

Dropout Decision Support System (DDSS): A Data-Driven Approach for Student Dropout Analysis

Jecca Claire Pepe¹; Shane Bonghanoy²; John Christian Azores³;
Ralph Renson Enisimo⁴; Emmie Faye Marione Matabile⁵; Jeffrey Calim⁶;
Johani Basaula⁷

¹College of Computer Studies, Our Lady of Fatima University, Metro Manila, Philippines

²College of Computer Studies, Our Lady of Fatima University, Metro Manila, Philippines

³College of Computer Studies, Our Lady of Fatima University, Metro Manila, Philippines

⁴College of Computer Studies, Our Lady of Fatima University, Metro Manila, Philippines

⁵College of Computer Studies, Our Lady of Fatima University, Metro Manila, Philippines

⁶College of Computer Studies, Our Lady of Fatima University, Metro Manila, Philippines

⁷College of Computer Studies, Our Lady of Fatima University, Metro Manila, Philippines

Publication Date: 2026/04/08

Abstract: Student dropout remains a major problem in universities in the Philippines. Although student data is available, many institutions lack automated systems for early identification of at-risk students. This study developed a Dropout Decision Support System using a threshold-based rule system combined with the Support Vector Machine algorithm. The web-based system classifies students into High, Medium, and Low risk categories. Data were collected from the Registrar and Guidance offices, covering several academic years. Key variables influencing dropout risk included grade point average, attendance, and financial balance. System evaluation followed a recognized software quality model and involved all our respondents. Results indicated strengths in usability, functionality, and reliability, showing that the system was accurate, intuitive, and reliable. Analysis of Variance results, no significant differences among evaluator groups. Model validation demonstrated strong performance across all risk categories, confirming the system's effectiveness as an early warning tool.

Keywords: Dropout Prediction, Machine Learning, Support Vector Machine, Decision Support System.

How to Cite: Jecca Claire Pepe; Shane Bonghanoy; John Christian Azores; Ralph Renson Enisimo; Emmie Faye Marione Matabile; Jeffrey Calim; Johani Basaula (2026) Dropout Decision Support System (DDSS): A Data-Driven Approach for Student Dropout Analysis. *International Journal of Innovative Science and Research Technology*, 11(3), 3537-3542.

<https://doi.org/10.38124/ijisrt/26mar1928>

I. INTRODUCTION

Even with its recent developments, student dropout still represents a problem of epic proportions in Philippine tertiary education, largely resulting from academic distress, money constraints, attendance problems, and absence of early-warning systems. The school collects a large amount of student data, but many still monitor them manually, leading to late identification of at-risk students. This poses the need for an automated, fact-based solution that will analyze many risk factors at once. To this end, the researchers created the Dropout Decision Support System (DDSS), a machine learning scheme capable of predicting dropout risk based on GPA, attendance, and financial status. The system classifies students' risk by using threshold-based

analysis and an SVM algorithm and alerts and reports to the school staff. DDSS uses a CSV based methodology that allows training information to be processed on a file to detect low-achieving students at a timely temporal rate to help guidance staff, faculty, and administrators offer appropriate interventions in academic, financial, or counseling aspects.

It will aim to (1) plan the DDSS using machine learning to act as a dropout identification solution, (2) build the system to effectively detect at-risk students, (3) design the DDSS to classify the risk levels accurately, (4) evaluate the system using ISO 25010 derived from feedback from Registrar Staff, Guidance Staff, and Experts, and (5) validate the results with academic and financial indicators

(GPA, attendance, balance) to determine the accuracy in predicting high-, medium-, and low-risk classifications. The researchers aim to predict dropout risk using the GPA, attendance, and financial data analyzed through the SVM and threshold-based values. Note that although DDSS may give ideas on interventions to consider such as financial aid, tutoring, and counseling, it is based on full records which may be inaccurate in the case of missing data. The system does not consider factors like mental health, family problems, or motivation, and it is simply human decision-making. To avert abuse of or leaking of student data, ethical and secure management of data must be maintained.

II. METHODOLOGY

➤ *Algorithm*

The developed system used Support Vector Machine (SVM) as the main algorithm. The way SVM works is by determining the best boundary to separate which students are expected to be dropping out from others, who are not likely to drop out. When separation does not form a line on the data, the SVM considers a kernel function, such as the RBF kernel, for further transformation of the data into a higher-dimensional space to facilitate separation. As a result, the SVM is used for educational datasets with complex relationships. Threshold values of GPA, attendance, and balance were also applied alongside the model to provide transparency and explainability of risk classification.

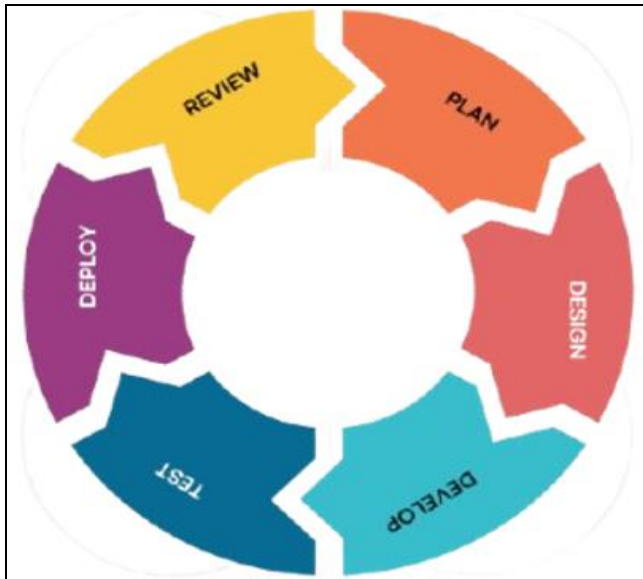


Fig 1 Agile Methodology

➤ *System Design*

For development of the Dropout Decision Support System (DDSS) the Agile methodology was used and helped ensure flexibility and continuous improvement. With Agile concepts, in this system, the DDSS became more effective, more accurate, and more flexible in predicting student dropout risk to allow organizations to better act in response to evolving conditions and improve on retention methods.

➤ *Plan*

In the plan phase, the researchers collected information from the registrar staff, guidance staff, and e-lite center to identify the main factors why students are dropping out. The researchers examined the existing processes and found out that monitoring was mostly manual and often delayed. This stage helped the system requirements, the datasets needed, and the specific indicators to be used for the SVM model and for the threshold-based analysis.

➤ *Design*

At this point of implementation the planning was transformed into a system structure for predicting and preventing student dropout. Front-end (HTML, CSS, JavaScript) delivered a simple GUI, back-end (Python) processed data, built SVM model, and identified the at-risk students. The database for student records and prediction outputs was handled by PHP. The dashboard allows schools to make proactive interventions to assist their student populations by showing real-time probability of dropouts based on a school's historical pattern and trend analysis as well as providing the school with charts that show the school's high-risk student populations. Predictive modeling made use of the SVM algorithm and produced high accuracy in predicting dropout risks.

➤ *Theoretical Framework*

This study is informed by data and machine learning to provide evidence-based recommendations for predicting and reducing student dropout. It is based on two key retention concepts—Tinto's Student Integration Theory (which associates dropout with low academic performance, financial problems, and poor engagement) and Bean's Student Attrition Model (which focuses on external and personal factors including economic challenges). Both underscore the early identification of at-risk students. Accordingly, in accordance with these theories, the proposed study uses the SVM algorithm that effectively handles multiple risk factors. Previous studies have found support for machine learning to predict student attrition [1]. These foundations support the DDSS, a predictive analytic tool that helps schools recognize students at risk at an early stage and facilitate proactive interventions to increase retention.

➤ *Conceptual Framework*

This diagram demonstrates data flow of a machine learning applied system called the Dropout Decision Support System (DDSS). It begins with uploading a CSV datasets into the system. Python first cleans and prepares the data once the file has been loaded to allow analysis of data. When the data is ready, it is input to the machine learning algorithm, which classifies every student in terms of potential dropout risk. The outcomes of these are then saved in the database. It generates the predictions that are displayed on the dashboard. This provides administrators with a visibility into which students are at risk, allowing them to make informed decisions and act on issues early.

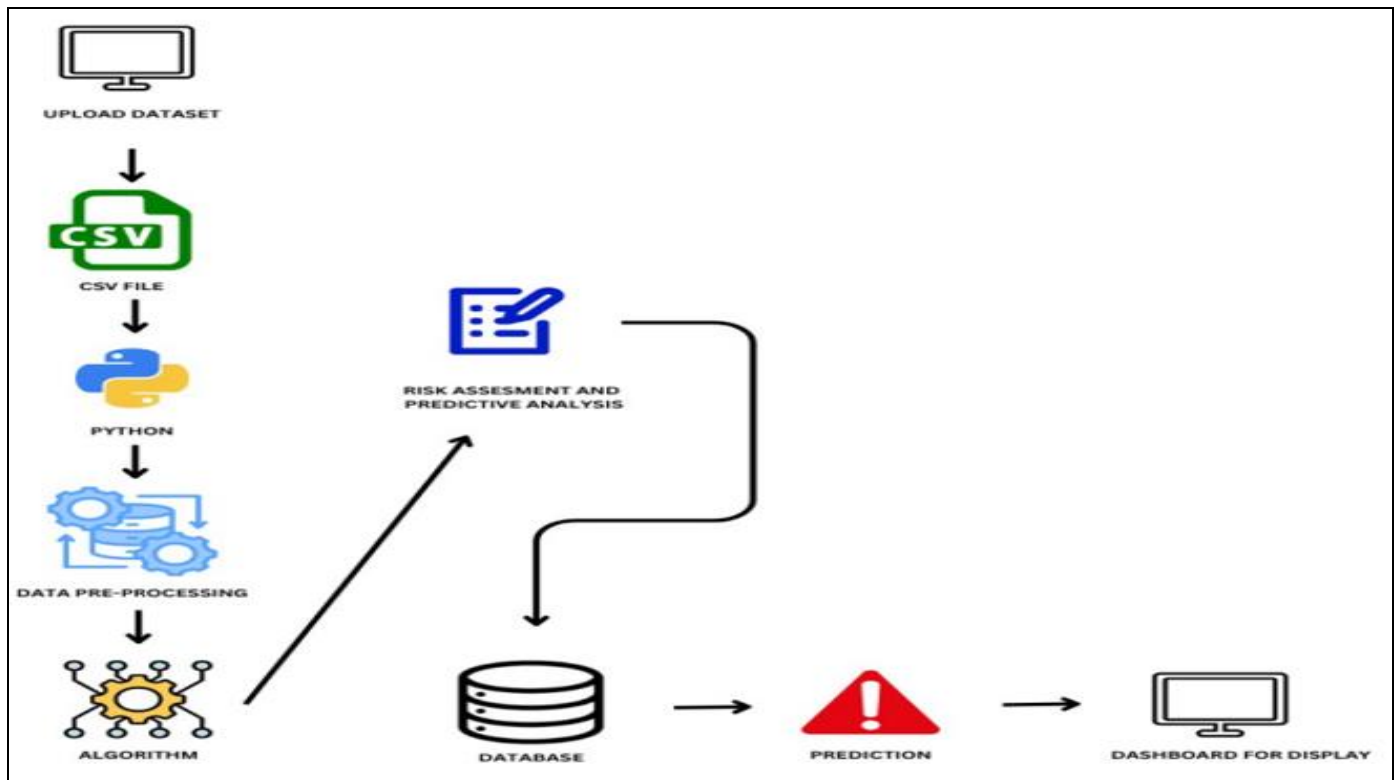


Fig 2 Conceptual Framework

➤ *Develop*

The researchers in this phase designed the web-based DDSS to predict dropout risks and help administrators design focused interventions. Its primary element is the student dashboard that offers student-specific risk analytics like probability scores, contributing factors, and trends. This allows for personalized intervention, depending on individual student characteristics. The SVM model evaluates multiple risk variables, while data preprocessing manages missing data and improves model accuracy and reliability.

➤ *Testing*

The researchers will perform tests to help detect any errors or issues in the DDSS that require fixing. The researchers conducted an analysis of student data to distinguish patterns among high, medium, and low dropout risk levels, allowing the system to generate accurate predictions based on academic and financial indicators. And that is a great way to validate that the DDSS is creating accurate, data driven insights that can be acted upon earlier.

➤ *Deployment*

The DDSS has reached its final working state. All functional requirements have been completed, and the developed system is now ready for institutional evaluation based on its intended purpose dropout risk and supporting early intervention.

➤ *Review*

A review of data from DDSS will assess its efficacy in decreasing dropout rates and enabling interventions through the student dashboard. Performance of the system will be assessed by continuous analysis with new datasets and by

comparing dropout rates before and after implementation. Review results will inform improvements to the algorithm, dashboard usability, and security features. This continuous review process maintains the DDSS’s enhanced predictive accuracy whilst also improving its ability to inform institutional interventions.

III. RESULTS AND DISCUSSION

➤ *Finding/Answers for Each Statement of the Problem*

- The researchers planned and created a solution to accurately identify students at risk of dropping out through the Dropout Decision Support System (DDSS).

The researchers planned the Dropout Decision Support System (DDSS) to accurately determine at-risk students. They obtained information from Registrar Staff, Guidance Staff and the E-Lite Center, datasets from 2021–2025 regarding student withdrawal profiles and reasons why students had left. Data were used to identify GPA, attendance levels, and financial balance as the key predictors of dropout risk, with course requirements and records of the institution acting as the thresholds. An SVM algorithm was chosen; it is an effective method to cope with complex data and classify students properly. After staff contributions, dataset manipulation, threshold estimation, and algorithm choice, the DDSS was introduced as a reliable tool to address the university’s dropout challenges.

- The researchers developed the DDSS with a user-friendly dashboard that processes the students’ data and provides predictive insights.

The DDSS is developed with an HTML/CSS and React.js front-end with a Flask back-end server to handle CSV files of student GPA, attendance, and financial balance from Registrar and Guidance Offices. It implements a Support Vector Machine (SVM) algorithm to output predictions, which appear on both the dashboard and the risk classification module. This allows administrators to monitor and support at-risk students based on classifiers and a user-friendly interface to identify high, medium, and low risk students based on their current performance.

The researchers designed the DDSS a systematic design with rule-based threshold analysis and machine learning classification. Threshold criteria were initially developed from historical student data from 2021 to 2024 based on university requirements providing the baseline conditions for risk assessment. The threshold results were then processed using the Support Vector Machine (SVM) algorithm. SVM is able to analyze this data and create more accurate classifications and assign every student to a risk level based on that interaction between multiple risk factors.

- The researchers designed the DDSS using the Support Vector Machine (SVM) algorithm to effectively classify students at high, medium, and low risk.

Table 1 Overall Evaluation Result of the System

Indicator	Weighted Mean	Interpretation
Functional Suitability	3.46	Agree
Usability	3.63	Strongly Agree
Reliability	3.89	Strongly Agree
Security	3.52	Strongly Agree
Performance Efficiency	3.56	Strongly Agree
Compatibility	3.84	Strongly Agree
Flexibility	3.50	Agree
Total Mean	3.63	Strongly Agree

Data has been gathered, and as seen in Table 1, it can be seen that the Dropout Decision Support System performs well on the ISO 25010 measures with weighted mean around 3.63 (Strongly Agree). This means that the system is reliable, easy to use, efficient, secure, and usable on multiple

systems, its users report. High reliability (3.89) and compatibility (3.84) are indications that the developed system processes the data consistently and is able to work smoothly across browsers.

Table 2 Descriptive Analysis Result Based on Experts, Registrar, and Guidance Staff Evaluation

Criteria	N	Minimum	Maximum	Mean	Std. Deviation
Functionality	23	2.75	4.00	3.543	0.443
Usability	23	2.50	4.00	3.641	0.493
Security	23	1.88	4.00	3.408	0.591
Total	69	2.38	4.00	3.53	0.51

The results of Descriptive Analysis of the DDSS, based on expert, Registrar Staff, and Guidance Staff evaluations are presented as Table 3. The Dropout Decision Support System (DDSS) gave an average overall score of 3.53 (Agree) for functionality, usability, and security,

revealing that the systems were functional, easy to operate, and secure to the users. The Usability (3.64), Functionality (3.54), and Security (3.41) mean scores similarly indicated similar positive evaluations across evaluator groups.

Table 3 Analysis of Variance (ANOVA) Result Based on Experts, Registrar, and Guidance Staff Evaluation

Variables Source of Variation	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	0.325	2	0.163	0.8120	0.458
Functionality Within Groups	4.006	20	0.200		
Total	4.332	22			
Between Groups	0.355	2	0.178	0.7110	0.503
Usability Within Groups	4.998	20	0.250		
Total	5.353	22			
Between Groups	0.151	2	0.076	0.2000	0.820
Security Within Groups	7.543	20	0.377		
Total	7.694	22			

- The Guidance Staff, Registrar Staff, and Experts evaluated the proposed system entitled "Dropout

Decision Support System (DDSS): A Data-Driven Approach for Student Dropout Analysis" based on (ISO

25010: 2011) systems and Software Quality Requirements and evaluation.

We see strong results in Usability (3.63) and Performance Efficiency (3.56), implying the system is user-friendly and responsive. These results align with [3], that such data-driven decision support systems obtain strong evaluations as they can guide educational decision-making and early intervention. Overall, the ratings indicate that the DDSS does its job and it is well-accepted by its users.

This was consistent with the view of [9] that high-performing systems have more favorable ratings across different user groups because it has consistent and predictable behavior. Similarly, the Dropout Decision Support System (DDSS) demonstrated no significant differences in its evaluations between groups, further supporting its reliable and effective performance in achieving its intended goals.

The results of Analysis of Variance (ANOVA) of the Dropout Decision Support System (DDSS), based on expert, Registrar Staff, and Guidance Staff evaluations are presented as Table 4 and all indicated significance values greater than 0.05 ($p = 0.818, 0.846, 0.873, \text{ and } 0.582$). That means there were no significant differences in how the groups appraised the system’s performance and quality. This phenomenon agrees with [5], whose ANOVA found no significant differences in stakeholder ratings ($F = 0.789, p = 0.457$).

The level of consistency throughout groups suggests that clear, effective systems tend to obtain similar ratings, and in the same way all reviewers perceived the Dropout Decision Support System (DDSS) as reliable, functional, and stable. These findings confirm the DDSS is consistent and reliable across different user roles, and boost user trust in the system by providing predictable system behavior.

Table 4 Overall Confusion Matrix Result for High Risk

TP	FN	TN	FP	Accuracy	Precision	Recall
18	14	40	3	77.33%	86%	56.25%

In the paper by [4], the researcher applied joint learning approaches to predicting student performance at various levels of risk and reached 81% accuracy, 83% precision, and 81% recall: multiclass models show the promise of reliably

classifying students into risk groups. Also, the DDSS obtained 77.33% accuracy, 86% precision, and 56.25% recall for high-risk students according to expert evaluation.

Table 5 Overall Confusion Matrix Result for Medium Risk

TP	FN	TN	FP	Accuracy	Precision	Recall
31	4	20	20	68%	61%	88%

[8] used a multiclass model to predict student performance and reported balanced accuracy and precision scores of 65–75%, noting that medium-risk students are harder to classify because their characteristics overlap with both high- and low-performing groups. For medium-risk students, according to expert evaluation, the Dropout

Decision Support System (DDSS) had 68% accuracy, 61% precision, and 88% recall. The high recall indicates that the system successfully identifies most medium-risk cases, allowing timely support for students whose performance may decline. The results indicate that the DDSS performs well in detecting students in the middle-risk category.

Table 6 Overall Confusion Matrix Result for Low Risk

TP	FN	TN	FP	Accuracy	Precision	Recall
2	6	66	1	91%	67%	25%

[7] applied a multiclass model for predicting student outcomes and managed more than 93% accuracy and remarked that low-risk students are typically easy to classify but frequently present low recall due to less evident warning signs. The DDSS provided expert results that were comparable, as it has a 91% accuracy, 67% precision, and 25% recall for low-risk students. The high accuracy suggests that the system reliably identifies students who are not at risk, and the low recall indicates the challenges of detecting all low-risk cases. This confirms that the DDSS performs consistently with existing research.

The researchers validated the accuracy of the developed system by comparing the predicted risk categories generated by the Dropout Decision Support System (DDSS) with the actual categories, which were identified by expert evaluators (2 Data Analyst and 1 System Analysts). The experts used predefined thresholds for GPA, Attendance, and Balance as the guide of their classification, while DDSS performed its own automated classification. By comparing the predicted categories and the actual categories results, a confusion matrix was constructed to measure the system’s performance.

- The researchers assessed and utilized the DDSS to accurately predict the dropout risks of students based on their academic and financial factors (GPA, attendance, and balance).

Results obtained for Dropout Decision Support System (DDSS) are close in precision and accuracy, showing that the DDSS recognizes high-risk students accurately, although lower recall also confirms findings that multiclass models

often fail to learn all high-risk cases because of small sample sizes. This verifies that the DDSS works within an acceptable

range of high-risk classification.

Table 7 Model Confusion Matrix Format

Classification Model Score Description	High Risk	Medium Risk	Low Risk
Correct classification	18 correctly classified	31 correctly classified	2 correctly classified
Incorrect classification	14 incorrectly classified	4 incorrectly classified	6 incorrectly classified
Accuracy (model)	77.33%	68%	91%
Recall	56.25%	89%	25%
Precision	86%	61%	67%
F1 Score	68%	72.4%	36.4%

Based on the confusion matrix, the Dropout Decision Support System (DDSS) performed the best among high-risk students with accuracy of 77.33%, precision of 86%, recall of 56.25%, thus decreasing the errors obtained from false positives but missing some cases of high risk. On medium-risk classes, it achieved 68% accuracy, 61% precision, and 89% recall, and exhibits strong agreement with certain misclassifications. While the low-risk group achieved 91% accuracy and 25% recall, indicating that it is conservative model in detecting low-risk cases. F1 Scores (68% high; 72.4% medium; 36.4% low) indicate balance of precision vs recall and indicate that there is need for sensitivity improvements, particularly for low-risk samples. These results are consistent with [6] which found different precision and recall distributions of the multi-class dropout model, which is evidence that the DDSS demonstrates typical multicast prediction performance.

IV. CONCLUSION

The Dropout Decision Support System (DDSS) was designed and implemented by the researchers to predict student dropout risk through GPA, attendance, and financial status. To achieve that, the system was implemented with the SVM algorithm with threshold-based analysis and was displayed on a simple dashboard which automatically classifies these students in risk groups, allowing the administrators to recognize trends and implement remedial measures at an early stage to increase retention. The Support Vector Machine and threshold DDSS design improved early detection and decision-making support of early dropout prevention strategies by enriching the data. Confusion matrix validation showed reliable classification accuracy in low-, medium-, and high-risk students.

Evaluation by the Guidance Staff with ISO 25010 confirmed that the system was functional, usable, reliable, and secure; Registrar Staff assured that it integrated with student records and processed data appropriately. Experts found the system reliable and identified areas for improvements for both its accuracy and security. Descriptive analysis reported high usability ($M = 3.64$), functionality ($M = 3.54$), and acceptable security ($M = 3.41$), while ANOVA did not differ between the participant groups ($p > 0.05$) and no differences were detected, which

confirmed the reliability of the perceptions of performance by evaluators. At a high level, the proposed DDSS, based on SVM classifiers and threshold analysis, correctly represents dropout risk in terms of GPA, attendance, and financial balance and confirms its valid accuracy and usability, providing universities with tools to prevent potential dropouts.

REFERENCES

- [1]. M. Y. Amare and S. Simonova, "Global challenges of students dropout: A prediction model development using machine learning algorithms on higher education datasets," SHS Web of Conferences, vol. 129, p. 09001, 2021.
- [2]. C. Dann, P. Redmond, M. Fanshawe, A. Brown, S. Getenet, T. Shaik, X. Tao, L. Galligan, and Y. Li, "Making sense of student feedback and engagement using artificial intelligence," Australasian Journal of Educational Technology, 2024.
- [3]. I. E. Livieris, T. A. Mikropoulos, and P. Pintelas, "A decision support system for predicting students' performance," 2016.
- [4]. M. R. Alzahrani, "Predicting student performance using ensemble models and learning analytics techniques," 2024.
- [5]. J. Guzman, "Stakeholders' participation in School Improvement Plan and school performance of secondary schools," International Journal of Arts, Sciences and Education, 2022.
- [6]. S. Kim, E. Choi, Y.-K. Jun, and S. Lee, "Student dropout prediction for university with high precision and recall," Applied Sciences, vol. 13, no. 10, 2023.
- [7]. J. A. C. Montes, M. C. B. Rodríguez, and M. A. Chamorro, "Drop-out prediction in higher education using imbalanced multiclass dataset," Journal for ReAttach Therapy and Developmental Diversities, 2023.
- [8]. S. S. Padmanavar and S. Begum, "Prediction of student performance using genetically optimized feature selection with multiclass classification," IJETT, 2022.
- [9]. K. Sagar and A. Saha, "Exploring the effect of task difficulty on usability scores of academic websites computed using SUS," 2020.