Karnatak Law Society's
GOGTE INSTITUTE OF TECHNOLOGY
Udyambag Belagavi -590008
Karnataka, India.

A Course Project Report on

# "Student Performance Analysis and Prediction"

Submitted for the requirements of 6th semester B.E. in CSE

for **"Artificial Intelligence and Machine Learning (21CS63)"**

**Submitted by**


| NAME | USN |
|---|---|
| 1) Suraj Savant | 2GI22CS415 |
| 2) Takshak Chavalagi | 2GI22CS416 |
| 3) Varun M Golai | 2GI22CS417 |
| 4) Vinayak R Lamani | 2GI22CS418 |

**Under the guidance of**

**Prof. Namitha Bhat**

**Professor, Dept. of CSE**


**Academic Year 2022-2023 (Even semester)**

Karnatak Law Society's

GOGTE INSTITUTE OF TECHNOLOGY

Udyambag Belagavi -590008

Karnataka, India.

**Department of Computer Science and Engineering**

# Certificate

This is to certify that the Course Project work titled **"Student Performance Analysis and Prediction"** carried out by **Suraj Savant,Takshak Chavalagi,Varun M Golai,Vinayak R Lamani** bearing **USNs 2GI22CS415, 2GI22CS4116, 2GI22CS417, 2GI22CS418** for AIML **(21CS63)** course is submitted in partial fulfilment of the requirements for 6th semester B.E. in **COMPUTER SCIENCE AND ENGINEERING,** Visvesvaraya Technological University, Belagavi. It is certified that all corrections/ suggestions indicated have been incorporated in the report. The course project report has been approved as it satisfies the academic requirements prescribed for the said degree.

Date: 07/06/2024

Place: Belagavi

Signature of Guide

Prof. Namitha Bhat

Asso., Prof., Dept. of CSE

KLS Gogte Institute of Technology, Belagavi

Karnatak Law Society's

GOGTE INSTITUTE OF TECHNOLOGY

Udyambag Belagavi -590008

**Academic Year 2022-23 (Even Semester)**

**Semester: IV**

**Course: AIML (21CS63)**

## Rubrics for evaluation of Course Project

| S. No | Project Component | Max. Marks | Marks Earned | | | |
|---|---|---|---|---|---|---|
| | | | 2GI22CS415 | 2GI22CS416 | 2GI22CS417 | 2GI22CS418 |
| | | | Suraj S | Takshak C | Varun M G | Vinayak R L |
| 1 | Relevance of the project and its objectives | 01 | | | | |
| 2 | Tools/Framework used | 01 | | | | |
| 3 | Methodology / Design | 02 | | | | |
| 4 | Implementation and Results | 03 | | | | |
| 5 | Project Report | 03 | | | | |
| | **Total** | **10** | | | | |

# ACKNOWLEDGEMENTS

We wish to express my gratitude to KLS Gogte Institute of Technology for providing us opportunity to Learn course entitled AIML. It was a great chance for industrial exposure and a great chance for learning and professional development. We would like to hereby acknowledge their valuable assistance in providing suggestions, constant guidance or abiding with us throughout our project.

We take this opportunity to express my profound gratitude and deep regard to all my guides for their exemplary guidance, monitoring and constant encouragement throughout the course of this project. The blessing, help and guidance given by them time to time have been a constant source of inspiration.

Further I would like to extend our gratitude **to Prof. Namitha Bhat** for her complete and constant unconditional support and her suggestions in maintaining the flow of this report. We thank herfor developing the analytical ability to a larger extent that can provide impetus to all future undertakings

# Table of contents

# 1. ABSTRACT

In the evolving landscape of education, the ability to accurately analyze and predict student performance has become increasingly vital. This report explores the use of data analytics and machine learning techniques to assess and forecast student outcomes based on various influencing factors. By leveraging a comprehensive dataset that includes demographics, academic records, and socio-economic indicators, we aim to identify key determinants of student success and develop robust predictive models.

Our approach begins with data collection and preprocessing, ensuring the integrity and usability of the data. This is followed by exploratory data analysis (EDA) to uncover underlying patterns and relationships within the dataset. Feature engineering techniques are then applied to enhance the predictive power of the models. A dedicated analysis of the Gender column is conducted to examine performance disparities between male and female students.

Multiple machine learning algorithms, including Linear Regression, Decision Trees, Random Forest, Support Vector Machines (SVM), and Neural Networks, are employed to build predictive models. These models are rigorously evaluated using metrics such as Mean Absolute Error (MAE), Root Mean Squared Error (RMSE), and R-squared ($R^2$) for regression tasks, as well as accuracy, precision, recall, and F1-score for classification tasks.

The results indicate that factors such as attendance, parental education level, and socio-economic status significantly impact student performance. Notably, the Random Forest model emerged as the most effective predictor, highlighting the importance of ensemble learning methods in educational data mining.

In conclusion, this study demonstrates the potential of data-driven approaches in enhancing educational strategies and outcomes. The insights derived from this analysis can guide educators in implementing targeted interventions and fostering an environment conducive to academic success. Future research should explore the integration of additional data sources and the application of more advanced machine learning techniques to further improve predictive accuracy.