reach a predetermined limit of movements and in a specific time, their level of Pattern Recognition by realizing that the level will be satisfied when they have all the blocks from the tower on the left to the tower on the right and their level of Abstraction to divide the problem into smaller, more achievable goals to reach the goal satisfactorily.

Figure 1: Minimal movement game - Recall. Source: Self-made.

3.3. Supervised learning algorithm

3.3.1. Support Vector Machine SVM

It is a supervised learning technique that is capable of classifying different categories of data from various disciplines. These have been used for two-class classification problems and are applicable in linear and nonlinear data classification tasks. SVM creates a hyperplane or multiple hyperplanes in a high-dimensional space, and the best hyperplane in them is the one that optimally divides the data into different classes with the largest separation between the classes. A nonlinear classifier uses several kernel functions to estimate the margins. The main objective of these kernel functions (i.e., linear, polynomial, radial basis, and sigmoid) is to maximize the margins between hyperplanes [18].

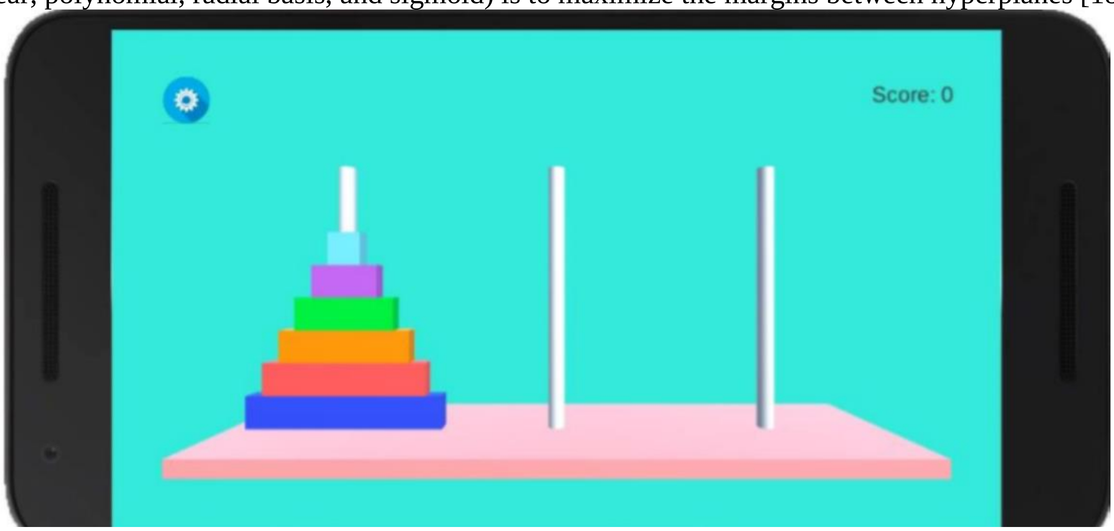

Figure 2: Sudoku game. Source: Self-made.

Table 1
Paired Samples T-Test

Measure 1	Measure 2	T	df	p	Cohen's d
Pre-treatment	Post-treatment	-23.36	119	<0.01	-2.13

4. Results

The population to be analyzed are the students of a school in Arequipa, which are approximately 120 students in elementary school. For our sample, we will use the following statistical formula.

$$n = \frac{Z^2 * p * q}{d^2}$$

Our confidence level (Z) will be 99%, and for our variability (p and q) we will use the maximum variability because we start from a starting point in which we do not know the classification of the computational thinking knowledge of our population, so the values will be p=q=0.5 or 50.

Finally, we will take a maximum permissible error (d) of 5 percent in terms of proportion and on an all-to-all basis. Replacing all these variables, a representative sample for our population would be approximately 50 students.

Applying a Shapiro Wilk normality test for two groups, we obtain normality of 0.98%. The P value, in this case, is 0.52 higher than the minimum 0.05 required, so it can be affirmed that normality is present.

In Table 2 the effect size, according to Cohen's d indicator (see Table 1) is 2.132, so the difference is large.

The study has a sample of 120 study subjects (see Table 3) where it can be seen that the mean score in the Pre-Test is 5.658 out of 14, this varies in the Post-Test up to 10.150 out of 14 indicating a significant increase in the average final score.

Table 2
Effect size

Test	Measure	Irrelevant	Small	Medium	Large
Pre-treatment	Post-treatment	<0.2	0.2	0.5	0.8

Table 3
Descriptive statistical values for the sample

	N	Mean	SD	SE
Pre-treatment	120	5.66	1.5	0.14
Post-treatment	120	10.15	1.94	0.17

5. Discussion

The descriptive statistical values show a higher mean in the scores obtained in the questionnaire given to the primary level students after the treatment. This shows a higher degree of computa- tional thinking after the treatment, and the Shapiro Wilk test shows that the sample is within statistical normality. The Student's t-test shows that there is a significant difference between the level of computational knowledge before and after the treatment and finally Cohen's index shows that this significant difference is large.

6. Conclusions

In recent years, we have seen the importance of computational thinking and gamification, and how it has improved students' skills within the STEM framework. In this research work, we propose the creation of a recommendation system for the development of computational thinking through gamification, finding great improvements in the learning process of students, going from obtaining a 5.658 out of 14 to 10.150, and noting the interest of students to this type of technologies, even more in our new reality that we are living all people.

7. References

[1] Instituto nacional de estadistica e informatica, https://m.inei.gob.pe/, ???? (Accessed on 08/01/2022).

[2] K. M. Kapp, The gamification of learning and instruction: game-based methods and strategies for training and education, John Wiley & Sons, 2012.

[3] M. Kalogiannakis, S. Papadakis, A.-I. Zourmpakis, Gamification in science education. a systematic review of the literature, Education Sciences 11 (2021) 22.

[4] K. O. V. Condori, Teaching formation to develop computational thinking, in: Global implications of emerging technology trends, IGI Global, 2018, pp. 59–72.
[5] G. F. Tondello, R. Orji, L. E. Nacke, Recommender systems for personalized gamifica- tion, in: Adjunct publication of the 25th conference on user modeling, adaptation and personalization, 2017, pp. 425–430.
[6] L. O. Colombo-Mendoza, R. Valencia-García, A. Rodríguez-González, G. Alor-Hernández, J. J. Samper-Zapater, Recommetz: A context-aware knowledge-based mobile recommender system for movie showtimes, Expert Systems with Applications 42 (2015) 1202–1222.
[7] J. Vera, S. Castro Cuba-Sayco, K. Villalba-Condori, Modelo de sistema de recomendación basado en el contexto a partir del análisis de código estático para el desarrollo del pen- samiento computacional: Caso de programación web (2018).
[8] Y. A. Rodríguez del Rey, I. N. Cawanga Cambinda, C. Deco, C. Bender, R. Avello-Martínez, K. O. Villalba-Condori, Developing computational thinking with a module of solved problems, Computer Applications in Engineering Education 29 (2021) 506–516. URL: https://onlinelibrary.wiley.com/doi/abs/10.1002/cae.22214. doi:https://doi.org/10.1002/cae.22214.
[9] K. O. Villalba-Condori, C. Deco, L. M. Oliva Córdova, S. E. Castro Cuba Sayco, C. Bender, Silva Sprock, A methodology to assign educational resources with metadata based on the purpose of learning, in: 2019 XIV Latin American Conference on Learning Technologies (LACLO), 2019, pp. 221–225. doi:10.1109/LACLO49268.2019.00045.
[10] M. Sailer, J. U. Hense, S. K. Mayr, H. Mandl, How gamification motivates: An experimental study of the effects of specific game design elements on psychological need satisfaction, Computers in human behavior 69 (2017) 371–380.
[11] M. Aparicio, T. Oliveira, F. Bacao, M. Painho, Gamification: A key determinant of massive open online course (mooc) success, Information & Management 56 (2019) 39–54.
[12] F. Ali, D. Kwak, P. Khan, S. H. A. Ei-Sappagh, S. R. Islam, D. Park, K.-S. Kwak, Merged ontology and svm-based information extraction and recommendation system for social robots, IEEE Access 5 (2017) 12364–12379.
[13] P. Pu, L. Chen, R. Hu, A user-centric evaluation framework for recommender systems, in: Proceedings of the fifth ACM conference on Recommender systems, 2011, pp. 157–164.
[14] D. R. Schrider, A. D. Kern, Supervised machine learning for population genetics: a new paradigm, Trends in Genetics 34 (2018) 301–312.
[15] J. D. Casal, Stem: Oportunidades y retos desde la enseñanza de las ciencias, Universitas Tarraconensis. Revista de Ciències de l'Educació (2019) 154–168.
[16] Anaconda navigator :: Anaconda.org, https://anaconda.org/anaconda/anaconda-navigator, (Accessed on 08/01/2022).
[17] Unity real-time development platform | 3d, 2d vr & ar engine, https://unity.com/, ???? (Accessed on 08/01/2022).
[18] Ahmad, M. Basheri, M. J. Iqbal, A. Rahim, Performance comparison of support vector machine, random forest, and extreme learning machine for intrusion detection, IEEE access 6 (2018) 33789–33795.