Journal of Advanced Artificial Intelligence
Volume 1 – No.1, October 2024

Predicting Level of Computer Science Education Students' Engagement in Programming Activities and Hackathon using Random Forest Supervised Machine Learning

Ajayi Olufunke Esther
Institute of Part-Time Programme,
Adekunle Ajasin University,
Akungba-Akoko

ABSTRACT
The study uses a Random Forest supervised machine learning model to predict computer science education students' level of participation in programming activities and hackathons. To categorize student engagement, 310 students' data were reviewed. Variables such as the amount of time spent programming, participation in online forums, test results, and hackathon attendance were all considered. With an R-squared value of -0.0014 and a Mean Squared Error (MSE) of 8.64, the model's prediction accuracy was found to be lacking, indicating the need for more varied data sources. The main conclusions revealed a moderate association between online engagement and programming time, and poor correlations between test results, programming time, and hackathon participation. The study underlines the need of adding psychological and social elements into future models and advocates broader integration of hackathons into the curriculum to boost student participation. The findings draw attention to the shortcomings of the current predictive methodology and suggest investigating more factors in order to raise the accuracy of the model.

Keywords
Computer Science Education, Student Engagement, Programming Activities, Hackathon, Machine Learning, Random Forest, SHAP, Predictive Model

1. INTRODUCTION
In computer science education, student engagement in practical activities such as hackathons and programming is critical for fostering computational thinking, problem-solving skills, and collaborative learning methods. However, not every student participates to the same extent, which might affect their academic achievement and readiness for the demands of the workplace. With the increasing availability of educational data, predictive models can provide early interventions to enhance engagement as well as insights into student behavior.

Hackathons have grown in popularity as dynamic, immersive events in which students work to build new solutions within a 24- to 48-hour timeframe. These events require students to collaborate in diverse teams to tackle real-world problems using industry-standard technologies and processes, such as Agile (Briscoe & Mulligan, 2014). Hackathons promote creativity, quick prototyping, and teamwork while also offering significant networking chances with professionals and mentors from other industries. They provide students with practical experience and exposure to potential employers, which improves their employability (Komssi, Pichlis, Raatikainen, Kindstrom, & Jarvinen, 2015).

In today's fast changing technology landscape, programming expertise has become critical for success in a variety of academic and professional settings. Educational institutions and organizations are increasingly incorporating programming activities and hackathons into their curricula to nurture and enhance these abilities. These initiatives offer students hands-on experiences that promote creativity, teamwork, and practical problem-solving skills. Participating in these activities allows students to build technical expertise, computational thinking, and critical soft skills, all of which are essential for preparing them for the future workforce (Grover & Pea, 2013) (Resnick, 2017)

Programming activities such as coding challenges, workshops, and project-based learning are critical for providing students with the technical and analytical abilities they need to succeed. These activities allow students to obtain hands-on experience with programming languages such as Python, JavaScript, Java, and C++ while also improving their problem-solving skills and algorithmic thinking (Hsu, Chang, & Hung, 2018). Furthermore, these activities promote soft skills such as creativity, logical thinking, and collaboration. Students also create a portfolio of projects to exhibit their skills to potential employers, bridging the gap between theoretical understanding and practical application (Wing, 2006).

Despite the various benefits, obstacles such as a lack of knowledge, time constraints, and limited access to resources can prevent students from participating in programming activities and hackathons (Fredricks, Blumenfeld, & Paris, School engagement: Potential of the concept, state of the evidence. ,, 2004). For newcomers, the intricacy of programming languages and the rapid pace of hackathons might be overwhelming. Furthermore, students frequently struggle to balance hackathon participation and academic obligations. Access to crucial resources, such as laptops and software licenses, is critical to inclusion. Addressing these problems while using machine learning approaches such as Random Forests might assist educators forecast and improve student involvement, allowing them to adapt interventions for more participation (Breiman, 2001) (Baek, Lee, & Yoon, 2021).

The goal of this study is to use machine learning to forecast how involved computer science education students will be in programming and hackathons. The research employs random forest machine learning to predict computer science students' interest in programming activities and hackathons. This adds to the current research by using a quantitative approach to predicting student engagement, complementing earlier studies that have mostly focused on theoretical and qualitative insights. Unlike typical studies that focus on engagement frameworks (Kuh, 2008) Fredricks et al., 2016), this research work adds a data-driven way for analyzing engagement, utilizing machine learning to model complicated links between student activities and engagement results. This technique provides educators and institutions with concrete knowledge into how to better adapt interventions and increase involvement in hackathons and programming events.

2. RELATED WORK

(Kuh, 2008), argues that student involvement is critical to first-year student success because it leads to improved academic achievement and retention rates. Institutions play an important role in offering support systems, high-impact instructional practices, and chances for students to engage meaningfully with teachers and peers. Colleges may assist students overcome first-year problems and set them on a road to graduation by creating an engaging environment.

Lawson and Lawson (2013) present a theory that sees student engagement as a dynamic, multidimensional phenomenon influenced by the interaction of personal, societal, and institutional elements. They advocate for a **holistic approach** to engagement, which combines supporting policies, practical tactics, and a better understanding of the student's social environment. Their strategy calls on schools to collaborate closely with families and communities to create environments that not only promote academic accomplishment but also improve overall student well-being and participation.

The critical function of Learning Management Systems in increasing student engagement through controlled online interactions was investigated by (Norris & Coutas, 2014). They argued that when LMS platforms are effectively incorporated into the educational process, they promote a more dynamic, engaging, and active learning environment, resulting in improved educational outcomes. However, the study highlights problems associated with technology use and advocates for improved support mechanisms to overcome these barriers.

(Henrie, Halverson, & Graham, 2015), discovered that when used properly, technology can dramatically increase student engagement, particularly through interactive platforms and educational tools. Their review found positive links with behavioral, emotional, and cognitive engagement, but they also identified issues connected to unequal access and inappropriate use of technology. Fredricks et al. (2016) developed a comprehensive model for analyzing student engagement in terms of its behavioral, emotional, and cognitive components. The study highlights the importance of motivation, autonomy, and goal-directed behaviors in encouraging meaningful student participation and achievement in educational environments by integrating self-determination and goal-setting theories. This multifaceted approach provides educators with insight into how to create greater involvement by meeting learners' emotional and cognitive demands while also encouraging good actions.

(Schindler, Burkholder, Morad, & Marsh, 2017), investigate the effects of digital tools including games, social media, and web-conferencing on student engagement, particularly in remote learning settings. Digital games encourage critical thinking and problem solving, social media promotes cooperation and information sharing, and web-conferencing technologies replicate real-time interactions to help maintain a classroom-like setting. These technologies are critical for keeping students engaged in remote learning, but their usefulness is dependent on how well they are integrated into the learning design. The study underlines the importance of using technology to complement educational aims in order to truly improve learning results.

A comprehensive evaluation of educational technology's function in increasing cognitive and behavioral engagement in undergraduate STEM education was given by (Webb, Gibson, & Forkosh-Baruch, 2017). The study discovers that, when used correctly, technology can enhance deep learning and active involvement. However, concerns like as guaranteeing equal access to technology and providing enough educator support must be addressed. The authors recommend more research on the various effects of technology on different learners and fields in STEM. (Kaliisa & Picard, 2017), investigated the function of mobile learning in increasing student engagement and discovered that it works best when combined with student-centered instructional practices. Mobile technology' flexibility and accessibility, along with learner-centered education, help to boost engagement and learning results. However, problems such as technology access and proper implementation must be addressed in order to achieve broad success.

Sociocultural model of student engagement provides a comprehensive framework that emphasizes the importance of both individual and institutional elements in determining student involvement, especially during the key transition to university life. The approach gives a sophisticated view of how involvement is fostered or hindered in higher education settings, taking into account human motivations, institutional practices, and sociocultural circumstances (Kahu & Nelson, 2018).

(Picton, Clark, & Judge, 2018), suggest that socio-emotional factors such as belonging and identification with the learning environment are important drivers of student engagement. They emphasize that students' sense of belonging, identification with the learning environment, and strong relationships with classmates and teachers all have a substantial impact on their motivation, involvement, and academic performance. As a result, schools and instructors should prioritize providing inclusive, supportive, and emotionally safe settings in order to improve students' academic and personal success.

(Hidi, Renninger, & Krapp, 2019), argued that self-relatedness—the degree to which tasks fit with an individual's sense of self - is an important driver of motivation and engagement. Their findings connect cognitive and emotional processes to brain mechanisms, demonstrating how self-relatedness promotes persistent engagement and motivated behavior. This paradigm emphasizes the significance of personal relevance in promoting long-term motivation and goal-directed behavior.

(Minkos & & Gelbar, 2020), investigated the significance of social-emotional learning and adaptable teaching strategies in mitigating the negative impacts of the COVID-19 pandemic on student participation. They advocate for a comprehensive strategy that prioritizes emotional well-being, promotes strong connections, and caters to each student's unique learning needs. The authors emphasize that by incorporating SEL and utilizing flexible teaching strategies, educators may help kids recover from the pandemic's interruptions and thrive in the new educational environment.

(Ndofirepi, 2020) examined the inequities that exist in African educational systems, focusing on underprivileged groups such as rural pupils, refugees, and impaired learners. They contend that these inequities are profoundly ingrained in colonial histories that molded educational frameworks across the continent. The authors argue that in order to attain social justice in education, inclusive practices and indigenous knowledge systems should be integrated. Their work is a call to action for African higher education institutions to not only give access to education, but also actively engage and integrate different student populations in learning environments. They argue that this method would help bridge the gap between past exclusions and current educational disparities. (Agai, 2020) offered a thorough examination of the historical and cultural underpinnings of education in West and North Africa. He examines the impact of Islam and colonialism on educational systems, highlighting the enormous epistemic alterations that occurred over time. Agai's research focuses on the relationship between ancient knowledge systems and the advent of Islamic learning, which eventually collided with European colonial educational paradigms. This study provides important insights into how historical legacies continue to impact contemporary African educational practices, notably in terms of knowledge dissemination and learning.

Fredricks, Blumenfeld, and Paris (2020), in *"Student Engagement: Potential of the Concept, State of the Evidence,"* provide a detailed examination of the different dimensions of student engagement and how it is conceptualized and measured across various educational settings. Although their study does not focus specifically on Africa, their emphasis on culturally responsive pedagogies and engagement strategies is relevant to addressing educational challenges in African contexts. By adapting their framework, African educators could enhance student engagement, particularly in areas where engagement is key to improving educational outcomes.

(Dinsmore & Ertmer, 2021), in their study published in *Educational Psychology Review*, explore the cognitive, motivational, and emotional factors influencing student engagement, offering insights relevant to global education systems, including Africa. They emphasize the importance of tailoring educational strategies to specific social and cultural contexts, acknowledging that historical and cultural factors significantly shape educational environments in Africa. Their theoretical framework provides educators with context-sensitive methods to enhance student engagement, making it highly applicable in diverse educational settings, particularly where localized approaches are needed to foster engagement.

(Cicha, Rizun, Rutecka, & Strzelecki, 2021), provided a complex view of how the COVID-19 pandemic impacted student participation. The switch to distant learning resulted in a mixed academic experience, with the flexibility of online learning balanced against a loss of motivation, focus, and direct social engagement. While technology constraints and a lack of preparation hampered learning, the emotional toll of isolation exacerbated the capacity to remain engaged intellectually and socially. The study underlines the importance of improved support systems in remote learning environments to promote both academic performance and social well-being.

(Gurcan, Ozturk, & Topuz, 2021), conducted a comprehensive review of the emerging trends in e-learning and student involvement, revealing a significant increase in research on the use of technology in education. They discovered rising trends using topic modeling, such as the utilization of interactive tools and data-driven engagement techniques. While technology shows promise for increasing involvement, the study also recognized the need to address issues such as equal access and content quality. (Cagiltay, Karakus, & & Ercan, 2021), underscore the critical role of interactive and adaptive learning technologies in boosting student engagement in online education. By tailoring content and providing immersive experiences, these tools create a more engaging learning environment. However, successful implementation requires overcoming technological, financial, and educational barriers. The research highlights the importance of continued investment and innovation in educational technology to sustain engagement in digital learning environments.

The methodological approach and implementation of this study differ from that of previous research. While earlier research has mostly focused on the theoretical, behavioral, cognitive, and socio-cultural components of student involvement (Kuh, 2008) (Lawson & Lawson, New conceptual frameworks for student engagement research: policy, and practice, 2013) 6), this study provides a quantitative, predictive model based on random forest machine learning. This transition from descriptive and explanatory frameworks to predictive, data-driven approaches is essential because it allows for real-time or near-real-time analysis of student participation in programming events and hackathons. Unlike typical research that look into engagement aspects, this work provides a tool for predicting engagement levels, allowing educators to make proactive adjustments.

3. METHODOLOGY

The study adopted a descriptive survey research design, which was appropriate for gathering detailed information on the engagement levels of computer science students in programming activities and hackathons. This design enabled the researcher to observe, describe, and analyze trends and relationships without manipulating variables. To further unravel the impact of participation in programming activities and hackathon on the performance of computer science education students, the study deploys random forest supervised machine learning approach to predict the level of participation of the students in the programming events and exercises.

3.1 Data Collection

The data which were retrieved from the student results database of the Department of Computer Science, Adekunle Ajasin University (Table 1 in Appendix I), have the following features:

i. **p_timetaken** (time spent on programming activities)
ii. **no_online_participate** (number of online activities)
iii. **hackathon_participation** (whether a student participated in hackathons)
iv. **test_score** (academic performance in relevant subject

Population and Sample

The population for this study comprised 310 students enrolled in the Computer Science Education program within the Department of Science Education at Adekunle Ajasin University, Akungba-Akoko (AAUA), spanning the 2016/2017 to 2022/2023 academic sessions. The students were purposefully selected from the 200 to 400 level classes because programming-based courses are core components of the curriculum at these levels. This ensures that participants have substantial exposure to programming, which is crucial for the objectives of the study.

A purposive sampling technique was adopted to ensure that only students who had completed mandatory programming courses were included. This method was employed to increase the relevance of the study's findings by focusing on a sample of students who have been directly engaged with programming coursework. By selecting this subset of students, the study aims to enhance the validity of its conclusions regarding students' programming engagement and participation in hackathons.

Instruments for Data Collection

Data were gathered by assessing students' participation in programming activities and hackathons, specifically as they relate to courses taught during the academic year. The study also examined the impact of these activities on students' academic performance, measured through their test scores. The data collection captured basic demographic information, including students' academic levels and the specific programming courses they completed at each level. The number of participants from each level was proportional to the total number of students enrolled in programming courses, ensuring that the sample accurately represented the distribution of students with relevant experience across different levels.

3.2 Data Preprocessing

The collected data were in normalized state with no missing data. The data were splitted into 80:20 percent for the training and test sets. The student engagement were defined such that the number of times a student participate in hackathon or online programming is expected to be greater or equal to 2, while the number of hours spent on practical/programming activities or hackathon is expected to be greater than or equal to 120minutes (2hours). Random forest supervised machine learning model was preferred for this analysis/computation because of its capabilities to handle extensive and intricate dataset. To however analyze the level of accuracy of this model, the study make use of r-square value and mean-square error as model performance metrics, while heatmap, boxplot, and scatterplot were used for the visualization of the result. Future prediction result were also computed using the formula: future_predictions = model.predict(future_data) (Appendix II - Figure 2).

4. RESULTS

With Mean Squared Error of 8.63683254438767, and R-squared Value of -0.0014195990252818813, the evaluation metrics indicate that the regression model is not performing well. The Mean Squared Error (MSE) of 8.64 suggests that there is a significant difference between the predicted and actual values. More concerning is the R-squared value of -0.0014, which means the model is performing worse than a simple baseline that predicts the mean of the target variable. This suggests the model is either underfitting or has issues with feature selection or data preprocessing, requiring further refinement to improve performance.

Also, Figure 1 (the heatmap), illustrates the correlations between four features: **test_score**, **p_tt**, **no_online_participate**, and **hackathon_participation**. Most of the relationships between the variables show weak correlations. The strongest correlation is a moderate positive relationship between **p_tt** and **no_online_participate** (0.48), suggesting that higher participation in one may be linked to the other. Other correlations, such as between **test_score** and **hackathon_participation** (0.033), and **test_score** and **p_tt** (-0.18), are very weak, indicating minimal linear relationships. Overall, the features appear to be largely independent of each other.

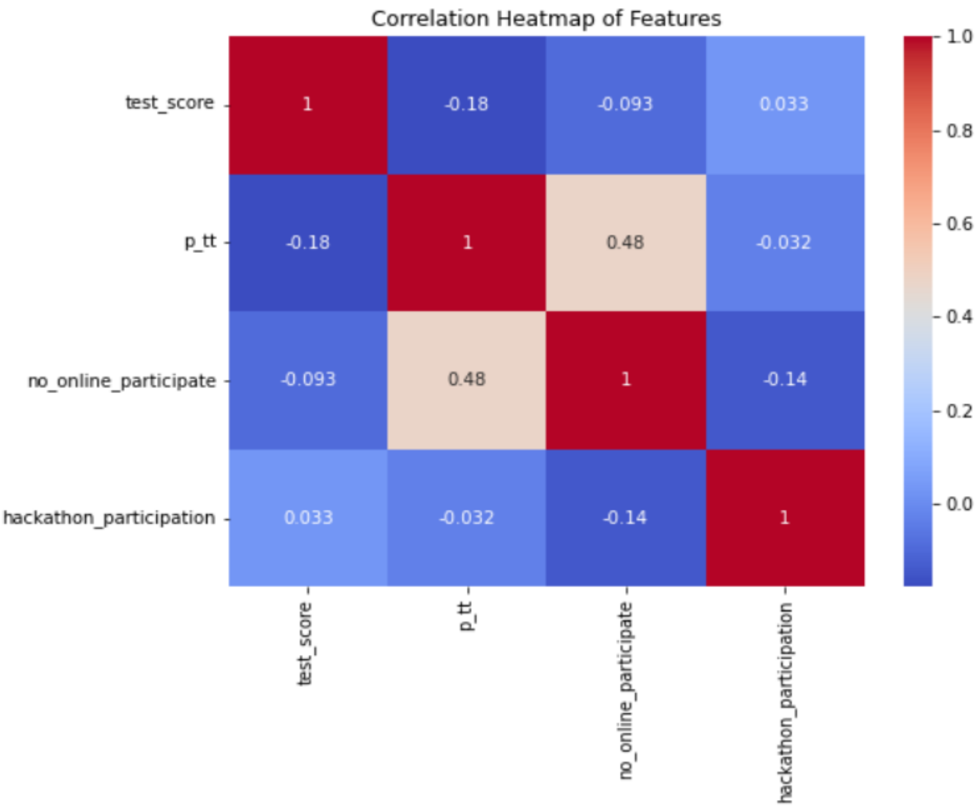

Figure 1: Heatmap showing correlation of features/variables

Findings

The followings findings could be deduced from the results:

i. In terms of accuracy of engagement prediction, the Random Forest model, which had an R-squared value of -0.0014 and a Mean Squared Error (MSE) of 8.64, performed poorly in predicting student participation in programming events and hackathons. This suggests that the model had difficulty predicting student involvement levels with any degree of accuracy.

ii. Weak connections were discovered by the study between important factors, including exam results, programming time, and hackathon participation. The variables that showed the greatest association (0.48) were programming time and online involvement, indicating a modest relationship between them.

iii. The majority of characteristics, such as involvement in hackathons and test scores, had little effect on predicting engagement. This may indicate the need for more complicated features than those included in the model at this time.

iv. Students who programmed more or attended hackathons more regularly were more likely to take part in other online activities. However, there was no significant correlation found between the total amount of time students spent programming and participating in hackathons and their academic achievement.

Implications of Findings

The followings are the implications of the findings:

i. Given the Random Forest model's poor prediction power, future models should incorporate more variables to better understand student engagement, such as psychological indicators, social interaction data, or other student behaviors.

ii. Intervention Strategies: Academic performance may not be a reliable indicator of programming engagement due to the limited link found between test scores and hackathon participation. This means that educators should consider other elements, such motivation or resource access, when creating interventions to promote engagement.

iii. The weak link between programming activities and academic outcomes shows that involvement in hands-on, collaborative events like hackathons would require deeper integration into the curriculum to maximize their impact on learning outcomes.

5. CONCLUSION

This study showed how machine learning may be used to forecast students' participation in computer science-related events like hackathons and programming. However, the findings imply that the existing model has limitations due to poor feature correlations and low prediction accuracy. In order to enhance the prediction of student engagement in educational activities, the study highlights the significance of investigating more robust data aspects. Furthermore, even though experiential programming activities are beneficial in the real world, it is still unknown how they affect students' academic achievement and needs more research.

6. RECOMMENDATION AND SUGGESTION FOR FURTHER RESEARCH

In order to improve prediction accuracy, future research should focus on improving the Random Forest model by including varied variables, such as psychological aspects (e.g., motivation, self-efficacy) and social interaction. Institutions must also provide more comprehensive assistance by addressing hurdles like resource access and academic scheduling problems to enable greater student participation in programming activities. Furthermore, a deeper integration of hackathons and programming events into the computer science curriculum will enhance their alignment with academic objectives and enhance their impact on student learning outcomes.

To further understand programming engagement, future research should look at psychological and social data by analyzing elements like student motivation, peer influence, and self-efficacy. Furthermore, longitudinal research will make it possible to monitor student performance and engagement over time, offering important insights into how these variables change and affect academic and professional results. In addition, it is critical to look into how technological availability and resources, especially in settings with limited resources, affect students' involvement in hackathons and programming activities.

When combined, these methods will provide a more thorough understanding of the factors that influence students' interest in computer science education.

7. REFERENCES

[1] Baek, Y., Lee, S., & Yoon, S. (2021). Machine learning approaches to predict student engagement in online learning environments. Sustainability, 13 (14), 7833.

[2] Breiman, L. (2001). Random forests. Machine Learning, 45 (1), 5-32.

[3] Briscoe, G., & Mulligan, C. (2014). Digital innovation: The hackathon phenomenon. Creativeworks London.

[4] Fredricks, J. A., Blumenfeld, P. C., & Paris, A. H. (2004). School engagement: Potential of the concept, state of the evidence. ,. Review of Educational Research, 74 (1), 59-109.

[5] Fredricks, J. A., Filsecker, M., & Lawson, M. A. (2016). Student engagement, context, and adjustment: Addressing definitional, measurement, and methodological issues. Learning and Instruction, 43, 1-4. doi:https://doi.org/10.1016/j.learninstruc.2016.02.00

[6] Grover, S., & Pea, R. (2013). Computational thinking in K–12: A review of the state of the field. Educational Researcher, 42 (1), 38-43.

[7] Henrie, C. R., Halverson, L. R., & Graham, C. R. (2015). Measuring student engagement in technology-mediated learning: A review. . Computers & Education, 90, 36-53. doi:https://doi.org/10.1016/j.compedu.2015.09.005

[8] Hidi, S., Renninger, K. A., & Krapp, A. (2019). The role of interest in learning and development. Cambridge: Cambridge University Press. doi:https://doi.org/10.10

[9] Hsu, T. C., Chang, S. C., & Hung, Y. T. (2018). How to learn and how to teach computational thinking:

Suggestions based on a review of the literature. Computers & Education, 126, 296-310.

[10] Kahu, E. R., & Nelson, K. (2018). Student engagement in the educational interface: Understanding the mechanisms of student success. Higher Education Research & Development, 37 (1), 58-71. doi:https://doi.org/10.1080/07294360.2017.1344197

[11] Kaliisa, R., & Picard, M. (2017). A systematic review on mobile learning in higher education: The African perspective. The Turkish Online Journal of Educational Technology, 16 (1), 1-18.

[12] Komssi, M., Pichlis, D., Raatikainen, M., Kindstrom, K., & Jarvinen, J. (2015). What are hackathons for? IEEE Software, 32(5), 60-67.

[13] Kuh, G. D. (2008). High-Impact Educational Practices: What They Are, Who Has Access to Them, and Why They Matter. Washington, D.C: Association of American Colleges and Universities.

[14] Lawson, M. A., & Lawson, H. A. (2013). New conceptual frameworks for student engagement research, policy, and practice. Review of Educational Research, 83 (3), 432-479. doi:https://doi.org/10.3102/0034654313480891

[15] Lawson, M. A., & Lawson, H. A. (2013). New conceptual frameworks for student engagement research, policy, and practice. . Review of Educational Research, 83 (3), 432-479. doi:https://doi.org/10.3102/0034654313480891

[16] Norris, C., & Coutas, R. (2014). Building student engagement through Learning Management Systems. Journal of Educational Technology, 25 (2), 127-142.

[17] Picton, I., Clark, C., & Judge, C. (2018). Teachers' and learners' perceptions of engagement and motivation in literacy learning. Journal of Educational Research, 61 (3), 156-169.

[18] Resnick, M. (2017). Lifelong Kindergarten: Cultivating Creativity through Projects, Passion, Peers, and Play. MIT Press.

[19] Schindler, L. A., Burkholder, G. J., Morad, O. A., & Marsh, C. (2017). Computer-based technology and student engagement: A critical review of the literature. International Journal of Educational Technology in Higher Education, 14 (1), 1-28. doi:https://doi.org/10.1186/s41239-017-0063-0

[20] Webb, M., Gibson, D., & Forkosh-Baruch, A. (2017). Challenges for IT-supported formative assessment of complex 21st-century skills. Technology, Pedagogy and Education, 26 (2), 135-146. doi:https://doi.org/10.1080/1475939X.2016.1164007

[21] Wing, J. M. (2006). Computational thinking. Communications of the ACM, 49 (3), 33-35.

APPENDIX I

Table 1: Student score sheets

s/n	course_code	session	matric_no	test_score	exam_score	p_tt	no_online_participate	hackathon_participation
1	csc212	2016_2017	120213012	25	33	120	2	1
2	csc212	2016_2017	130213001	24	19	111	1	1
3	csc212	2016_2017	130213012	23	25	90	3	1
4	csc212	2016_2017	130213014	25	32	87	4	1
5	csc212	2016_2017	140213002	23	35	120	3	0
6	csc212	2016_2017	140213004	24	31	120	3	0
7	csc212	2016_2017	140213011	25	22	120	3	0
8	csc212	2016_2017	140213012	24	17	120	2	0
9	csc212	2016_2017	140213016	24	30	110	3	0
10	csc212	2016_2017	140213017	25	35	109	3	1
11	csc212	2016_2017	140213018	24	38	108	4	1
12	csc212	2016_2017	140213019	25	33	109	1	1
13	csc212	2016_2017	140213021	23	20	107	2	1
14	csc212	2016_2017	140213022	24	38	105	2	1
15	csc212	2016_2017	140213023	23	29	112	3	1
16	csc212	2016_2017	150213003	23	23	120	4	0
17	csc212	2016_2017	150213004	25	21	130	4	0
18	csc212	2016_2017	150213006	24	20	111	4	0
19	csc212	2016_2017	150213008	23	33	120	4	0
20	csc212	2016_2017	150213009	24	18	114	3	1
21	csc212	2016_2017	150213010	20	16	117	3	1

22	csc212	2016_2017	150213011	22	34	116	3	1
23	csc212	2016_2017	150213012	22	13	115	3	1
24	csc212	2016_2017	150213013	25	17	87	4	1
25	csc212	2016_2017	160213001	23	18	120	3	0
26	csc212	2016_2017	160213015	24	33	120	3	0
27	csc212	2016_2017	160213023	25	16	120	3	0
28	csc212	2016_2017	160213033	24	36	120	2	0
29	csc305	2016_2017	120213021	18	23	110	3	0
30	csc305	2016_2017	130213011	21	19	109	3	1
31	csc305	2016_2017	130213016	19	22	108	4	1
32	csc305	2016_2017	140213001	19	39	109	1	1
33	csc305	2016_2017	140213002	18	40	107	2	1
34	csc305	2016_2017	140213004	17	30	105	2	1
35	csc305	2016_2017	140213005	18	33	112	3	1
36	csc305	2016_2017	140213006	17	36	120	4	0
37	csc305	2016_2017	140213007	15	0	130	4	0
38	csc305	2016_2017	140213010	19	45	87	4	1
39	csc305	2016_2017	140213011	20	36	120	3	0
40	csc305	2016_2017	140213012	18	24	120	3	0
41	csc305	2016_2017	140213013	15	2	120	3	0
42	csc305	2016_2017	140213016	18	40	120	2	0
43	csc305	2016_2017	140213017	17	37	110	3	0
44	csc305	2016_2017	140213018	18	40	109	3	1
45	csc305	2016_2017	140213019	18	25	108	4	1
46	csc305	2016_2017	140213021	19	25	109	1	1
47	csc305	2016_2017	140213022	21	47	107	2	1
48	csc305	2016_2017	140213023	17	24	105	2	1
49	csc305	2016_2017	140213024	20	38	112	3	1
50	csc305	2016_2017	140213025	17	31	120	4	0
51	csc305	2016_2017	150213001	17	23	130	4	0
52	csc305	2016_2017	150213006	19	31	111	4	0
53	csc305	2016_2017	150213013	21	29	87	4	1
54	csc212	2017_2018	150213012	28	28	120	3	0
55	csc212	2017_2018	160213002	26	24	120	3	0
56	csc212	2017_2018	160213003	21	11	120	3	0
57	csc212	2017_2018	160213004	27	26	120	2	0
58	csc212	2017_2018	160213005	25	32	110	3	0
59	csc212	2017_2018	160213006	24	16	109	3	1
60	csc212	2017_2018	160213007	26	39	108	4	1
61	csc212	2017_2018	160213008	25	18	109	1	1
62	csc212	2017_2018	160213009	28	29	107	2	1
63	csc212	2017_2018	160213010	28	13	105	2	1
64	csc212	2017_2018	160213011	27	25	112	3	1
65	csc212	2017_2018	160213012	24	18	120	4	0
66	csc212	2017_2018	160213013	25	31	130	4	0
67	csc212	2017_2018	160213014	28	37	87	4	1

68	csc212	2017_2018	160213016	27	24	120	3	0
69	csc212	2017_2018	160213017	26	26	120	3	0
70	csc212	2017_2018	160213018	26	14	120	3	0
71	csc212	2017_2018	160213019	27	14	120	2	0
72	csc212	2017_2018	160213020	27	13	110	3	0
73	csc212	2017_2018	160213021	27	23	109	3	1
74	csc212	2017_2018	160213022	28	23	108	4	1
75	csc212	2017_2018	160213024	26	14	109	1	1
76	csc212	2017_2018	160213025	28	23	107	2	1
77	csc212	2017_2018	160213026	28	15	105	2	1
78	csc212	2017_2018	160213027	19	11	112	3	1
79	csc212	2017_2018	160213028	27	44	120	4	0
80	csc212	2017_2018	160213029	24	13	130	4	0
81	csc212	2017_2018	160213030	25	18	87	4	1
82	csc212	2017_2018	160213031	28	16	120	3	0
83	csc212	2017_2018	160213032	27	24	120	3	0
84	csc212	2017_2018	160213034	25	13	120	3	0
85	csc212	2017_2018	160213036	28	16	120	2	0
86	csc212	2017_2018	160213037	20	11	110	3	0
87	csc212	2017_2018	160213038	21	11	109	3	1
88	csc212	2017_2018	160213039	28	12	108	4	1
89	csc212	2017_2018	160213040	28	15	109	1	1
90	csc212	2017_2018	160213042	27	26	107	2	1
91	csc212	2017_2018	160213043	24	46	105	2	1
92	csc212	2017_2018	160213044	25	45	112	3	1
93	csc212	2017_2018	160213045	28	23	120	4	0
94	csc212	2017_2018	160213046	27	11	130	4	0
95	csc212	2017_2018	160213047	26	19	111	4	0
96	csc212	2017_2018	160213049	27	13	120	4	0
97	csc212	2017_2018	160213050	20	11	114	3	1
98	csc212	2017_2018	160213052	27	25	87	4	1
99	csc212	2017_2018	160216012	27	24	120	3	0
100	csc212	2017_2018	170213006	27	31	120	3	0
101	csc212	2017_2018	170213007	26	18	120	3	0
102	csc212	2017_2018	170213010	25	25	120	2	0
103	csc212	2017_2018	170213011	28	29	110	3	0
104	csc212	2017_2018	170213013	25	39	109	3	1
105	csc212	2017_2018	170213014	28	22	108	4	1
106	csc212	2017_2018	170213015	27	35	109	1	1
107	csc212	2017_2018	170213018	26	29	107	2	1
108	csc212	2017_2018	170213021	26	20	105	2	1

109	csc212	2018_2019	170213022	25	26	112	3	1
110	csc212	2018_2019	170213009	24	26	120	4	0
111	csc212	2018_2019	170213020	21	22	130	4	0
112	csc212	2018_2019	170213017	19	13	111	4	0
113	csc212	2018_2019	170213019	26	35	120	4	0
114	csc212	2018_2019	170213004	22	33	114	3	1
115	csc212	2018_2019	170213023	23	22	87	4	1
116	csc212	2018_2019	170213005	26	35	120	3	0
117	csc212	2018_2019	170213002	23	37	120	3	0
118	csc212	2018_2019	170213001	21	42	120	3	0
119	csc212	2018_2019	170213003	20	3	120	2	0
120	csc212	2018_2019	170213016	18	1	110	3	0
121	csc212	2018_2019	170213012	25	28	109	3	1
122	csc212	2018_2019	170213008	23	27	108	4	1
123	csc212	2018_2019	110213033	24	8	109	1	1
124	csc212	2018_2019	140213024	21	11	107	2	1
125	csc212	2018_2019	150213010	24	20	105	2	1
126	csc212	2018_2019	160213027	26	16	112	3	1
127	csc212	2018_2019	160213050	17	1	120	4	0
128	csc212	2018_2019	160213038	17	1	130	4	0
129	csc212	2018_2019	180213012	25	20	111	4	0
130	csc212	2018_2019	180213020	22	25	87	4	1
131	csc305	2019_2020	130213012	24	31	120	3	0
132	csc305	2019_2020	160213011	25	16	120	3	0
133	csc305	2019_2020	160213025	21	19	120	3	0
134	csc305	2019_2020	160213032	24	19	120	2	0
135	csc305	2019_2020	170213001	24	36	110	3	0
136	csc305	2019_2020	170213002	24	38	109	3	1
137	csc305	2019_2020	170213003	23	18	108	4	1

138	csc305	2019_2020	170213004	23	22	109	1	1
139	csc305	2019_2020	170213005	26	30	107	2	1
140	csc305	2019_2020	170213008	21	9	105	2	1
141	csc305	2019_2020	170213009	22	29	112	3	1
142	csc305	2019_2020	170213012	25	41	120	4	0
143	csc305	2019_2020	170213016	24	26	130	4	0
144	csc305	2019_2020	170213017	26	24	111	4	0
145	csc305	2019_2020	170213019	24	26	120	4	0
146	csc305	2019_2020	170213020	24	26	87	4	1
147	csc305	2019_2020	170213022	21	36	120	3	0
148	csc305	2019_2020	170213023	23	30	120	3	0
149	csc305	2019_2020	180213012	25	25	120	3	0
150	csc305	2019_2020	180213020	24	27	120	2	0
151	csc411	2019_2020	160213002	20	32	110	3	0
152	csc411	2019_2020	160213003	20	0	109	3	1
153	csc411	2019_2020	160213004	23	28	108	4	1
154	csc411	2019_2020	160213005	23	37	109	1	1
155	csc411	2019_2020	160213005	20	22	107	2	1
156	csc411	2019_2020	160213006	22	32	105	2	1
157	csc411	2019_2020	160213007	22	32	112	3	1
158	csc411	2019_2020	160213007	23	23	120	4	0
159	csc411	2019_2020	160213008	25	20	130	4	0
160	csc411	2019_2020	160213009	24	52	111	4	0
161	csc411	2019_2020	160213010	26	30	120	4	0
162	csc411	2019_2020	160213011	24	28	114	3	1
163	csc411	2019_2020	160213012	23	33	87	4	1
164	csc411	2019_2020	160213013	25	43	120	3	0
165	csc411	2019_2020	160213014	23	63	120	3	0
166	csc411	2019_2020	160213017	26	48	120	3	0

167	csc411	2019_2020	160213018	23	38	120	2	0
168	csc411	2019_2020	160213019	25	25	110	3	0
169	csc411	2019_2020	160213020	20	32	109	3	1
170	csc411	2019_2020	160213021	24	26	108	4	1
171	csc411	2019_2020	160213022	23	18	109	1	1
172	csc411	2019_2020	160213024	20	32	107	2	1
173	csc411	2019_2020	160213025	22	19	105	2	1
174	csc411	2019_2020	160213026	22	42	112	3	1
175	csc411	2019_2020	160213027	23	10	120	4	0
176	csc411	2019_2020	160213028	25	64	130	4	0
177	csc411	2019_2020	160213029	24	24	111	4	0
178	csc411	2019_2020	160213030	26	37	120	4	0
179	csc411	2019_2020	160213031	24	33	87	4	1
180	csc411	2019_2020	160213032	23	32	120	3	0
181	csc411	2019_2020	160213034	26	15	120	3	0
182	csc212	2020_2021	190213001	26	47	120	2	0
183	csc212	2020_2021	190213002	26	39	110	3	0
.	.	.	.	.	.	.	.	.
.	.	.	.	.	.	.	.	.
.	.	.	.	.	.	.	.	.
281	csc305	2022_2023	200213030	24	33	112	3	1
282	csc305	2022_2023	200213033	26	25	120	4	0
283	csc419	2022_2023	180213001	20	12	130	4	0
284	csc419	2022_2023	180213002	20	20	111	4	0
285	csc419	2022_2023	180213003	25	20	120	4	0
286	csc419	2022_2023	180213004	21	10	114	3	1
287	csc419	2022_2023	180213005	26	19	87	4	1
288	csc419	2022_2023	180213006	27	34	120	3	0
289	csc419	2022_2023	180213007	23	31	120	3	0
290	csc419	2022_2023	180213008	21	40	120	3	0

291	csc419	2022_2023	180213009	22	41	120	2	0
292	csc419	2022_2023	180213011	23	37	110	3	0
293	csc419	2022_2023	180213013	25	15	109	3	1
294	csc419	2022_2023	180213014	21	29	108	4	1
295	csc419	2022_2023	180213015	23	40	109	1	1
296	csc419	2022_2023	180213017	27	13	107	2	1
297	csc419	2022_2023	180213018	21	21	105	2	1
298	csc419	2022_2023	180213019	20	25	112	3	1
299	csc419	2022_2023	180213021	26	14	120	4	0
300	csc419	2022_2023	180213022	23	20	130	4	0
301	csc419	2022_2023	180213023	24	36	87	4	1
302	csc419	2022_2023	180213024	25	40	120	3	0
303	csc419	2022_2023	180213025	23	25	120	3	0
304	csc419	2022_2023	180213026	25	15	120	3	0
305	csc419	2022_2023	180213028	22	24	120	2	0
306	csc419	2022_2023	180213029	20	20	110	3	0
307	csc419	2022_2023	180213033	23	17	109	3	1
308	csc419	2022_2023	180213035	21	22	108	4	1
309	csc419	2022_2023	180213038	20	30	109	1	1
310	csc419	2022_2023	180213039	21	30	107	2	1

Appendix II

Figure 2

future prediction.... ¶

In [23]:

```
future_data = pd.DataFrame({
    'p_tt': [130, 140, 120, 125, 150, 135, 128, 110, 115, 145, 132, 138, 120, 124, 140],  # Hours spent on activities
    'no_online_participate': [3, 4, 2, 3, 5, 3, 2, 4, 3, 4, 3, 2, 2, 4, 5],  # Number of online/lab participation
    'hackathon_participation': [1, 0, 1, 1, 1, 0, 0, 1, 1, 0, 1, 1, 0, 0, 1]  # Hackathon participation (1 = Yes, 0 = No)
})

# Predict future performance for the 15 students
future_predictions = model.predict(future_data)

# Output the predictions
for i, score in enumerate(future_predictions, start=1):
    print(f"Predicted test score for student {i}: {score:.2f}")
```

```
Predicted test score for student 1: 23.30
Predicted test score for student 2: 23.42
Predicted test score for student 3: 24.28
Predicted test score for student 4: 23.93
Predicted test score for student 5: 23.42
Predicted test score for student 6: 23.30
Predicted test score for student 7: 23.64
Predicted test score for student 8: 24.64
Predicted test score for student 9: 23.93
Predicted test score for student 10: 23.42
Predicted test score for student 11: 23.30
Predicted test score for student 12: 23.64
Predicted test score for student 13: 24.28
Predicted test score for student 14: 24.64
Predicted test score for student 15: 23.42
```