Exploring Gamification Elements in Educational Applications for Python Programming*

Nathan Dihh[1], and Christian Lopez[1,2]
[1]Computer Science Department
[2]Mechanical Engineering Department
Lafayette College,Easton, PA 18042
lopezbec@lafayette.edu

Abstract

This paper explores the gamification elements and user sentiment in gamified educational applications aimed at teaching Python programming. Programming skills have become increasingly valuable. Particularly, Python has gained significant popularity due to its simplicity and versatility but learning it can still be difficult under traditional teaching methodologies. To address these challenges, gamification has been explored as a method to make programming more accessible and enjoyable. This paper analyzes users' reviews, both in English and Spanish, of gamified applications that teach Python programming to understand how game elements are leveraged and perceived. The results indicate the prevalence of game elements such as progress bars and point systems, with positive perceptions from users. Although concerns were raised regarding certain game elements and the need for internet access. Moreover, the analysis highlights the effectiveness of gamification strategies while suggesting areas for improvement, like the importance of considering language preferences and addressing language limitations for non-English speaking users.

1 Introduction

With its vast array of applications in fields as diverse as scientific research, business, and software development, programming has become the backbone of many contemporary disciplines and industries [29]. However, the learning curve associated with programming could be steep and intimidating for many beginners. The abstract nature of programming concepts, compounded by the need for precision and logical thinking, might make it a challenging discipline to grasp and excel at [3].

Amidst an expansive range of programming languages, Python, in particular, has gained significant traction due to its simplicity, versatility, and widespread adoption in areas like Data Science, Machine Learning, and Web Development [26]. Even with Python's relative simplicity (i.e., high-level language), programming in it is still challenging to learn, especially under traditional teaching methodologies, which often fail to effectively engage students or cater to their learning needs [17].

To mitigate these challenges and make programming more accessible and enjoyable, educators have been exploring the application of gamification in teaching programming. Gamification, or the application of game-design elements in non-game contexts, has been shown to increase user engagement, motivation, and ultimately, learning outcomes [15]. By transforming complex programming concepts into interactive challenges and rewarding students for their progress with multiple game elements, gamified applications can make learningprogramming an immersive and enjoyable experience [15].

Crucial to successful gamification are the game elements and their implementation. Engaging game elements (such as Level, Challenges, Points, etc.) not only enhance the user experience but also aid the learning process by engaging learners in different ways than traditional methods. These elements, in their effective implementation, can tap into users' intrinsic motivations, promoting continued engagement and facilitating deep learning [18]. However, the reception and effectiveness of these gamified elements could vary between individuals, influenced by their personal experiences and the way these elements are implemented [13]. Thus, the effectiveness of gamification can be substantially enhanced by tailoring them to individual users' needs, preferences, and abilities.

In order to personalize these experiences effectively, it is crucial to identify the features of gamified applications that are most conducive to learning. There are popular and effective educational applications currently in use (e.g.,SoloLearn and Mimo), that effectively use gamification to enhance their users' learning. By analyzing the online user reviews of existing gamified applications that aim to teach Python programming, this paper delves into the process of exploring the game elements across these applications and attempts to

gain an understanding of how they leverage them and how they are perceived by users. Through understanding the specific game elements that resonate with users and contribute to successful applications, a better understanding of the effects of gamification elements on learning Python programming could be gained with the goal of contributing towards more personalized and effective programming gamified educational applications.

2 Literature Review

The concept of gamification has become a central theme in modern education systems, owing to its demonstrable potential to improve learner engagement, motivation, and achievement [37]. Gamification refers to the incorporation of game elements such as points, badges, leaderboards, quests, and progress tracking into educational settings. These game elements could serve as performance indicators and provide a system of rewards that incentivize continued engagement [15].

Taking inspiration from the Self-Determination Theory (SDT) [27, 8], existing literature posits that in the context of learning, intrinsic motivation (i.e., where the interest or desire stems from within the individual) tends to yield the most beneficial outcomes [20]. This type of motivation typically results in higher levels of engagement and superior retention of knowledge [27]. The potential of gamification to address motivation-related challenges is often emphasized, as it is frequently employed to enhance motivation [38].

2.1 Gamification in Programming Education

The field of programming education has been increasingly influenced by the principles of gamification. Gamification techniques provide a fresh perspective on learning by making it a less daunting and more enjoyable experience [18, 30, 21]. The inherent complexities and abstract nature of programming concepts make it a challenging discipline for learners [31, 24]. Therefore, the incorporation of playful and game-like elements can significantly contribute to breaking down these complexities and promoting a deeper understanding.

As such, the implementation of gamified platforms in this domain presents an opportunity to transform programming education. These platforms, by using game elements such as challenges, quests, puzzles, and immediate feedback, stimulate a proactive learning approach. They serve to push students beyond passive consumption of information to active engagement with programming concepts, hence fostering problem-solving skills and creativity [30].To further reinforce the learning process and celebrate the incremental achievements of learners, numerous platforms have incorporated game elements like progress bars, badges, and points systems. These tangible signs of progression provide

learners with a sense of accomplishment, thereby motivating them to continue their learning journey [30, 16].

In an effort to evaluate the influence of different game elements on student outcomes, Imran [14] conducted a study with 450 undergraduate students in an "introduction to programming" 13-week course. Besides testing different game elements, they also change the gamification level (number of game elements implemented at once). Their findings indicate that gamification had an impact on student's motivation, engagement, and performance. Moreover, the gamification level was shown to impact motivation and performance. Thus, concluding that to improve student motivation and performance it is key to use the 'right level' of gamification.

Mihai et al. [4] conducted a systematic review of gamification studies in e-learning platforms. Their key findings were tasks, levels, and avatars affected satisfaction and attitude towards the gamified e-learning platform; achievements, badges, and rewards positively affected performance; badges, rewards, trophies, and leaderboards had a positive impact on attitude toward gamification, performance, engagement, and enjoyment. Regarding the factors for developing effective gamification, the study noted that the effectiveness of gamified e-learning platforms varied across disciplines. In the context of programming education, these findings demonstrate that gamification can be an effective tool to engage students, enhance their learning experience, and improve their performance. However, despite these promising developments, current gamified programming platforms face certain challenges. Notably, they often fall short when it comes to personalization and catering to the diverse learning needs and preferences of individual learners. As a result, such a lack of personalization may result in decreased engagement and sub-optimal learning outcomes [34].

Current research on the impact of gamification underpins the notion that for gamification to be truly efficacious, it ought to be customized for its users. Specifically, within the realm of education, the focus of adaptation has predominantly been on tailoring educational content to suit learners and their specific contexts - a line of inquiry that is extensively researched [33] and has proven successful. This growing recognition of the limitations of existing platforms underscores the need for more personalized approaches in gamified programming education.

A systematic exploration of adaptive gamification in the context of e-learning was presented in [2]. This study offers a comprehensive overview of adaptive gamification by synthesizing key findings from existing literature. The findings of this literature review underscore the potential of adaptive gamification in e-learning. Personalizing gamification elements and designs based on learner preferences, traits, and learning styles have been found to signifi-

cantly enhance engagement and motivation. However, several challenges, including the need for effective personalization, engagement strategies, technological infrastructure, resource allocation, assessment mechanisms, ethical considerations, and optimal design approaches, must be addressed to ensure the successful implementation of adaptive gamification. Therefore, to help address issues of engagement strategies and optimal design approaches in programming educational gamified applications, it is key to have a good understanding of what game elements users would prefer and how the elements impact their interactions. Toward this end, this work will use Data-mining techniques to analyze online users' reviews of existing gamified applications designed to teach Python programming.

2.2 Mining Online User Reviews

In the digital age, online user reviews have become a significant source of valuable data for decision-making. In today's era of rapidly evolving digital technologies, consumers have easy access to products and services through online platforms and digital marketplaces [9]. As part of their purchasing process, potential buyers often consider the reviews and recommendations of others available on e-commerce websites or specialized review platforms. They offer authentic insights into users' experiences, needs, and feedback regarding products and services [32]. Reviews often contain constructive criticisms or suggestions for improvement. They can highlight potential areas of concern or dissatisfaction, whether related to the product, service, or the overall user experience, uncovering how users feel about specific aspects of a product or service [6].

Within the context of app and software development, mining online user reviews has emerged as a critical tool for developers. User reviews can provide a wealth of insights about the performance of an application, including user satisfaction, areas of difficulty, and potential areas for improvement [7]. Developers can use these insights to identify and address bugs, understand user preferences, and tailor app features to better meet user needs. Recent research has demonstrated that integrating feedback from user reviews into the software development process can significantly improve both the application's quality and user satisfaction levels [10, 19].

Various tools and methods have been employed to analyze user reviews. Natural Language Processing (NLP) techniques, such as sentiment analysis, have proven to be effective in extracting sentiment polarity from text [28]. Sentiment analysis, an automated process of discerning attitudes or emotions, has been applied across various domains. Machine learning algorithms, including supervised and unsupervised learning approaches, have been utilized to classify reviews into different categories and identify patterns within the data

[22]. Topic modeling algorithms, such as Latent Dirichlet Allocation (LDA), have been used to uncover latent themes and topics within the review corpus [11, 36]. Furthermore, text mining techniques have facilitated the extraction of relevant keywords and phrases from reviews [25]. Lastly, Semantic Networks have been used as a visual tool to identify structure and pattern on large volumes of text data [35, 12]. These tools and methods enable researchers to gain valuable insights from user reviews and contribute to the improvement of products and services.

By examining the current literature on gamification in education and programming, it is clear that a personalized, user-centric approach can significantly enhance the effectiveness of gamified programming platforms, and that a vital part of this effort is the identification of effective gamification features. Hence, this work applied Sentiment and Semantic Network Analyses on the online user's reviews of existing gamification applications that teach Python Programming, with the goal of better understanding how different game elements are perceived by users and how they could impact users' perception of the applications.

3 Method and Results

For this work, the team first identified a list of gamified educational applications designed to teach Python programming available in the Google Play store and subsequently proceeded to use the Google Play API to collect users review both in English and Spanish. Table 1 shows the list of the applications analyzed in this work along with the average start rating and number of downloads . For each of the applications, the team identified and grouped the different game elements implemented, following Yu-kai Chou's framework [5]. A total of 19 different game elements were identified. The most common ones were Progress Bar and Points.

Table 1: Summary of Gamified Applications Analyzed

App name	Start Rating	Downloads [Millions]
Mimo	4.7	10
Sololearn	4.6	10
Programming Hub	4.7	5
Codecademy	4.7	1
DataCamp	4.7	1
Programming Hero	4.6	1
Learn Python	4.7	0.5

After the reviews were obtained, Sentiment Analysis was performed for both

English [1] and Spanish [23] reviews. Table 2 shows the summary statistics of the reviewers and the sentiment analysis. The "Proportion of Positive Reviews" (Pro. Pos. Reviews) was calculated based on the total percentage of reviews that were classified as positive (i.e., a sentiment value of 0 or greater based on the sentiment algorithms used).

Table 2: Summary statistics of reviews

Name	Lang.	No. Reviews	Avg. Rating	Avg. Review Length	Prop. Pos. Reviews
Mimo	EN	86724	4.55	59.22	0.86
	SP	9253	4.38	85.07	0.88
Sololearn	EN	130851	4.72	59.93	0.91
	SP	28876	4.75	77.07	0.95
Programming Hub	EN	48682	4.59	59.55	0.86
	SP	1633	4.26	90.69	0.88
Codecademy	EN	3100	3.85	70.24	0.66
	SP	141	4.04	78.04	0.77
DataCamp	EN	7378	4.50	69.90	0.84
	SP	389	4.66	87.12	0.91
Programming Hero	EN	16407	4.78	56.69	0.90
	SP	221	4.52	97.77	0.89
Learn Python	EN	4728	4.66	59.31	0.86
	SP	28	4.32	92.46	0.96

To explore if there was any correlation between the applications' star rating, number of downloads, number of reviews, avg. review rating, and proportion of positive reviews for both English and Spanish, a series of non-parametric Spearman correlation tests were performed. As expected, the results indicated a positive correlation between the total number of downloads and the total number of reviews per application for both English (ρ=0.88, p-value$<$0.01) and Spanish reviews (ρ=0.95 p-value$<$0.001). This could be attributed to the fact that the more downloads an application has, the more users it has reached, which could potentially post a review online. Similarly, there was a positive correlation between avg. review ratings and proportion of positive reviews per application, indicating that on average the more highly rated reviews an application had, the more positive the sentiment of the reviews was. However, the results only show a statistically significant correlation for English reviews (ρ=0.86 p-value$<$0.05), and not for the Spanish reviews (ρ=0.64 p-value$>$0.05). This might indicate a potential limitation of the Spanish Sentiment algorithms, or a potential disconnect between the rating and the reviews given by Spanish-speaking users (e.g., give high start rating review, but provided a review with

mostly negative sentiment).

Moreover, to explore if there was any difference between the number of reviews, avg. review rating, avg. review length, and proportion of positive reviews for both English and Spanish reviews, a series of non-parametric t-tests were performed. The results indicate that there were significantly more English reviews than Spanish reviews (μ_{EN}=42,552.8 vs μ_{SP}=5,791.6, p-value$<$0.05). This can be attributed to the fact that these applications are designed for English-speaking users (i.e., the text of the applications is all in English) and consequently, there are more English than Spanish-speaking users providing reviews. Nonetheless, the avg. rating between the English and Spanish reviews (μ_{EN}=4.52 vs μ_{SP}=4.42, p-value$>$0.05) and the proportion of positive reviews (μ_{EN}=0.84 vs μ_{SP}=0.87, p-value$>$0.05) were not significantly different. Only the avg. review length was significantly different between the English and Spanish reviews (μ_{EN}=62.12 vs μ_{SP}=86.89, p-value$>$0.001). This indicates that the Spanish reviews on average had similar ratings than the English reviews, but the Spanish reviews tended to be longer.

To better understand if users expressed any opinion about the game elements of the applications in their reviewers, the reviewers that mention a word related to a game element were identified. To achieve this, all the synonyms and stem words of the 19 game elements identified in all the applications were generated using ChatGPT [https://chat.openai.com/]. Similarly, the game elements were translated into Spanish and their synonyms and stem were also generated using ChatGPT. This list of words was used to filter any reviews containing them. For example, for the game elements of Leaderboard, words such as "leaderboard", "leader", "ranking", "rank", "scoreboard", and "board" were used for filtering the reviews.

Table 3 shows a summary statistic of the sentiment of the reviews based on the game elements. The table is sorted in descending order based on the difference between the number of positive and negative reviews. For example, the difference between the number of positive and negative reviews using words related to the game elements of Challenges was larger than any other game elements (i.e., 2357-604=1753). Moreover, the proportion of positive reviews is shown in Table 3 alongside the p-value of a proportion test for each game element (i.e., null hypothesis: proportion of positive reviews is equal to 0.5). The results from Table 3 indicate that users tended to post more positive than negative reviews about all the game elements except for the game element of Torture Breaks. This game element refers to the mechanics in which users are forced to take a "break" after taking an action. For example, in the free version of SoloLearn, after users make three consecutive mistakes in an assessment, they needed to way several hours to be able to try again.

Lastly, to gain a better understanding of the discourse and what aspects

Table 3: Game elements review sentiment

Game Elements	No. Pos. Reviews	No. Neg. Reviews	Prop. Pos. Reviews	p-value
Challenges	2357	604	0.796	0.001
Boosters	1884	247	0.884	0.001
Daily Streaks	987	178	0.847	0.001
Avatar	445	61	0.879	0.001
points system	284	50	0.850	0.001
Competition/Duel	247	68	0.784	0.001
Last mile drive	179	11	0.942	0.001
status quo sloth	1277	1109	0.535	0.001
Voting	249	107	0.699	0.001
Virtual economy	142	34	0.806	0.001
Leaderboard/Ranking	148	44	0.770	0.001
Visual Grave	121	20	0.858	0.001
Win States	95	24	0.798	0.001
Social Discovery	131	90	0.59	0.007
High Five	27	0	1	0.001
Progress Bar	39	21	0.65	0.028
Crowning	21	5	0.807	0.003
Anticipation parade	11	0	1	0.003
Torture Breaks	507	543	0.482	0.28

of the applications were discussed in the reviews, Semantic Network analyses were performed. For this, all the reviewers were preprocessed by removing URL, punctuation, and stop words (for both English and Spanish reviews), as well as converting to lowercase. After this processing step, all the bigrams (e.g., pair of consecutive words) were identified, and their frequency, as well as the average sentiment of the reviews in which they were present, were calculated.

Figure 1 and 2 shows the Semantic Network of the top 50 bigrams for all the English and Spanish reviews. The nodes of the networks represent the bigram words, while the color of the edges represent the average sentiment of the reviews in which the bigram was present. From these figures, and by looking at the top 5 most frequent "positive" bigrams in both English ('learn'-'coding', 'easy'-'learn', 'learn'-'programming', 'easy'-'use', 'way'-'learn'), and Spanish ('buena'-'aplicacion', 'excelente'-'aplicacion', 'aprender'-'programar', 'aplicacion'-'aprender', 'lenguajes'- 'programacion') it is clear users, both English and Spanish speaking, reviewed positively the fact that the applications are good for learning programming.

Similarly, by looking at the top 5 most frequent "negative" bigrams in both

English ('please'-'fix', 'please'-'make', 'waste'-'time', 'internet'-'connection', 'buy'-'pro'), and Spanish ('cambiar'-'idioma', 'conexion'-'internet', 'puede'-'cambiar', 'ingles'-'entiendo', 'si'- 'pagas') it is clear users, both English and Spanish speaking, reviewed negatively the fact that some applications need to have internet access to work correctly and that some have pro-version that would enable users to access more features (e.g., SoloLearn in the pro-version does not have the torture break game element). More interestingly, Spanish-speaking users negatively reviewed the fact that the applications are all in English, and the language cannot be changed.

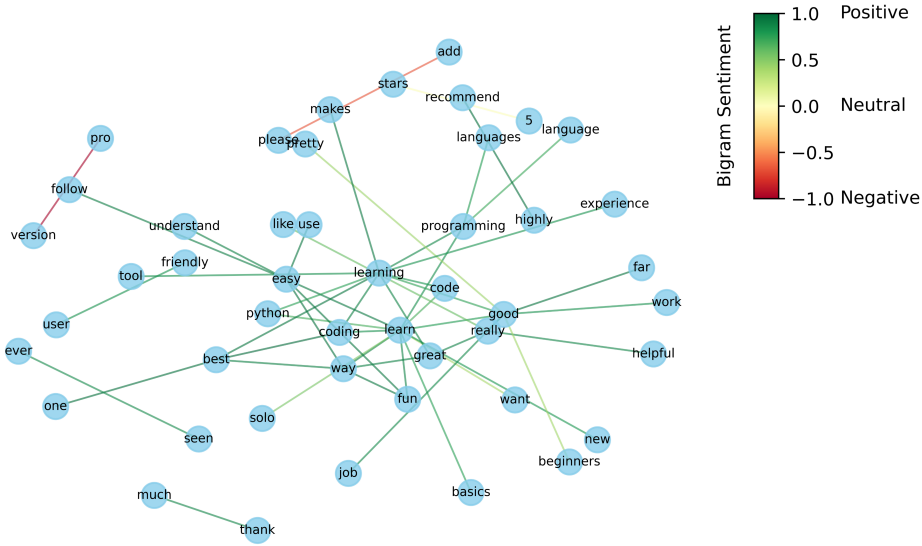

Figure 1: Semantic Network with Top 50 Bigrams of English reviews

In summary, all these findings indicate:

- On average users perceived these educational gamified applications as positive.
- On average users had a positive perception of the game elements implemented, except for the game elements of "Torture Breaks."
- Users indicated these are good applications for learning programming, but some areas of improvement still exist.
- The requirement to have an internet connection to function properly is an aspect of the application that users dislike

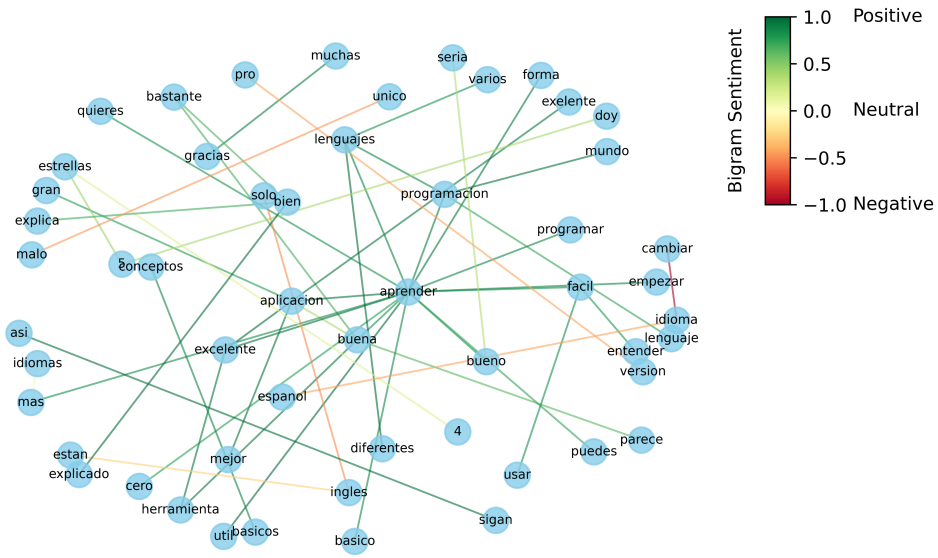

Figure 2: Semantic Network with Top 50 Bigrams of Spanish reviews

- Spanish-speaking users would prefer to have an application design for them, in which the content and text are not all in English.

4 Conclusion and Future Works

The findings of the study indicate that gamification elements, such as progress bars and point systems, are prevalent in these applications and align with previous research on gamification in educational contexts. However, concerns were raised regarding specific game elements, such as "Torture Breaks," and the requirement of internet access, which can limit accessibility for some users. The analysis of semantic networks provided valuable insights into the aspects of the applications discussed in the user reviews, highlighting positive sentiments towards learning programming and negative sentiments towards language limitations and internet requirements.

These findings emphasize the importance of addressing language preferences and ensuring language diversity in gamified educational applications to cater to a diverse user base. Hence, to enhance the effectiveness of gamified applications for programming education, future efforts should focus on addressing user concerns and improving the user experience. This includes refining game elements to ensure they align with users' preferences and removing potential

barriers like language limitations and the need for internet access.

This study provides insights into the implementation of game elements in popular gamified applications for learning Python. By exploring the elements that resonate with users and contribute to successful applications, educators and technologists can design more personalized and effective programming gamified educational applications, thereby advancing programming education in a rapidly digitizing world. Nevertheless, there exist several areas of improvement. For example, The Spanish sentiment algorithm used might not be designed to handle online user reviews. Similarly, the approach to filter reviews based on game elements related words might have introduced additional noise since some words might be used in different contexts. Hence, more advanced methods based on Machine Learning approaches (like Large Language Models) should be implemented to reduce false positives.

To enhance the effectiveness of gamified applications for programming education, future efforts should focus on addressing user concerns and improving the user experience. This includes refining game elements to ensure they align with users' preferences and removing potential barriers like language limitations and the need for internet access.

Acknowledgement

This research was funded by FONDOCYT 2022-3A1-112 and by the EXCEL scholar program at Lafayette College. Any opinions, findings, or conclusions found in this paper are those of the authors and do not necessarily reflect the views of the sponsors.

References

[1] Alan Akbik et al. "FLAIR: An easy-to-use framework for state-of-the-art NLP". In: *NAACL 2019, 2019 Annual Conference of the North American Chapter of the Association for Computational Linguistics (Demonstrations)*. 2019, pp. 54–59.

[2] Souha Bennani, Ahmed Maalel, and Henda Ben Ghezala. "Adaptive gamification in E-learning: A literature review and future challenges". In: *Computer Applications in Engineering Education* 30.2 (2022), pp. 628–642.

[3] Jens Bennedsen and Michael E Caspersen. "Failure rates in introductory programming: 12 years later". In: *ACM inroads* 10.2 (2019), pp. 30–36.

[4] Mihai Burlacu, Claudiu Coman, and Maria Cristina Bularca. "Blogged into the System: A Systematic Review of the Gamification in e-Learning before and during the COVID-19 Pandemic". In: *Sustainability* 15.8 (2023), p. 6476.

[5] Yu-kai Chou. *Actionable gamification: Beyond points, badges, and leaderboards*. Packt Publishing Ltd, 2019.

[6] Jacek Dąbrowski et al. "Analysing app reviews for software engineering: a systematic literature review". In: *Empirical Software Engineering* 27.2 (2022), p. 43.

[7] Jacek Dąbrowski et al. "Mining and searching app reviews for requirements engineering: Evaluation and replication studies". In: *Information Systems* 114 (2023), p. 102181.

[8] Edward L Deci and Richard M Ryan. "Self-determination theory." In: (2012).

[9] Yogesh K Dwivedi et al. "Setting the future of digital and social media marketing research: Perspectives and research propositions". In: *International Journal of Information Management* 59 (2021), p. 102168.

[10] Giovanni Grano et al. "Exploring the integration of user feedback in automated testing of android applications". In: *2018 IEEE 25Th international conference on software analysis, evolution and reengineering (SANER)*. IEEE. 2018, pp. 72–83.

[11] Yue Guo, Stuart J Barnes, and Qiong Jia. "Mining meaning from online ratings and reviews: Tourist satisfaction analysis using latent dirichlet allocation". In: *Tourism management* 59 (2017), pp. 467–483.

[12] Hyoji Ha et al. "An improved study of multilevel semantic network visualization for analyzing sentiment word of movie review data". In: *Applied Sciences* 9.12 (2019), p. 2419.

[13] Stuart Hallifax et al. "Adaptive gamification in education: A literature review of current trends and developments". In: *Transforming Learning with Meaningful Technologies: 14th European Conference on Technology Enhanced Learning, EC-TEL 2019, Delft, The Netherlands, September 16–19, 2019, Proceedings 14*. Springer. 2019, pp. 294–307.

[14] Hazra Imran. "An Empirical Investigation of the Different Levels of Gamification in an Introductory Programming Course". In: *Journal of Educational Computing Research* (2022), p. 07356331221144074.

[15] Zahra Jamebozorg and Seyyed Mehdi Sarkeshikiyan. "The effectiveness of Education design based on gamification on students' motivational strategies and social skills". In: *Rooyesh-e-Ravanshenasi Journal (RRJ)* 12.1 (2023), pp. 51–62.

[16] Michail Kalogiannakis, Stamatios Papadakis, and Alkinoos-Ioannis Zourmpakis. "Gamification in science education. A systematic review of the literature". In: *Education Sciences* 11.1 (2021), p. 22.

[17] Firas Layth Khaleel et al. "Programming learning requirements based on multi perspectives". In: *International Journal of Electrical and Computer Engineering* 7.3 (2017), p. 1299.

[18] Richard N Landers, Michael B Armstrong, and Andrew B Collmus. "How to use game elements to enhance learning: Applications of the theory of gamified learning". In: *Serious Games and Edutainment Applications: Volume II* (2017), pp. 457–483.

[19] Sachiko Lim, Aron Henriksson, and Jelena Zdravkovic. "Data-driven requirements elicitation: A systematic literature review". In: *SN Computer Science* 2 (2021), pp. 1–35.

[20] Katerina Mangaroska et al. "Exploring students' cognitive and affective states during problem solving through multimodal data: Lessons learned from a programming activity". In: *Journal of Computer Assisted Learning* 38.1 (2022), pp. 40–59.

[21] Dwi Maryono, Sajidan Budiyono, and Muhammad Akhyar. "Implementation of Gamification in Programming Learning: Literature Review". In: *Int. J. Inf. Educ. Technol* (2022).

[22] Behrooz Noori. "Classification of customer reviews using machine learning algorithms". In: *Applied Artificial Intelligence* 35.8 (2021), pp. 567–588.

[23] Juan Manuel Pérez, Juan Carlos Giudici, and Franco Luque. *pysentimiento: A Python Toolkit for Sentiment Analysis and SocialNLP tasks*. 2021. arXiv: `2106.09462 [cs.CL]`.

[24] Yizhou Qian and James Lehman. "Students' misconceptions and other difficulties in introductory programming: A literature review". In: *ACM Transactions on Computing Education (TOCE)* 18.1 (2017), pp. 1–24.

[25] Rodrigo Romero-Silva and Sander De Leeuw. "Learning from the past to shape the future: a comprehensive text mining analysis of OR/MS reviews". In: *Omega* 100 (2021), p. 102388.

[26] AS Saabith, MMM Fareez, and T Vinothraj. "Python current trend applications-an overview". In: *International Journal of Advance Engineering and Research Development* 6.10 (2019).

[27] Laila Al-Safoury, Akram Salah, and Soha Makady. "Integrating User Reviews and Issue Reports of Mobile Apps for Change Requests Detection". In: *International Journal of Advanced Computer Science and Applications* 13.12 (2022).

[28] Zaher Salah et al. “A systematic review on opinion mining and sentiment analysis in social media”. In: *International Journal of Business Information Systems* 31.4 (2019), pp. 530–554.

[29] K. Schwab and S. Zahidi. *“The Future of Jobs Report 2020,”*. `https://www.weforum.org/reports/the-future-of-jobs-report-2020`.

[30] Mahwish Shahid et al. “A review of gamification for learning programming fundamental”. In: *2019 International Conference on Innovative Computing (ICIC)*. IEEE. 2019, pp. 1–8.

[31] Tze Ying Sim and Sian Lun Lau. “Review on Challenges and Solutions in Novice Programming Education”. In: *2022 IEEE International Conference on Computing (ICOCO)*. IEEE. 2022, pp. 55–61.

[32] Itamar Simonson. “Imperfect progress: An objective quality assessment of the role of user reviews in consumer decision making, a commentary on de Langhe, Fernbach, and Lichtenstein”. In: *Journal of Consumer Research* 42.6 (2016), pp. 840–845.

[33] Jallal Talaghzi et al. “Online adaptive learning: a review of literature”. In: *Proceedings of the 13th International Conference on Intelligent Systems: Theories and Applications*. 2020, pp. 1–6.

[34] Gustavo F Tondello, Alberto Mora, and Lennart E Nacke. “Elements of gameful design emerging from user preferences”. In: *Proceedings of the annual symposium on computer-human interaction in play*. 2017, pp. 129–142.

[35] Marcello Trovati and Suleman Awan. “Semantic Network: A Brief Review of its Datasets”. In: *Advances in Intelligent Networking and Collaborative Systems: The 14th International Conference on Intelligent Networking and Collaborative Systems (INCoS-2022)*. Springer. 2022, pp. 224–233.

[36] Wenxin Wang, Yi Feng, and Wenqiang Dai. “Topic analysis of online reviews for two competitive products using latent Dirichlet allocation”. In: *Electronic Commerce Research and Applications* 29 (2018), pp. 142–156.

[37] Zamzami Zainuddin et al. “The impact of gamification on learning and instruction: A systematic review of empirical evidence”. In: *Educational Research Review* 30 (2020), p. 100326.

[38] Zamzami Zainuddin et al. “The impact of gamification on learning and instruction: A systematic review of empirical evidence”. In: *Educational Research Review* 30 (2020), p. 100326.