EAI Endorsed Transactions
on Internet of Things

Research Article EAI.EU

Interlink Platform for School, Higher and Technical Education in India: Design Platform

Nitesh Ghodichor[1,*], Prathm Chopde[2] , Mansi Choudhari[3], Saloni Rangari[4] and Pratham Badge[5]

[12345]Department of Computer Technology, Priyadarshini College of Engineering, Nagpur, M.S. 440019 (India)

Abstract

INTRODUCTION: aims to fill the knowledge gap in understanding the factors that contribute to academic performance and dropout rates in the Indian education system. The study proposes a “Interlinked platform for school, higher and technical education in India”, a unified digital space for students, educators, administrators and government agencies.
OBJECTIVES: The platform is designed to track students' academic performance across different educational levels and visualize dropout rates.
METHODS: The research uses a comprehensive methodology that integrates data analysis techniques and visualization frameworks and uses Python libraries such as NumPy, Pandas, Matplotlib and Scikit-Learn. Student academic outcomes are analyzed using linear regression and K-means clustering, and dropout rates are predicted using logistic regression.
RESULTS: Successfully developed a web-based system with a user-friendly interface.
CONCLUSION: The aim of the research is to provide institutions with valuable insights to understand the factors that contribute to dropout rates and to develop targeted interventions to address potential triggers of dropout.

Keywords: Performance Analysis, Interlinked Platform, Dropout Prediction, Educational

Received on 23 May 2024, accepted on 22 March 2025, published on 13 April 2025

doi: 10.4108/eetiot.6152

*Corresponding author. Email: niteshgho@gmail.com

1. Introduction

Any country's progress depends heavily on education, and India is no exception. The standard of education directly impacts the nation's social and economic development. India enrols a large number of students at all levels, from elementary to university. Prior studies have demonstrated that several characteristics, such as family background, socioeconomic level, and resource availability, affect academic achievement. Family history, socioeconomic level, and resource availability influence academic achievement and contribute to India's high dropout rates. Research on the factors influencing academic achievement and dropout rates in India is lacking, despite the significance of education. Students do not have access to an integrated platform that contains all of their academic achievement results from elementary school through higher education under the current system. The purpose of this study is to fill this knowledge gap by examining data on students' academic achievement in various areas and their past and current class grades. We are also using this data to visualise the rates of student dropout. A unified platform that guarantees smooth access to resources and fills in the gaps between different educational levels is necessary in the quickly changing field of education. By providing a single digital platform for interaction, knowledge sharing, and success in the educational ecosystem, the "Interlinked Platform for School Education, Higher Education, and Technical Education in India" aims to meet this requirement. The platform serves as a conduit between parents, instructors, students, and a country's government. The platform assists government agencies in identifying the benefits and grey areas when draughting policies related to education and the betterment of the student body. This website allows us to track students entering different educational sectors and determine the dropout rate. Through this website, we can monitor students' entry into various educational sectors and calculate the dropout rate. The agencies in charge of school, higher, and technical education are now more in sync with one another as a result of the unification of these three educational levels. By understanding current industry trends, the educational system will improve alignment with industry demands and assist students in enhancing their skills while also benefiting the industry. This platform allows us to condense the curriculum and concentrate on the students' skill development.

2. Research Objective

The aim of this research is to explore the potential uses of dynamic visualisation techniques and real-time data streams in educational environments. The emphasis is on their capacity to offer thorough insights into students' academic development. By monitoring a range of metrics in real time, such as attendance, grades, engagement levels, and other relevant indicators, we aim to create a comprehensive picture of students' academic careers. We also want to investigate how these technologies can help administrators and teachers recognise early signs of potential dropout risks. Our goal is to reduce dropout rates and create an environment that promotes student success and retention by developing efficient, timely intervention strategies based on these insights.

3. Literature Review

Previous research has extensively examined the application of clustering and data mining algorithms to assess student performance and classify students according to their academic performance. Researchers who looked at how non-cognitive traits like grit, self-efficacy, and goal orientation relate to academic success also found that affective variables are very useful for predicting how well a student will do in school [1]. Researchers have also studied a variety of factors, such as mental health, health-related behaviours, and sociodemographic indicators in relation to university students' academic performance [1, 2]. These factors also impact the retention rates of students. Predictive models and machine learning techniques have also been envisaged in the literature as being crucial for evaluating and predicting students' academic performance, more so in relation to their grades [3]. In this context, Educational Data Mining (EDM) techniques have been essential in analysing and predicting academic success, which was possible due to past grade data [4]. In a remarkable way, using performance data from courses, machine learning algorithms like deep artificial neural networks (DNN), decision trees, and logistic regression classifiers have been utilised to detect at-risk students early in their academic journeys [2, 4].Researchers have applied different resampling techniques, like SMOTE, ROS, ADASYN, and SMOTE-ENN, to the problem of imbalanced class distributions in datasets for student grade prediction, which has improved the predictive modelling performance [5, 6, 7]. Predictive models, which employ previous course grades, have also been found to have better performance in evaluating the success of students and, more importantly, identifying students who might have trouble facing academic difficulties at the outset of the semester [8, 9]. In the domain of visualising dropout, there have been varied methodologies and tools applied to get insights into student attrition dynamics. Many factors have been used to show how Educational Data Mining (EDM) techniques can be used to find patterns in academic grades and performance data that are linked to student dropout [11, 12]. For example, random forests, neural networks, and logistic regression are machine learning models that can be used to predict student dropout using academic performance as a predictor variable [13, 14]. In addition to these, survival analysis methods have proved useful in identifying those students who are at risk of dropping out from school over time, tracing academic grade trends over a period of time to look for patterns that point to such dropout behaviour. Sophisticated visualisation tools, like SDA-Vis, provide more detailed insights into the dropout rates that result from several areas of academic performance [15]. If you use a mix of these different approaches, you can fully understand why students drop out of school and come up with custom interventions that will lower the risk of dropping out and help students stay in school [16].

4. Methodology

This article aims to examine student academic performance and illustrate dropout trends using a comprehensive methodology that combines data analysis tools with visualisation frameworks. The emphasis is on utilising student grades, or scores, as the principal measures of academic performance gathered from dispersed assessment courses. The methodology has four primary phases: data collection, preprocessing, data analysis, and result interpretation. Python frameworks and modules, including NumPy, Pandas, Seaborn, and Scikit-Learn, are used for data preprocessing, analysis, and visualisation. We employ five classification techniques from Python's Scikit-learn module to construct prediction models: Decision Tree CART, Extra Trees Classifier, Random Forest Classifier, Logistic Regression, and C-Support Vector Classification. We evaluated the efficacy of several developed models on a dataset. The methodology involved the following steps:

Data Collection

Since the data collection procedure relates to a wealth of valuable information, it follows that it is a crucial component in any effort to understand and evaluate the student's performance within an educational context. Numerous characteristics of this dataset are pertinent to the students, including demographics like age, gender, ethnicity, and socioeconomic position as well as study-related information like grades or results attained by students in different subjects or courses. The amount of time spent studying or the type of interactions with peers and professors are a couple of other characteristics. Information on extracurricular activity involvement, prior academic performance, and occasionally even cohort peers are also appropriately identified. An in-depth picture of each student's academic career is provided by contextually recording other variables, such as course enrolment, the institution's location, and institutional features. Additionally, data collection and validation mechanisms will ensure data quality and reliability. It is also vital to note that every effort will be made to ensure that student data is secure and private per legal and ethical standards. Carefully sorting through this giant dataset might reveal subtle patterns and correlations that could shed light on the main factors that affect students' academic performance and guide research into ways to improve school results. It is crucial to keep in mind that stringent protocols have been established to guarantee the confidentiality and security of student information in compliance with legal and ethical requirements. Researchers are carefully choosing which parts of this huge dataset to look at to find complex patterns and connections that will help them figure out what is affecting students' academic success and how to help them do better.

Data Preprocessing

Any kind of alteration to the original data collection to get it ready for use by a data mining technique is referred to as "data preprocessing." During the data-gathering phase, data integration was done to produce a single data set. The data cleaning procedure included methods like filling in missing values for some numeric predictor attributes with the attribute mean. The process corrects data inconsistencies; we have removed records with characteristics unrelated to the problem. For instance, a student's name may not be crucial when evaluating their academic progress. Making early predictions about students' performance is a major obstacle to successfully satisfying their requirements. Gender, academic year, grade point average, total points gained by each student in the prior year, and course completion status were common features taken into account across all datasets. These are additional characteristics that define the academic success of the students. "Data preprocessing" refers to the process of cleaning, merging, transforming, and reducing the original data before analysing and visualising it. We do this to ensure that we meet the minimal needs and criteria set by research procedures for acquiring knowledge.

- **Data cleaning**: Data cleansing is the process of eliminating duplicate, missing, or insufficient data from a data set. There are several approaches to deal with missing values in attributes, including neglecting tuples, substituting a universal constant for the missing values, or using the attribute mean. The goal of this study is to add less missing data to course grade records while removing course grade records where data is frequently lacking. There are severe restrictions on the methodology employed in this study. The elimination of grade entries for more than two courses for which data is lacking is one example of this. Furthermore, the approach involves projecting these figures based on the average grade for the pertinent topic in cases when students are still missing course grades. It is also acknowledged that a zero in a course indicates that a student did not attend the exam and that the student's credit record will be removed. ·
- **Data Integration**: Because it resolves the issue of data redundancy that arises in every data set, the data integration process is crucial to the field of data management. During the data integration step, connected courses are purposefully combined to streamline the overall analysis process and get rid of any potential duplication. This pairing of related courses is especially crucial when the courses overlap, span multiple semesters, and hence call for distinct instructional materials. These classes can be taken together, and averaging the outcomes over several semesters prepares the data set for more intelligent and fruitful research in the future. In addition to simplifying the data set by removing superfluous attributes, this methodological approach ensures that each unique history is accurately recorded, which serves as the foundation for additional research. However, because the K-Means technique only handles numerical data, it necessitates some adjustments. The transformation of student identification numbers into character types occurs when merging data in a single database. To guarantee data format consistency and make additional analysis easier, the data type for each topic assessment is also converted to a numeric category with all decimal places set to zero.
- **Data reduction**: The process involves removing unnecessary components, such as credits and class time, to maximise the results of the presentation. We note only the total number of students enrolled and their corresponding course grades, focusing on the most crucial information. Additionally, a unique serial number is assigned to every student file, guaranteeing consistency and transparency in the clustering process and eventually improving statistical convenience.

Data analysis

It entails eliminating extraneous elements to optimize the presentation of results, including credits and class duration. Merely recording the total number of enrolled students and the corresponding course grades helps to focus attention on the most crucial information. Additionally, every student file is assigned a unique serial number, which guarantees consistency and transparency in the clustering process and eventually improves statistical convenience.

The following equation represents the linear regression model:

$$Y = \beta_0+\beta_1 x_1+\beta_2 x_2+\cdots+\beta_n x_n+\varepsilon$$

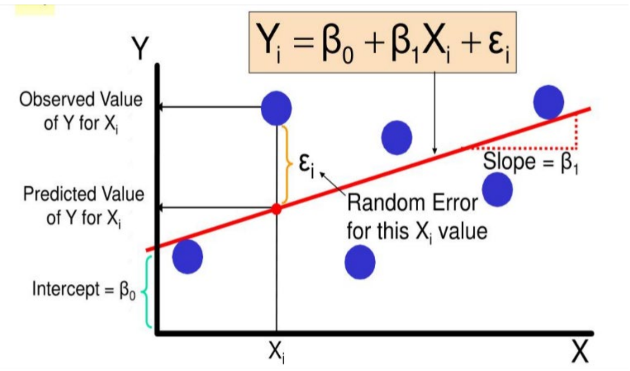

Figure 1. Linear Regression model

Where:

Y is the depending variable (e.g., final grade)

β_0 is the term for intercept.

β_1, β_2,…, β_n are the degree of the predictor terms x_1, x_2, …, x_n.

ε is the term of error, representing the difference between the observed and predicted values.

K-Means Clustering is an unsupervised machine learning approach that is highly effective at classifying materials into groups, or clusters, for the purpose of analysing student performance. An effective analytical technique for determining a number of students with similar performance on academic performance assessments is K-means clustering. This technique sorts data points based on their properties. As a result, instructors and educational institutions are better able to create student profiles and implement focused interventions. The input data is split up into K distinct clusters using the K-means clustering technique. Each cluster is defined by the average of its constituent parts. The procedure iteratively allocates each data point to the cluster with the closest mean until the assignments cease to fluctuate.

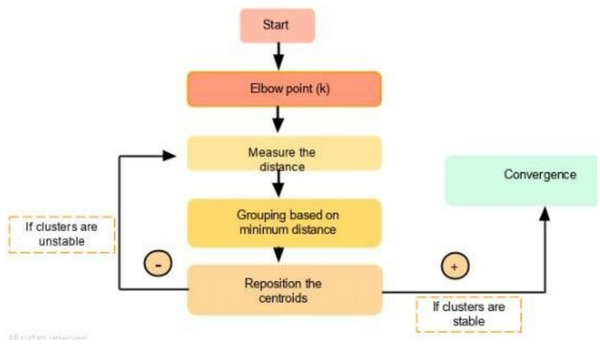

Figure 2. Working of k-means algorithm for clustering

Algorithm Overview:

i. Randomly initialize the centroids of K clusters.

ii. Assign every data point to the cluster whose centroid is nearest.

iii. Calculate the new centroid as the average of every point of data in the cluster.

iv. Continue with steps 2 and step 3 until there are no more changes occur.

K-means clustering, when applied to students' academic outcomes, can identify student groups with comparable academic performance. It may identify a class that excels in science but struggles in humanities, or vice versa. Students can then receive focused support by utilising these clusters. For instance, extra tutoring in science could be made available to pupils in a cluster who struggle in those disciplines. TTeachers can track the development of student groups over time via K-Means clustering. Teachers can assess the success of support interventions and make necessary modifications by regularly reassigning cluster assignments and performance trends.

Logistic regression for predicting at-risk students: It is a popular and significant supervised learning approach for predictive modelling applications, particularly those involving binary classification. It gives the probability of a binary result conditioned on a small number of outcome-predicting independent variables. The model adjusts its output to fall between 0 and 1, which represents the likelihood of the affirmative outcome, using the logistic function. Logistic regression can be used to estimate dropout risk in educational contexts. To ascertain which kids are more likely to drop out, it examines attendance records, grades, and academic performance. Timely intervention can be facilitated by using this information to establish early intervention and support measures.

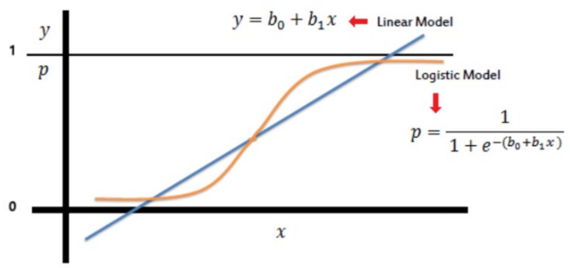

Figure 3. Linear Regression model

The equation for logistic regression is as follows:

$$P(Y = 1) = \frac{1}{1 + e^{-(b_0 + b_1 x)}}$$

Here:

- P(Y=1) is the probability of the event Y=1 (e.g., a student dropping out).
- X is the input variable (e.g., grades or attendance data).

- b0 and b1 are parameters of the model that are learned from the data.

The dependent variable in a logistic regression dropout prediction is whether or not the student leaves school. A student's grades, attendance history, involvement in extracurricular activities, etc., can all be considered independent variables. For instance, it is acceptable to use logistic regression to forecast a student's likelihood of dropping out if we know information on their attendance and grades. This model discovers the connection between a student's attendance, grades, and dropout status. Based on attendance and grades, the model can be used to forecast a new student's likelihood of dropping out after training.

Data Visualisation: In educational institutions, performance trends and dropout rates are analysed and displayed using data visualisation. The purpose of this graphical data representation is to facilitate the interpretation and comprehension of complex information. In order to effectively communicate two important parts of this research paper, performance trends and dropout rates, we employ data visualisation techniques.

To give insight into student grades and performance measures over time and across disciplines, performance trends are shown. We offer a thorough summary of student performance through the use of pie charts and bar charts, making it simple to identify strengths and weaknesses. Participants and administrators can monitor performance trends at the individual and cohort levels with this visualisation, pinpoint areas that need work, and modify interventions accordingly.

Interactive visualisations that analyse dropout rates by demographics, educational attainment, and other pertinent variables are used to display dropout rates. Using dynamic charts and dashboards, we generate dynamic dropout rate displays and let users examine and evaluate data according to various parameters. Using this visualisation method, institutions can create targeted interventions to deal with potential dropout triggers and gain important information about the factors that affect dropout rates.

5. Working

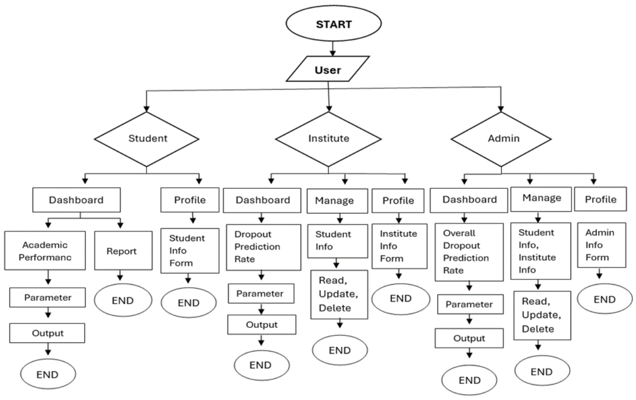

Figure 4. The working flow of website

The section describes the operational flow of the proposed interlinked educational data analysis platform. This platform is designed to serve students, educators, and institutions and uses data visualisation and machine learning to enable informed decision-making within the educational system.

1. User Authentication and Access Control: Through a secure login process, users gain access to the platform, which includes students, institutions and administrators. The system authenticates the credentials and differentiates user roles based on them. Upon successful login, the system redirects users to dashboards tailored to their specific roles. Student dashboards emphasise features related to accessing their profile information and viewing their grades, test scores, and other types of information. The institution dashboard contains overall performance analysis visualisations for students and dropout rates. Admin dashboards provide more complex functionality, including managing student information and profile updates.

2. **Student Functionalities**: Students can access and view their personal data and academic records via their profiles. This section gives them the opportunity to stay updated on their academic progress and may allow for limited profile updates within institutional policy.The platform provides students with interactive visualisations that transform their academic performance data into engaging and informative formats. These visualisations, possibly including pie charts and bar charts, can show grades, test scores, and other relevant metrics over time. This allows students to identify their strengths and weaknesses and track their progress across semesters or courses.

3. **Institution Functionalities**: Institute users have access to all student information management features through the system. Such functions may include adding new students, updating existing records, and possibly even deleting students in accordance with the institution's policies and data retention requirements. The system may also support bulk data management features for effective handling of student information. Using historical data on student performance, attendance, and possibly demographics, the system calculates the dropout risk score for an individual student. Institute users have access to the predictions to proactively determine which students are at risk and implement focused measures to increase retention. Institutional users gain insight into student retention by looking at the overall dropout rate at the institution. Institute users can create study information reports or study dropout reports. You can view the reports for further analysis.

4. **Administrator Functionalities**: Administrators have comprehensive capabilities to manage student information in the system. These capabilities can include adding new students, editing existing records, and even deleting student data based on institutional policies and data retention requirements. The platform enables efficient mass management of student information, such as uploading or exporting student data files. This simplifies the administrative process and data processing. Student profiles can be managed, e.g., adding, editing or deleting profile

information. This capability allows for the maintenance of accurate student information and would potentially allow for the inclusion of other areas of interest. The system uses algorithms, such as logistic regression, to predict students' risk of dropping out of school. The system uses historical data about student performance, attendance, and other potentially demographic factors to create dropout risk scores for individual students. Using such predictions, administrators can determine whether students are at risk and take specific action.

6. Proposed System

The proposed system is an interlinked educational platform that aims to deliver a comprehensive overview of a student's academic career, from primary school to higher and technical education. This platform is designed to track, analyse, and visualise academic data, providing valuable insights into a student's performance across different subjects and education levels. The platform serves as a central repository for academic data, making it easier to monitor a student's long-term academic development. Also, it allows analysing student performance in different subject categories; therefore, students can make informed decisions about the level of education and careers they would like to engage in. Students, teachers, and educational institutions can benefit from the platform's design. Students can receive insights about their academic performance and find areas for development. Teachers can use the data to tailor their teaching strategies to better support their students. Institutions can use the insights to implement better educational practices and policies.

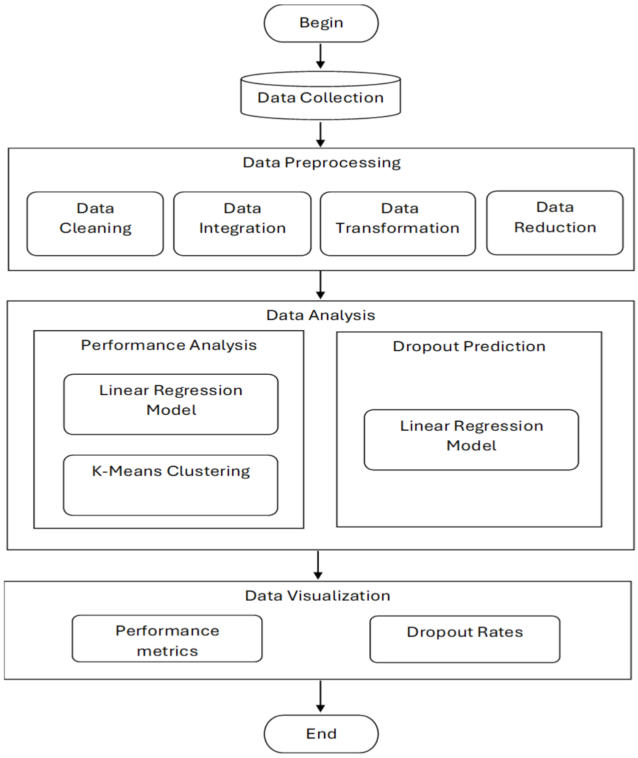

Figure 5. Data flow diagram

Functionality and Features:

- Data Tracking and Analysis: Students' academic data, including grades, attendance and performance in various subjects, are tracked and analysed by the system. This feature provides a comprehensive overview of a student's progress throughout their academic career.
- Visualisation of Data: The system uses Python libraries and functions to visualise data in the form of graphs, charts, and tables. This technique enhances comprehension and interpretation of the data.
- Custom Reports: The system generates personalised reports based on the analysed data that can be used by teachers and institutions to evaluate student performance and identify opportunities for improvement.
- Student Profile: Each student has a personalised profile on the platform that includes their academic achievements, attendance and other relevant information. This feature serves as a digital portfolio for students and can be accessed at any time.

7. Result and Discussion

The project successfully developed a web-based system for analysing student performance with a user-friendly and easily accessible homepage (Figure 3) aimed at students, institutes, and admin users. The homepage is designed in a clean and intuitive style with an emphasis on simple navigation. This page also has info for new users and a summary of the system's features and advantages.

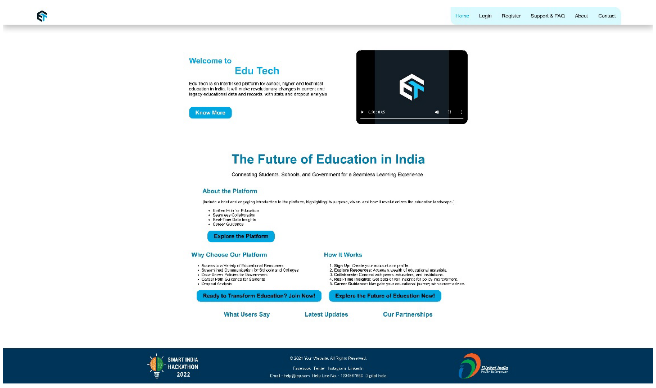

Figure 5. Home Page of Edutech

The student registration page provides a user-friendly form for new student enrolment. This form captures essential student information.
Institute registration contains details about the institute user, such as name, role within the institution, and contact information. Registration for institute users requires authorization from an existing administrator

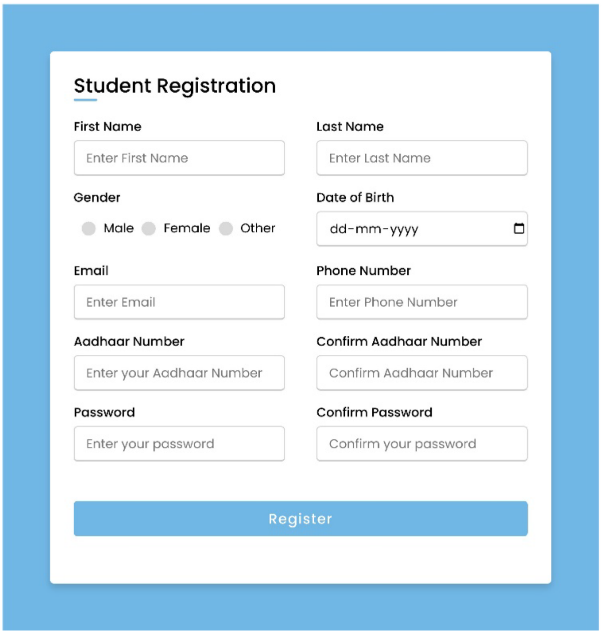

Figure 7. Student registration form

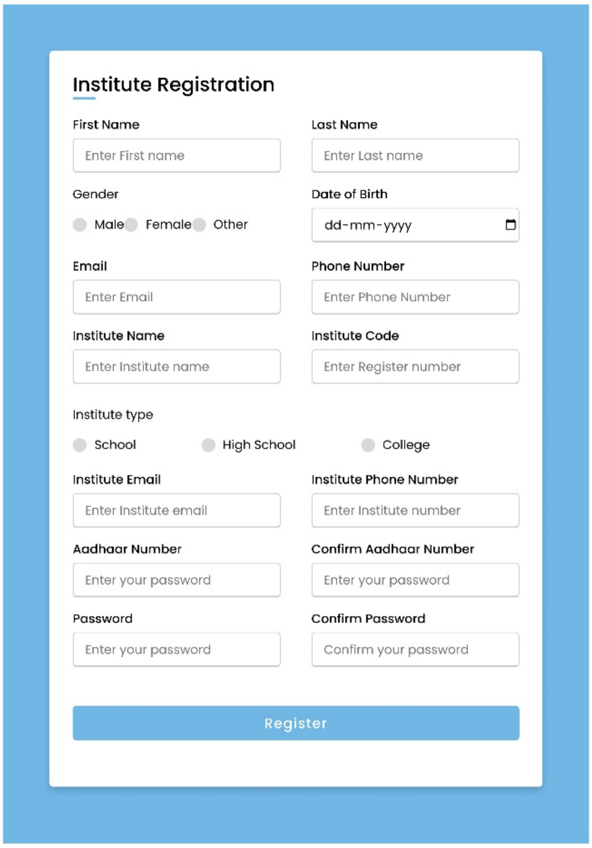

Figure 8. Institute registration form

The login page requires users (students, institutes, and admins) to provide valid credentials for authentication. The login page clearly distinguishes between student, institute and admin user logins. This feature facilitates a streamlined login process and ensures users are directed to the appropriate dashboard upon successful authentication.

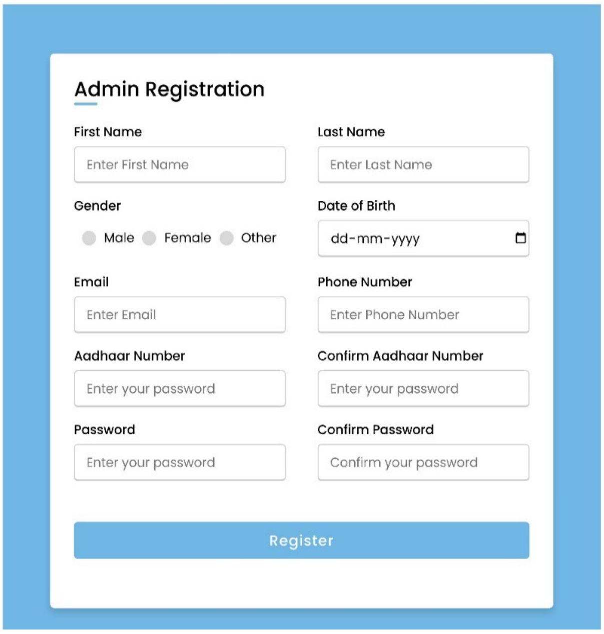

Figure 9. Admin registration form

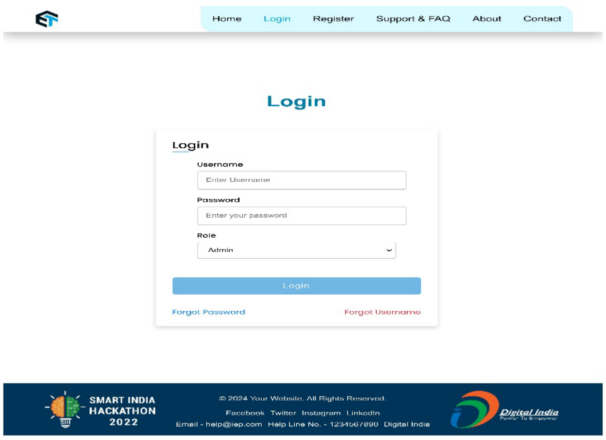

Figure 10. Login page

The About Us page outlines the project's fundamental principles, including its vision and mission.

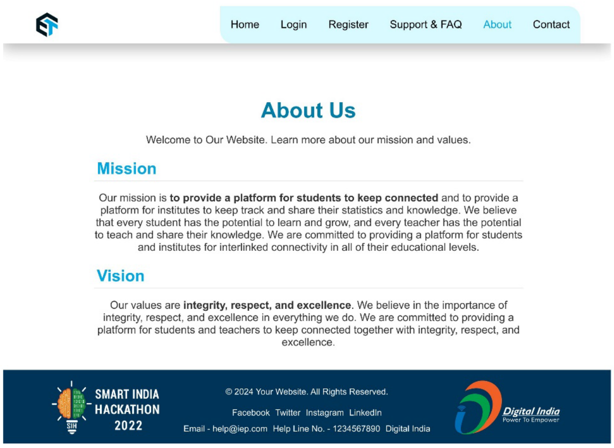

Figure 11. About us page

The Contact Us page includes a user-friendly contact form with an email address and message for feedback, enquiries, and suggestions for improvement.

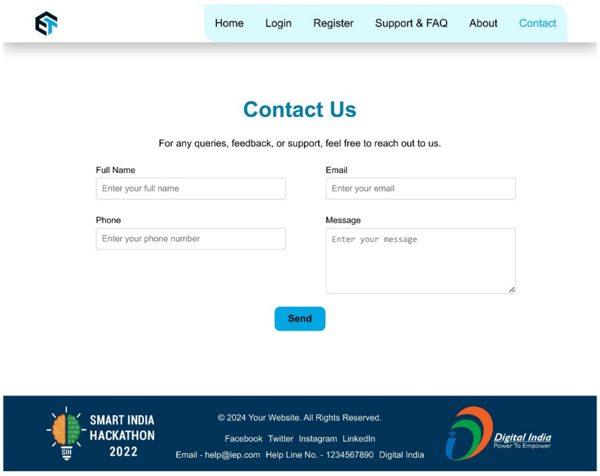

Figure 12. Contact us page

After registration, you successfully create your profile at Edutech. The profile includes your personal information and which module you belong to. You have full access to your details so you can read, update, and delete them.

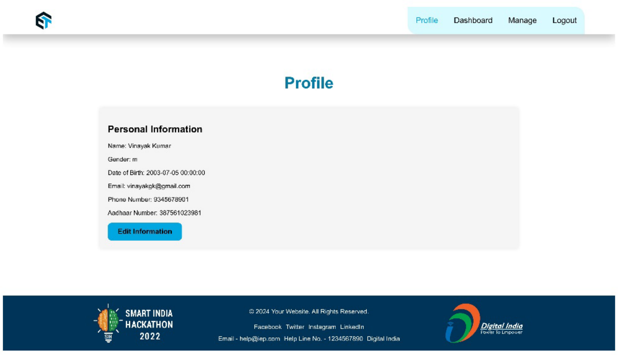

Figure 13. Profile page

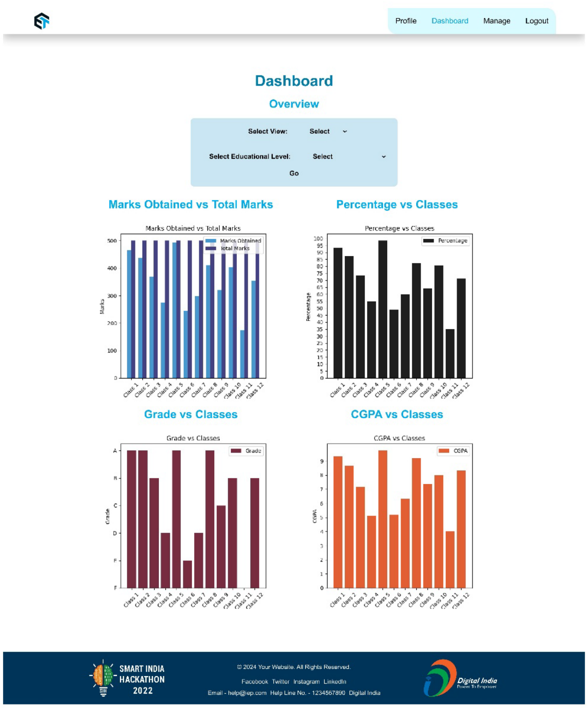

Figure 14. Dashboard

Upon successful login, the system directs the respective users (student, institute, and admin) to their personalised dashboards. The dashboard prominently displays a dedicated section for performance overview. This section displays a student's overall GPA and academic standing in the class using simple visuals, such as bar charts.

The system empowers student users with report generation functionalities. Reports generated for students provide a comprehensive overview of their performance in each enrolled class. This information is presented in a clear and concise format, potentially including personal details, key performance metrics (grades, marks obtained, total marks, and percentage).

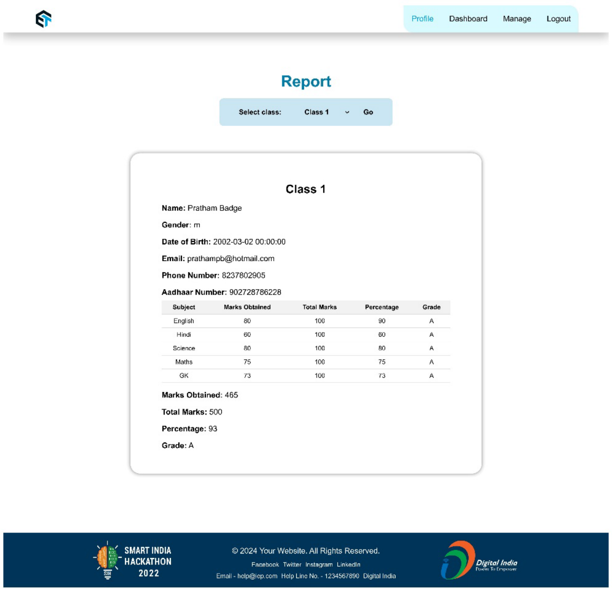

Figure 15. Report page

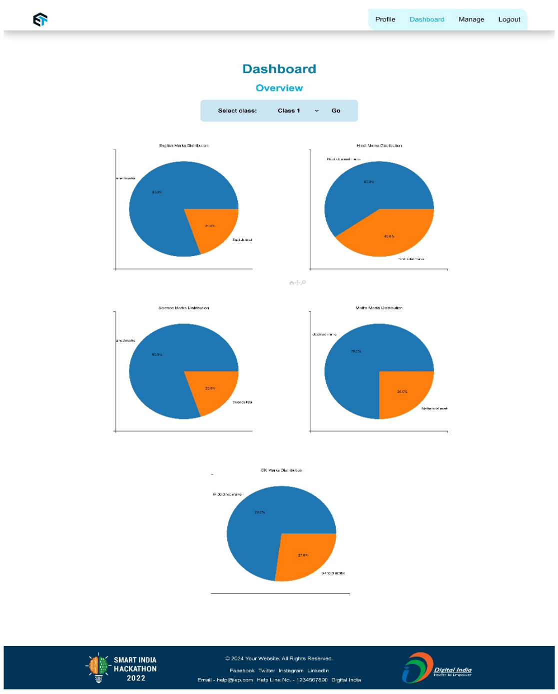

Figure 16. Analysis page

A pie chart can effectively depict the breakdown of marks obtained by a student in a particular subject relative to the total possible marks for that subject. The pie chart is divided into slices, each representing a specific portion of the whole. In this case, one slice would represent the total marks obtained by the student in the subject, while the remaining slice would represent the difference between the obtained marks and the total possible marks. This approach raises awareness and fosters a data-driven approach to learning, ultimately leading to improved student achievements.

In the described scenario, the figure shows the comparison graph plot of prediction vs. actual data. When the predicted academic performance from the linear regression model was evaluated with the actual data that was to be predicted for testing, the accuracy was about 70%–87% accurate and close to the actual data that was predicted based on a specific database for class and undergraduate and postgraduate students' data. MSE (Mean Squared Error), RMSE (Root Mean Squared Error), MAE (Mean Absolute Error), and R-Squared were used as evaluation metrics to assess the precision of the predictions.

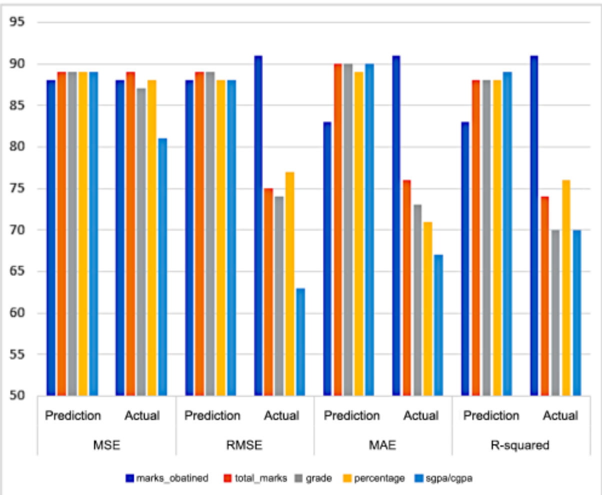

Figure 17. Prediction vs Actual data

For dropout analysis: We use a K-Means Clustering Model to divide student clusters into two groups: potential dropouts and non-dropouts. By analysing the clusters formed by K-means, we understand which group corresponds to potential dropouts (i.e., the group with the lower predicted grades). We use the predicted grades and convert them into binary labels that indicate whether they are less than 50% (1 for less than 50% and 0 for 50% or higher). Assign each student to the potential dropout group or the non-dropout group based on the cluster they belong to. Check how well the K-means clustering model predicts dropouts by looking at the predicted dropout labels and comparing them to the real dropout status (gap years), if that information is available. We also examine other characteristics of the clusters to gain insights into factors contributing to dropout likelihood shown in the figure. Furthermore, we analyse the dropout rates based on gender to obtain a more comprehensive understanding.

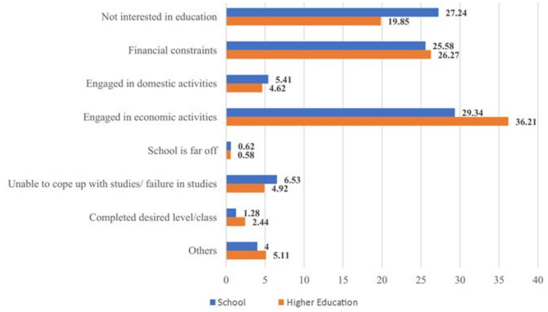

Figure 18. Primary Causes of Male Student Dropout

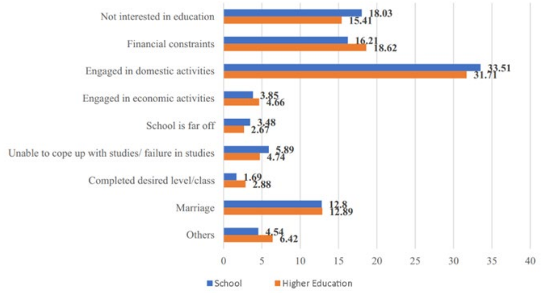

Figure 19. Primary Causes of Female Student Dropout

8. Conclusions and future work

In conclusion, this study proposes an innovative educational system that monitors, analyses, and visualises students' academic progress. The platform covers a student's journey from primary school to higher and technical education and provides a comprehensive overview of their academic performance. It provides valuable insights for students, teachers and educational institutions, empowering them to make knowledgeable decisions regarding their educational and career paths. Teachers can adapt their teaching strategies, while institutions can implement evidence-based practices. Key features include data tracking and analysis capabilities facilitated by Python libraries, as well as data visualisation capabilities. The platform also offers custom reports tailored to individual student profiles, providing actionable insights for educators and institutions. This connected educational platform represents a significant advancement in student-centred education and promotes a data-driven approach to academic success.

Future work: The suggested system should investigate machine learning techniques, create algorithms for personalised learning paths, and use sophisticated predictive analytics models to forecast student performance and dropout risk. We should investigate intervention strategies to reduce the risk of dropout, conduct long-term research, and collaborate with stakeholders to gather feedback. All students, including those with specific learning needs or disabilities, should be able to access the platform, and it should be compatible with the current learning management systems (LMS). The platform will encourage policy changes and assist in identifying trends. With this application, we can incorporate and implement the new National Education

Policy to improve education in a way that is more culturally rooted, skill-based, and inclusive.

Acknowledgements.
It is our pleasure to acknowledge our sincere thanks with a deep sense of gratitude to our project guide, "Dr. Nitesh Ghodichor," Assistant Professor, Computer Technology Department, for his continuous support in conducting this dissertation work. He has wholeheartedly helped us in this endeavour at all stages of the work. We are thankful to Dr S. A. Dhale, Principal, Priyadarshini College of Engineering, Nagpur, for providing the facilities at the institute. We thank Dr N. M. Thakare, Professor and Head of the Computer Technology Department of Priyadarshini College of Engineering, Nagpur. We herewith express our immense thanks to “Dr Nitesh Ghodichor”, Assistant Professor, Computer Technology Department, Priyadarshini College of Engineering, Nagpur, for giving us suggestions and coordinating with us from time to time as the project in charge. We also take the opportunity to thank all who have directly or indirectly extended help and encouragement in executing this project.

References

[1] Amal Alhadabi & Aryn C. Karpinski (2020) Grit, self-efficacy, achievement orientation goals, and academic performance in University students, International Journal of Adolescence and Youth, 25:1, 519-535, DOI: 10.1080/02673843.2019.1679202

[2] Nabil, Aya & Seyam, Mohammed & Abou-Elfetouh, Ahmed. (2021). Prediction of Students’ Academic Performance Based on Courses’ Grades Using Deep Neural Networks. IEEE Access. PP. 1- 1. 10.1109/ACCESS.2021.3119596

[3] Tadese M, Yeshaneh A, Mulu GB. Determinants of good academic performance among university students in Ethiopia: a cross-sectional study. BMC Med Educ. 2022 May 23;22(1):395. doi:10.1186/s12909-022-03461-0 PMID: 35606767; PMCID: PMC9125903.

[4] Chin-Wei Teoh, Sin-Ban Ho, Khairi Shazwan Dollmat, and Ian Chai. 2022. An Evolutionary Algorithm-Based Optimization Ensemble Learning Model for Predicting Academic Performance. In Proceedings of the 2022 11th International Conference on Software and Computer Applications (ICSCA '22). Association for Computing Machinery, New York, NY, USA, 102–107. https://doi.org/10.1145/3524304.3524320.

[5] Albreiki, B., Zaki, N., & Alashwal, H. (2021, September 16). A Systematic Literature Review of Student’ Performance Prediction Using Machine Learning Techniques. Education Sciences, 11(9), 552. ttps://doi.org/10.3390/educsci11090552

[6] G. Feng, M. Fan and Y. Chen, "Analysis and Prediction of Students’ Academic Performance Based on Educational Data Mining," in IEEE Access, vol. 10, pp. 19558-19571, 2022, doi: 10.1109/ACCESS.2022.3151652.

[7] Abdul Bujang, S. D., Selamat, A., Krejcar, O., Mohamed, F., Cheng, L. K., Chiu, P. C., & Fujita, H. (2023). Imbalanced Classification Methods for Student Grade Prediction: A Systematic Literature Review. IEEE Access, 11, 1970–1989. https://doi.org/10.1109/access.2022.3225404

[8] Chamberlin, Kelsey & Yasué, Maï & Chiang, I-Chant. (2018). The impact of grades on student motivation. Active Learning in Higher Education. 24. 10.1177/1469787418819728.

[9] Bucos, Marian & Drăgulescu, B.. (2018). Predicting student success using data generated in traditional educational environments. TEM Journal. 7. 617-625. 10.18421/TEM73-19.

[10] Sadasivam, R., Paramasivam, S., Prakash raj, N., & Saravanan, M. (2022). Students career prediction. International Journal of Health Sciences, 6(S5), 1357–1365. https://doi.org/10.53730/ijhs.v6nS5.8883

[11] Garg, Mausam & Chowdhury, Poulomi & Sheikh, Illias. (2023). Determinants of school dropouts in India: a study through survival analysis approach. Journal of Social and Economic Development. 10.1007/s40847-023-00249-w.

[12] Mr. Basavaraj Biradar, Dr. S.Y. SWADI, Reviews on school dropouts and research gap. International Journal of Research and Analytical Reviews (IJRAR).

[13] Kirazoğlu, Cem. (2009). The investigation of school-dropout at the secondary level of formal education: the stated reasons by the school administrators and school counselors: a preliminary study. Procedia - Social and Behavioral Sciences. 1. 905-914. 10.1016/j.sbspro.2009.01.161.

[14] Buchhorn, J., & Wigger, B. U. (2021). Predicting Student Dropout: A Replication Study Based on Neural Networks. SSRN Electronic Journal. https://doi.org/10.2139/ssrn.3929194

[15] Garcia-Zanabria, G.; Gutierrez-Pachas, D.A.; Camara-Chavez, G.; Poco, J.; Gomez-Nieto, E. SDA-Vis: A Visualization System for Student Dropout Analysis Based on Counterfactual Exploration. Appl. Sci. 2022, 12, 5785. https://doi.org/ 10.3390/app12125785

[16] Ladeira, Pedro & de Lima, Leandro & Krohling, Renato. (2022). A visualization tool for data analysis on higher education dropout: a case study at UFES.