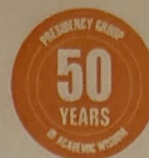




PRESIDENCY UNIVERSITY

Private University Estd. in Karnataka State by Act No. 41 of 2013

Itgalpura, Rajankunte, Yelahanka, Bengaluru – 560064



STUDENT DROPOUT ANALYSIS FOR SCHOOL EDUCATION

A PROJECT REPORT

Submitted by

TANUSHREE R - 20221CBD0029

KAVYA J - 20221CBD0023

KAVAYA S - 20221CBD0021

Under the guidance of,

Dr. Abdul Majid

BACHELOR OF TECHNOLOGY

IN

**COMPUTER SCIENCE AND TECHNOLOGY,
BIG DATA**

PRESIDENCY UNIVERSITY

BENGALURU

DECEMBER 2025



PRESIDENCY UNIVERSITY

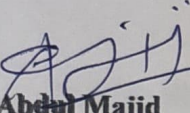
Private University Estd. in Karnataka State by Act No. 41 of 2013
Itgalpura, Rajankunte, Yelahanka, Bengaluru - 560064



PRESIDENCY SCHOOL OF COMPUTER SCIENCE AND ENGINEERING

BONAFIDE CERTIFICATE

Certified that this report "STUDENT DROPOUT ANALYSIS FOR SCHOOL EDUCATION" is a Bonafide work of "TANUSHREE R (20221CBD0029), KAVYA J (20221CBD0023), KAVYA S (20221CBD0021)", who have successfully carried out the project work and submitted the report for partial fulfilment of the requirements for the award of the degree of BACHELOR OF TECHNOLOGY in COMPUTER SCIENCE AND TECHNOLOGY, BIG DATA during 2025-26.


Dr. Abdul Majid

Project Guide

PSCS

Presidency University


Dr. Manjula H M

Program Project

Coordinator

PSCS

Presidency University


Dr. Sampath A K

School Project

Coordinators

PSCS

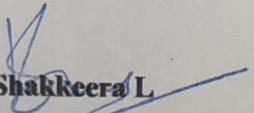
Presidency University


Dr. Pravinth Raja

Head of the Department

PSCS

Presidency University


Dr. Shakkeera L

Associate Dean

PSCS

Presidency University

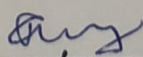
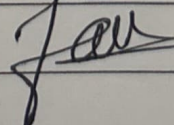

Dr. Durairam N

Dean

PSCS & PSIS

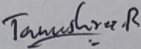
Presidency University

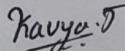
Examiners

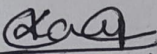
Sl. no.	Name	Signature	Date
1.	Ramannurthy. K		3/12/2025
2.	Pakurababu. B		3/12/25

DECLARATION

We the students of final year B.Tech in COMPUTER SCIENCE AND TECHNOLOGY, BIG DATA at Presidency University, Bengaluru, named TANUSHREE R, KAVYA J, KAVYA S, hereby declare that the project work titled "**STUDENT DROPOUT ANALYSIS FOR SCHOOL EDUCATION**" has been independently carried out by us and submitted in partial fulfilment for the award of the degree of B.Tech in COMPUTER SCIENCE AND TECHNOLOGY, BIG DATA during the academic year of 2025-26. Further, the matter embodied in the project has not been submitted previously by anybody for the award of any Degree or Diploma to any other institution.

Tanushree R USN: 20221CBD0029 

Kavya J USN: 20221CBD0023 

Kavya S USN: 20221CBD0021 

PLACE: BENGALURU

DATE: 3 DECEMBER 2025

ACKNOWLEDGEMENT

For completing this project work, We/I have received the support and the guidance from many people whom I would like to mention with deep sense of gratitude and indebtedness. We extend our gratitude to our beloved **Chancellor, Pro-Vice Chancellor, and Registrar** for their support and encouragement in completion of the project.

I would like to sincerely thank my internal guide **Dr. Abdul Majid, Professor**, Presidency School of Computer Science and Engineering, Presidency University, for his/her moral support, motivation, timely guidance and encouragement provided to us during the period of our project work.

I am also thankful to **Dr. Pravinth Raja, Professor, Head of the Department**, Presidency School of Computer Science and Technology Presidency University, for his mentorship and encouragement.

We express our cordial thanks to **Dr. Duraipandian N**, Dean PSCS & PSIS, **Dr. Shakkeera L**, Associate Dean, Presidency School of computer Science and Engineering and the Management of Presidency University for providing the required facilities and intellectually stimulating environment that aided in the completion of my project work.

We are grateful to **Dr. Sampath A K**, and **Dr. Geetha A**, PSCS Project Coordinators, **Dr. Manjula H M**, Program Project Coordinator, Presidency School of Computer Science and Engineering, for facilitating problem statements, coordinating reviews, monitoring progress, and providing their valuable support and guidance.

We are also grateful to Teaching and Non-Teaching staff of Presidency School of Computer Science and Engineering and also staff from other departments who have extended their valuable help and cooperation.

TANUSHREE R

KAVYA J

KAVYA S

ABSTRACT

School dropout continues to represent a critical obstacle to the realization of inclusive and equitable education across India. With national dropout rates standing alarmingly high at the secondary level the traditional reliance on retrospective analysis is insufficient for timely intervention. This project addresses this critical challenge by proposing and validating a novel, data-driven methodology: a dynamic, three-tiered web-based platform for predictive school dropout analysis.

The system is architected around clear functional separation: the Presentation Layer utilizes python framework Streamlit that provides an accessible user interface; the Application Layer hosts the business logic and the predictive engine, specifically a Random Forest classifier; and the Data Layer manages the persistent storage of both training and user-submitted data. The platform offers dual modes of intelligence: displaying comprehensive pre-trained analytics segmented by school, age, area, and caste for trend analysis, and crucially, allowing users to input new student data via direct forms for real-time risk prediction. The choice of the Random Forest algorithm is justified by its inherent robustness and, more importantly, its capability for feature importance analysis, providing diagnostic insights into the causes of attrition.

The model evaluation demonstrated a high degree of predictive reliability, achieving a Test Accuracy of 78.95% and a Cross-validation AUC of 0.8280 on the test dataset. In the binary classification task, the model exhibited excellent performance in identifying enrolled students and acceptable performance for the minority, at-risk class. Feature Importance analysis revealed that the academic and engagement factors are the most significant predictors of student attrition: the composite variable Attendance Score Interaction (0.182), Previous Academic Score (0.151), and raw Attendance (0.097) dominated the risk landscape. This finding is further supported by feature correlation analysis, which showed that Attendance Score Interaction ($r=-0.371$) and Previous Score ($r=-0.326$) have the strongest inverse relationship with dropout, underscoring that declining academic engagement is the primary precursor to leaving school.

By transforming raw data into targeted, actionable risk profiles, this web-based platform empowers educators and policymakers to allocate resources efficiently, design personalized interventions, and thus accelerate the journey towards a truly inclusive educational environment.

TABLE OF CONTENT

Sl. No.	Title	Page No.
	Declaration	iii
	Acknowledgement	iv
	Abstract	v
	List of Figures	vi
	List of Tables	vii
	Abbreviations	viii
1.	Introduction 1.1 Background 1.2 Statistics of project 1.3 Prior existing technologies 1.4 Proposed approach 1.5 SDGs 1.6 Overview of project report	1 to 5
2.	Literature review 2.1 Review of existing models 2.2 Research Gaps 2.3 Objectives	6 to 12
3.	Methodology	13 to 16
4.	Project management 4.1 Project timeline 4.2 Risk analysis 4.3 Project budget	17 to 22
5.	Analysis and Design 5.1 Requirements 5.2 Block Diagram 5.3 System Flow Chart 5.4 Choosing devices 5.5 Designing units 5.6 Standards	23 to 31

	5.7 Mapping with IoTWF reference model layers 5.8 Domain model specification 5.9 Communication model 5.10 IoT deployment level 5.11 Functional view 5.12 Mapping IoT deployment level with functional view 5.13 Operational view 5.14 Other Design	32 to 36
6.	Hardware, Software and Simulation 6.1 Hardware 6.2 Software development tools 6.3 Software code 6.4 Simulation	37 to 40
7.	Evaluation and Results 7.1 Test points 7.2 Test plan 7.3 Test result 7.4 Insights	41 to 43
8.	Social, Legal, Ethical, Sustainability and Safety Aspects 8.1 Social aspects 8.2 Legal aspects 8.3 Ethical aspects 8.4 Sustainability aspects 8.5 Safety aspects	44 to 51
9.	Conclusion	52
	References	54
	Base Paper	56
	Appendix	57

LIST OF FIGURES

Figure	Caption	Page No
Fig 1.5	Sustainable development goals	4
Fig 3.1	V-Model Methodology	13
Fig 3.4	Conceptual Diagram of the Three-Tier Web Application Architecture	16
Fig 5.2	Functional Block Diagram	25
Fig 5.3	System Flow Chart for Real-Time Dropout Prediction	27
Fig 5.7	The Request-Response Model involves three core entities	32
Fig. 5.9	IoT Deployment	34
Fig. 6.4	Dashboard with Results	40
Fig. 7.3	Feature Importance	42
Fig. 7.4	Prediction and Recommendation Results	43

LIST OF TABLES

Table	Caption	Page No
Table 1.5	Sustainable development goals Alignment	4
Table 2.1	Summary of Literature reviews	10
Table 3.3	Mapping to the V-Model	14
Table 4.1.1	Project Planning and Design Timeline (Weeks 1-7)	17
Table 4.1.2	Project Implementation and Validation Timeline (Weeks 6-15)	18
Table 4.2	Example of PESTEL analysis	19
Table 5.1.1	Summarizing requirements	23
Table 5.1.2	SW and HW requirements	24
Table 5.5	Mapping Project layers with IoTWFRM	29
Table 5.10	Operational view	35
Table 7.1	Identifying Test Points	41
Table 7.2	Observations of Model Unit	42
Table 7.3	Performance Evaluation	43
Table 8.4	Economic and Ecological Sustainability Principles	49

ABBREVIATIONS

Abbreviation	Full Form
AUC	Area Under the Curve
API	Application Programming Interface
ANN	Artificial Neural Network
CSV	Comma-Separated Values
DFD	Data Flow Diagram
EDA	Exploratory Data Analysis
F1	F1-score
GPU	Graphics Processing Unit
HTTPS	Hypertext Transfer Protocol Secure
IDE	Integrated Development Environment
IoT	Internet of Things
IoTWF	Internet of Things World Forum
ML	Machine Learning
NoSQL	Not Only SQL (Database Type)
RTE Act	Right to Education Act
SDG	Sustainable Development Goal
SSL/TLS / TLS	Transport Layer Security
UI	User Interface
VCS	Version Control System
VS Code	Visual Studio Code