A Data-Driven Gamification Approach to Monitor and Predict the Students' Academic Performance

Sherif Abdelhamid[1], Kolby Quigg[1], Mona Aly[2]

Virginia Military Institute[1], *Independent Researcher*[2]
United States of America

Abstract

Learning environments can be more stimulating by incorporating game design elements into the curriculum. Several studies have shown that gamification improves student motivation, learning, and academic performance. According to Karl M. Kapp, gamification is "the ideal process for creating engaging learning environments." Gamification can be implemented in schools at different grade levels, from kindergarten to 12 years of basic education. Additionally, the use of gamification in education has been demonstrated in several fields, including computer science, mathematics, astronomy, physics, medicine, and law. Current research works have extensively explored how gamification can improve students' engagement and motivation, but few studies have examined how it can track and predict students' academic performance. This study investigates how data collected from gamification activities can help instructors monitor and predict students' performance in the classroom. We used Quizizz, a web-based tool that delivers quiz questions in a game-like manner. By answering questions interactively, students earn points and rewards. Students can answer at the instructor's or their own pace and earn points based on their answering speed. Additionally, instructors can use team modes to place students in teams for scoring. For this study, we identified two experimental groups representing sophomore students in two computer science courses: Database Management Systems and Data Structures & Applications. We explored the causal relationship between the final course grade of students and their scores, interactions, and timings during the weekly gamified activities. The study employed a data-driven exploratory and correlational methodology that involves regression analysis to forecast and predict patterns in the course grades.

Keywords: *Gamification, Game-based Learning, Academic Performance, Machine Learning*

1. Introduction

Gamification in education refers to using game design elements and principles in non-game contexts, such as classrooms or online learning environments. Gamification aims to engage and motivate learners and improve learning outcomes. There are many ways that gamification can be used in education. A few examples include: (i) using rewards and incentives for completing tasks or reaching certain goals, (ii) utilizing quizzes and challenges in the learning material to make it more interactive and engaging, (iii) creating leaderboards to show students how they compare to their peers in terms of progress or performance, and (iv) applying virtual or simulations in which students can learn and explore in a more immersive and interactive way.
While most existing research works have focused on using gamification to increase students' engagement and motivation, few studies have examined how it can track and predict students' academic performance. In this work, we explore and analyze the data collected from gamification activities to identify at-risk students and anticipate their academic success represented by their final grades. The product of this research work can be implemented as a software service that can be integrated into learning management systems to help administrators and instructors anticipate the students' final grades and take the needed actions as early as possible to maintain their success and retention.

2. Related work

Several studies have shown that gamification improves student motivation, learning, and academic performance [1, 2]. Gamification is considered one of the best methods for creating engaging learning

environments [3]. Teaching through gamification can be accomplished in all grades, from kindergarten to elementary school [4]. Several fields, including computer science, mathematics, astronomy, physics, medicine, and law, have demonstrated the use of gamification in education [5, 6]. In [7], researchers studied gamification by introducing game design elements, including experience points, levels, leaderboards, challenges, and badges. The results showed greater student engagement and participation in course activities. Another research work [8] used gamification in a web-based learning system called "Classroom Live" for undergraduate students. The researchers included game design elements like experience points, levels, and in-game rewards. The results showed that students were more enjoying and engaged in class. In [9], researchers studied the role of badges with points and leaderboards as powerful ways of creating competitions, achievements, and status.
Based on the current literature review, we found underrepresented research areas related to the use of gamification in education, including 1) performance prediction (the focus of this paper), 2) long-term effectiveness, 3) individual differences, 4) gamification integration with other teaching methods and 5) ethical considerations and concerns.

3. Methodology

We followed a quantitative correlational analysis to examine the relationships between multiple variables representing students' performance metrics during each gamified activity (e.g., *scores, interactions, and timings)*. Additionally, our multi-method approach involved a linear regression analysis. The study's main steps are summarized in Fig. 1. In phase 1 (data collection), we generated all quiz reports from the *Quizizz* platform for the two courses (*Database Management Systems and Data Structures & Applications)*. Each report contains information about each student, including total time taken, accuracy, number of correct/incorrect answers, and number of answered/unanswered questions. This data represents our variables of interest. Following this step, we aggregated data by calculating the average of all students' records across all the reports. After this, we cleaned our data from any missing or inconsistent data.

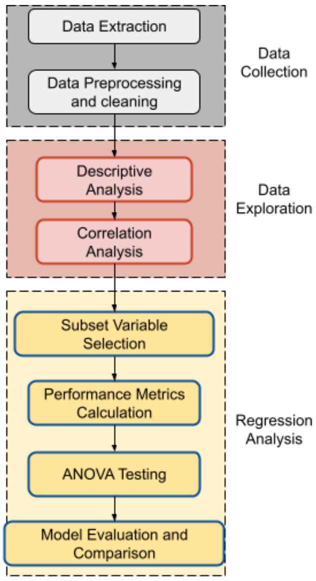


Fig. 1. The research methodology's main steps.

In phase 2, we conducted a correlational analysis among multiple variables of interest. Creswell [10] defines quantitative research as gathering, analyzing, interpreting, and presenting data. Furthermore, Creswell emphasizes using correlational study design to create predictive models. Shirish [11] explains that this design is suitable as it aims to measure the correlation between multiple variables using statistical data (p. 71). During the final phase (regression analysis), to build and identify the best model, the most straightforward method for variable selection is simply to try all of the subsets. Unfortunately, if the number of predictors is k, we must try all 2^k subsets. Thus, the number of subsets grows exponentially, causing exponential running times, which are considered extremely inefficient. To avoid this computationally expensive approach, we produced the best subsets regression with only four predictor variables. Best subsets regression is an exploratory model-building regression analysis that compares all possible models that can be created based on our four predictor variables. We use the R function regsubsets() from the leaps package to identify different best models of different sizes. The regsubsets function returns separate best models of all sizes up to 4 and uses different model selection criteria such as adjusted r^2, BIC (Bayesian information), and Cp. Finally, to further compare the fits of the models, we used the anova() function in R to conduct ANOVA testing [12] and determine whether the more complex model is significantly better at capturing the data than the simpler models.

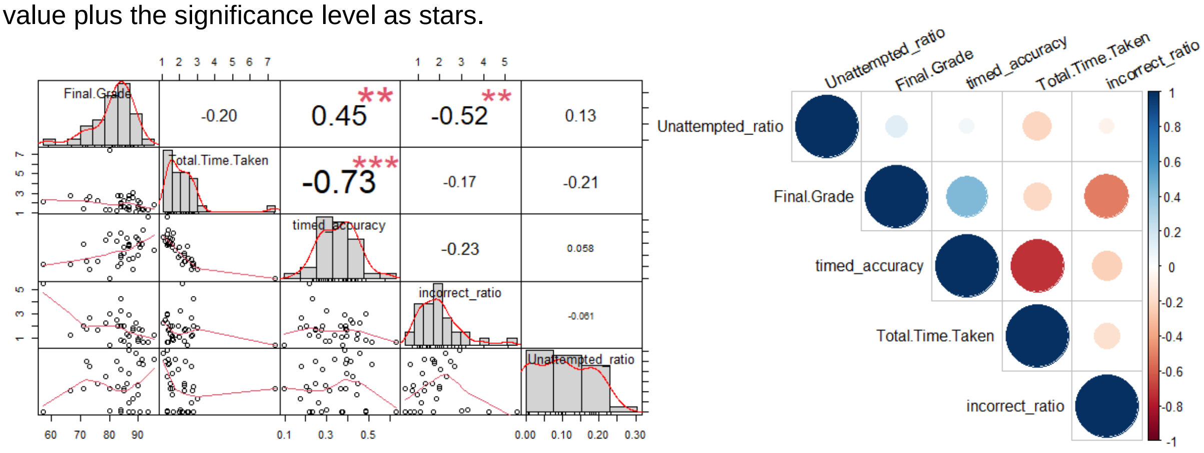
4. Results and Findings

The students' average final grade is 83.54 (SD = 8.23), with a minimum value of 57 and a maximum value of 95.66. A description of each variable in the student data record is presented in table 1.

Variable	Description	Min	Max	Mean
Final.Grade	Student final test grade in the course	57.0	95.66	83.54
Accuracy	Ratio of correct answers to the total number of questions	0.46	0.91	0.62
Total.Time.Taken	Average time taken by student to complete a quiz	1.09	7.43	2.09
Attempted	Average number of attempted questions in all quizzes	9.10	17.50	10.80
Unattempted	Average number of unattempted questions in all quizzes	0.00	2.80	1.26
Unattempted.Ratio	Ratio of answered to unanswered questions	0.00	0.30	0.13
Score	Average points collected by a student based on speed and correctness	4743.3	10402.0	6574.6
Timed.Accuracy	Ratio of accuracy to average time taken to complete a quiz	0.1	0.63	0.34
Incorrect.Ratio	Ratio of incorrect answers to correct ones	0.41	5.52	1.82

Table 1. Description of the student performance data used in the study. Selected variables were used as predictors, including the "Total.Time.Taken", "Unattempted.Ratio", "Timed.Accuracy" and "Incorrect.Ratio".

The distributions of the predictor variables are shown in fig 2.a. The bivariate scatter plots with a fitted line are displayed on the bottom of the diagonal. A bivariate plot lets us see the degree and pattern of relationships between each two variables. A large amount of scatter around the line indicates a weak relationship. Little scatter represents a strong relationship. If all points fall directly on a straight line, we have a perfect linear relationship between our two variables. The direction of the line indicates whether the relationship is positive, negative, or zero. On the top of the diagonal is the correlation value plus the significance level as stars.



(a) (b)

Fig. 2. (a) Variable distributions, scatter plots, and the correlation value. (b) A correlogram showing the correlations between the predictor variables and the independent variable (Final.Grade).

From figure 2.a, we notice that the "Final.Grade" and "Timed.Accuracy" variables are positively correlated, while the "Final.Grade" and "Incorrect.Ratio" are negatively correlated. A better view of variable correlation is achieved using the correlogram (see figure 2.b). The correlogram is a correlation matrix graph highlighting the most correlated variables in a data table. Correlation coefficients are colored according to the value.

Based on the Pearson's and Kendall's correlation coefficients (0.45 and 0.34, respectively), there is an evident positive correlation between the student's "Final.grade" and "timed.accuracy". Additionally, there is evidence of a negative correlation between the student's "Final.Grade" and the "Incorrect.Ratio", with Pearson's and Kendall's coefficients equal -0.32 and -0.52, respectively.

A simple way to select variables for a model is to consider all possible combinations of the available variables. However, this approach can be extremely time-consuming when the number of variables is large. We used best subsets regression to identify the best models using four predictor variables. This technique compares all possible models created based on our four predictor variables. The results showed that the model with two predictor variables (Incorrect.Ratio and Timed.Accuracy) has the highest adjusted r^2, which measures how well the data fit the regression line. Additionally, the same model has the lowest BIC and Mallows' Cp values. BIC is a criterion for model selection that penalizes models with more parameters. A lower BIC value indicates a better model. Similarly, Cp is a measure of unexplained error (smaller Cp values are better). The results of the best subset analysis are shown in Table 2.

Model	Variables	R^2	BIC	Cp
1-variable	Incorrect.Ratio	0.24	-3.31	4.43
2-variable	Incorrect.Ratio+ Timed.Accuracy	0.33	-5.32	1.30
3-variable	Incorrect.Ratio+ Timed.Accuracy+ Unattempted.Ratio	0.32	-2.16	3.01
4-variable	Incorrect.Ratio + Timed.Accuracy + Unattempted.Ratio + Total.Time.Taken	0.29	1.31	5.0

Table 2: Independent variables used for each model. The 2-variable regression model (highlighted in grey) has shown better results regarding adjusted r^2, BIC, and Cp values.

The final step in our study was conducting the ANOVA test between the four models. Three pairwise comparisons were performed: 1) 1-variable vs. 2-variable models, 2) 2-variable vs. 3-variable models, and 3) 2-variable vs. 4-variable models. The first pairwise comparison showed a degree of freedom of 1 (indicating that the 2-variable model has one additional parameter over the 1-variable model, and a p-value equals 0.0265 (< 0.05). Adding the "Timed.Accuracy" led to a significantly improved fit. The second comparison showed a degree of freedom of 1 (indicating that the 3-variable model has one additional parameter over 2-variable), and a p-value equals 0.5878 (> 0.05). Adding the "Unattempted.Ratio" variable to the model led to no significant improvement fit. Finally, the last comparison showed a degree of freedom of 2 (indicating that the 4-variable model has two additional parameters over the 2-variable one) and a p-value equals to 0.8606 (> 0.05). Adding the "Unattempted.Ratio" and "Total.Time.Taken" variables to the 2-variable model led to no significant improvement fit. These results indicate that the 2-variable model has the best fit for the data, the lowest prediction error, and the lowest complexity compared to the other models.

5. Conclusion and Future Work

Previous research have extensively studied how gamification can increase student engagement and motivation. Still, there is limited research on how it can be used to track and predict academic performance. This study examined how gamified classroom activities can be used to track and predict students' academic performance represented by their final course grades. We specifically used Quizizz, a web-based tool that delivers quiz questions in a game-like manner. We focused on two groups of sophomore computer science students in Database Management Systems and Data Structures & Applications courses. The study explored the relationship between students' final course grades and their scores, interactions, and timings during the weekly gamified activities through data-driven correlational analysis. Results showed positive correlations between the students' accuracy in the gamified activities and their final grades. Students with a higher incorrect-to-correct answers ratio were found to have lower final course grades. Finally, we built various regression models within a comparative analysis. Results indicated that the 2-variable ("Incorrect.Ratio" and "Timed.Accuracy") model has the best fit for the data, the lowest prediction error, and the lowest complexity compared to the other models. Our future goal is to increase the dataset sample size to achieve more robust results. Our future work will include other academic performance metrics and will examine more gamification tools and platforms. We are planning to Integrate the research work as a software service within CyEd [13], a cyberinfrastructure for computer education.

Acknowledgment

This work was supported in part by the Commonwealth Cyber Initiative, an investment in the advancement of cyber R&D, innovation, and workforce development. For more information about CCI, visit cyberinitiative.org.

References

[1] Liu, T., & Lipowski, M. (2021). Sports gamification: Evaluation of its impact on learning motivation and performance in higher education. International Journal of Environmental Research and Public Health, 18(3), 1267.
[2] Kalogiannakis, M., Papadakis, S., & Zourmpakis, A. I. (2021). Gamification in science education. A systematic review of the literature. Education Sciences, 11(1), 22.
[3] Kapp, K. M. (2012). The gamification of learning and instruction: game-based methods and strategies for training and education. John Wiley & Sons.
[4] Varannai, I., Sasvári, P. L., & Urbanovics, A. (2017). The use of gamification in higher education: an empirical study. International Journal of Advanced Computer Science and Applications, 8(10), 1-6.
[5] Kasahara, R., Sakamoto, K., Washizaki, H., & Fukazawa, Y. (2019). Applying gamification to motivate students to write high-quality code in programming assignments. 2019 ACM Conference on Innovation and Technology in Computer Science Education (pp. 92-98).
[6] Nevin, C. R., Westfall, A. O., Rodriguez, J. M., Dempsey, D. M., Cherrington, A., Roy, B. & Willig, J.H. (2014). Gamification as a tool for enhancing graduate medical education. Postgraduate medical journal, 90(1070), 685-693.
[7] Barata, G., Gama, S., Jorge, J., & Gonçalves, D. (2013). Engaging engineering students with gamification. In 2013 5th International Conference on Games and Virtual Worlds for Serious Applications (VS-GAMES) (pp. 1-8). IEEE.
[8] De Freitas, A. A., & de Freitas, M. M. (2013). Classroom Live: a software-assisted gamification tool. Computer Science Education, 23(2), 186-206.
[9] Wilson, B. G., Gasell, C., Ozyer, A., & Scrogan, L. (2016). Adopting digital badges in higher education: Scoping the territory. In Foundation of digital badges and micro-credentials (pp. 163-177). Springer, Cham.
[10] Creswell, J. W. (2002). Educational research: Planning, conducting, and evaluating quantitative (Vol.7). Prentice Hall Upper Saddle River, NJ.
[11] Shirish, T. S. (2012). Research methodology in education. Lulu. com.
[12] St, L., & Wold, S. (1989). Analysis of variance (ANOVA). Chemometrics and intelligent laboratory systems, 6(4), 259-272.
[13] Abdelhamid, S., Mallari, T., & Stower, T. (2022). CyEd: A Cyberinfrastructure for Computer Education. In Innovative Approaches to Technology-Enhanced Learning for the Workplace and Higher Education: Proceedings of 'The Learning Ideas Conference'2022 (pp. 3-14).