



**IJIRCCCE**

e-ISSN: 2320-9801 | p-ISSN: 2320-9798



# INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH


IN COMPUTER & COMMUNICATION ENGINEERING

Volume 12, Issue 5, May 2024

**ISSN** INTERNATIONAL  
STANDARD  
SERIAL  
NUMBER  
INDIA

**Impact Factor: 8.379**

 9940 572 462

 6381 907 438

 [ijircce@gmail.com](mailto:ijircce@gmail.com)

 [www.ijircce.com](http://www.ijircce.com)

# Placement and Salary Predictor with Job Recommendation App

Hrushikesh Dunde<sup>1</sup>, Mohd Salman<sup>2</sup>, Umaid Shaikh<sup>3</sup>, Tanmany Lone<sup>4</sup>

Department of Information Technology, Terna Engineering College, Mumbai University, Navi, Mumbai, India

**ABSTRACT:** The Placement and Salary predictor with job recommendation is an android app which is useful for students as well as the college faculty/TPO member and other academic staff.

In paper based system all the works are done manually. Paper based work is so time consuming. The main objective behind the use of this project is to reduce time. Through our application TPO/Faculties can predict the placement of students using their overall grades in different course and on the basis of various features of college dataset. It can be done using AIML algorithms and NLP. All information is stored in the secured database. Placement of students is one of the most important objectives of an educational institution. Students' academic achievement and their placements in companies selection is a difficult issue in the current manual system. Reputation and yearly admission of an institution are dependent upon the placement chances of a student. It also improves the placement percentage rate. The objective of the project is to analyze the previous year's dataset and predict the capability of the current year students. According to their performance the study used Naïve Bayes, Decision Tree, and Random Forest Algorithm to build the prediction model for placement of students. The model is built by both training and test set which gives accuracy in prediction. Placement of scholars is one of the vital activities in academic establishments. Admission and name of establishments primarily depend on placements. Hence all institutions strive to Strengthen the placement department.

## I. INTRODUCTION

The primary aim of this project is to analyze historical data of previous years' students to forecast current students' placement prospects, thereby enhancing the overall placement percentage of institutions. This project introduces a recommendation system that predicts whether a current student will secure placement, and if so, predicts the potential company based on data from previously placed students. We utilize two distinct data science classification algorithms, namely Random Forest and Decision Tree, independently predicting outcomes, and subsequently comparing their efficiencies based on the dataset. This model assists placement cells within organizations to identify potential students, focusing on enhancing their technical and social skills.

Placements are undeniably crucial for every college, as the success of a college is often gauged by its campus placement statistics. Prospective students often base their college choices on these placement percentages. Thus, predicting and analyzing placement trends is vital for colleges to enhance their reputation and improve students' chances of placement. The Placement Prediction system forecasts the probability of a student securing placement in a company through the application of classification algorithms like Decision Tree and Random Forest. The primary objective is to predict whether a student will secure placement during campus recruitment, utilizing data such as academic history and skill sets. These algorithms are applied to historical data of previous years' students. Therefore, predicting and analyzing placement trends remains crucial for colleges to enhance their reputation and improve students' placement prospects.

This project also presents a salary prediction system using a profile of graduated students as a model. A data mining technique is applied to generate a model to predict a salary for individual students who have similar attributes to the training data. In this work, we also made an experiment to compare four data mining techniques including Decision trees, Random Forest, LinearRegression, Logistic Regression, to find the suitable technique to the salary prediction. In the experiment, 600 records of graduated student data were used. Results showed that Decision Tree provided the best efficiency to be used as a model for salary prediction. The result also informed that they found they satisfied with the implemented system since the system was easy to use, and the prediction results were simple to understand

## II.MOTIVATION

The biggest motivation for the design of the Placement and Salary Predictor with Job Recommendation App is for increasing students' motivation in studying, and to improve is done on the paper and those records are stored in a file. The retrieval of data becomes so difficult due to this approach. Mobile and Web application being scalable and accessible from everywhere due to the Internet motivated us to develop a system which would be able to achieve the goals for the development of this system and handle recruiting process easily.

## III.SCOPE

This system consists of Web API and Web app and Mobile-based applications and will be designed to predict the Placement of students and generate/predict salary also with various job suggestions according to the skills of students by using technology. According to previous year dataset the model is divided into trained and test dataset and model is trained on these datasets and of all the algorithms we obtained 100% accuracy using Decision Tree algorithm which is remarkable comparing to previous works and surveys of various research papers. Only after the student is predicted to be placed then only the system will predict the salary also based on the trained previous year data of Terna college and suggest the jobs using NLP After extracting skills from the resume of students and matching it with the Job Descriptions of various Companies through web scraping. Students can get to know if they can be placed or not having current skillsets and at what salary they can obtain and can apply also to their favorite companies in one single go/click. The system will have 4 types of users:

1. Students: They can register themselves.
2. Faculties: They can use this app to improve the skillset of students by teaching what is needed in the current industry.
3. TPOs: To shortlist the students based on the academic criteria or company criteria.
4. Admin: The admin team who has central control over the system.

## IV.OBJECTIVE

Upon successful registration, students, faculties, and TPOs have their details added to the database. They can then utilize the system to predict student placements and salaries, receive job suggestions tailored to their skills and interests, and apply for jobs seamlessly.

NLP and the Python library ResumeParser are employed for job recommendations, while various machine learning algorithms such as Decision Tree, Random Forest, KNN, Linear Regression, and Logistic Regression contribute to accurate predictions.

The system's backend, featuring ML algorithms and datasets, is deployed on the Heroku website via a Python API, interacting with the Android app through Flask. Additionally, a web app hosted on PythonAnywhere on the cloud enhances accessibility.

## V.LITERATURE SURVEY

Several research studies highlight the significance of educational data mining (EDM) in predicting students' academic performance using various machine learning techniques. These studies focus on analyzing past academic data to predict future performance accurately, aiding in early intervention and academic improvement strategies. Techniques such as decision trees, over-sampling, and sparse linear models have shown promise in enhancing prediction accuracy and informing personalized educational pathways.

[1] **Al-Barrak et al.** proposed that educational data mining is the process of applying data mining tools and techniques to analyze data at educational institutions. In this paper, we used educational data mining to predict students' final GPA based on their grades in previous courses. In our case study, we collected students' transcript data that included their final GPA and their grades in all courses. After pre-processing the data, we applied the J48 decision tree algorithm to discover classification rules. We extracted useful knowledge for final GPA, and identify the most important courses in the students' study plan based on their grades in the mandatory courses.

[2] **Abana et al** proposed that The educational data mining (EDM) can be specified as one of the main fields related to



high-quality research that involves mining datasets to address research questions related to education; such questions examine the ways in which people learn and teach. a large amount of data, including education data, are being collected, and much of them are unprocessed. The success of EDM was examined in this paper, and nine data mining techniques were explored including: bagging, multilayer perception (MLP), naïve Bayes (NB), K-nearest neighbours (KNN), logistic regression (LR), support vector machine (SVM), XGBoost, decision tree (DT), and random forest (RF).

[3] **Ahmed et al** proposed that In this paper, a multi-class prediction model using five machine learning algorithms was compared to predict final students' grades based on the previous student final examination result of the first-semester course. We successfully presented the importance of F.S. and SMOTE methods that can efficiently give different prediction results to improve student grade prediction. Interestingly, we discovered that the prediction accuracy performance result was improved consistently when using SMOTE than F.S. for all predictive models. Performing F.S. with SMOTE significantly impacts the performance of the student grade prediction.

[4] **Khan et al.** proposed that large amount of digital data is being generated across a wide variety of fields and Data Mining (DM) techniques are used transform it into useful information so as to identify hidden patterns. One of the key areas of the application of Education Data Mining (EDM) is the development of student performance prediction models that would predict the student's performance in educational institutions. We build a model which can notify students (in introductory programming course) about their probable outcomes at an early stage of the semester (when evaluated for 15% grades). We applied 11 Machine Learning algorithms (from 5 categories) over a data source using WEKA and concluded that Decision Tree (J48) is giving higher accuracy in terms of correctly identified instances, F-Measure rate and true positive detections. This study will help to the students to identify their probable final grades and modify their academic behavior accordingly to achieve higher grades.

[5] **Anderson et al.** proposed that With ever growing pressure to increase student performance, as well as the need to support first generation students and those from underrepresented backgrounds, never before has there been a greater need to quickly identify students who are on track to perform poorly in a class. Traditionally, instructors have relied mainly on intuition and rough averaging of exam scores to predict what grade a student who is on track to receive.

[6] **Iqbal et al** proposed that In higher educational institutes, many students have to struggle hard to complete different courses since there is no dedicated support offered to students who need special attention in the registered courses. Machine learning techniques can be utilized for students' grades prediction in different courses. Such techniques would help students to improve their performance based on predicted grades and would enable instructors to identify such individuals who might need assistance in the courses.

[7] **Polyzou et al.** proposed that the accurate estimation of students' grades in future courses is important as it can inform the selection of next term's courses and create personalized degree pathways to facilitate successful and timely graduation. This paper presents future-course grade predictions methods based on sparse linear models and low-rank matrix factorizations that are specific to each course or student-course tuple.

[8] **Jishan et al.** proposed that the data can be preprocessed using a discretization method called the Optimal Equal Width Binning and an over-sampling technique known as the Synthetic Minority Over-Sampling (SMOTE) to improve the accuracy of the students' final grade prediction model for a particular course. In order to validate our method we have used data from a course offered at North South University, Bangladesh. The result obtained from the experiment gives a clear indication that the accuracy of the prediction model improves significantly when the discretization and over-sampling methods are applied.

Author	Paper Name	Sample Size	Data Source	Attributes	Algorithm	Best Performance	Limitations
Barrak et al	Predicting Students Final GPA Using Decision Trees: A Case Study (2016)	236	A female student from Computer Sciences College at King Saud University 2012	Student name, Student ID, final GPA, the semester of graduation, major, nationality, campus, courses are taken and course grade	DT (J48)	DT (J48)	Lack of experimental techniques for prediction
Abana	characteristics of data mining by classification educational dataset to improve student's evaluation(2021)	133	Students of Computer Engineering program in 4 years	Research Method (RM) grade, Research Project (RP) grade, gender, backlog, programming proficiency	RT, RepTree DT (J48)	RT Accuracy 75.2%	Lack of experimental techniques for prediction
Ahmad et al	Multiclass Prediction Model for Student Grade Prediction Using Machine Learning (2021)	399	First-year bachelor students in Computer Science at UniSZA from 2006/2007 to 2013/2014	GPA, race, gender, family income, university entry mode, Malaysia Certificate of Education (SPM) grade in 3 subjects	DT, NM and Rule Based (PART)	RB Accuracy 71.3%	A small number of dataset due to incomplete and missing value
Khan et al	Tracking Student Performance in Introductory Programming by Means of Machine Learning. (2019)	50	Student of Buraimi University College, 50 Oman	Test 1 marks, CGPA, Attendance, Major, Gender, Year	NB, MLP, SVM, Lazy (IBK), Rules-Based (Decision Table, JRIP, OneR, PART and ZeroR) D.T. ((J48), R.F.. RT	DT (J48) (Feature Selection+SMOTE) Accuracy 88%	Small number of dataset
Anderson et al	APPLICATIONS OF MACHINE LEARNING TO STUDENT GRADE PREDICTION IN QUANTITATIVE BUSINESS COURSES	683	Students of Craig School of Business at California State University, Fresno from 2006 to 2015	Historical grade data from 18 Semester	NB, KNN, SVM	SVM	Some of the dataset are not available due to significant changes.



	(2017)						
Iqbal et al.	Machine Learning Based Student Grade Prediction: A Case Study (2017)	225	Undergraduate students of the Electrical 225 Engineering Program from 2013 to 2015	Grades, GPA	CF, M.F., RBM	RBM	Use limited attribute for analysis
Polyzou et al.	Grade Prediction with Course and Student Specific Models (2016)	76,748	Student-course grade from 2002 to 2014	Historical student course grade information	LinReg, MF	LinReg	Not support a large number of latent factors
Jishan et al.	Improving accuracy of students' final grade prediction model using optimal equal width binning and synthetic minority over-sampling technique (2015)	180	Student Core Course offered at North South 180 University, Bangladesh	CGPA, Quiz, Midterm, Lab, Attendance, Final grade	N.B., D.T., Neural Network Backpropagation with oversampling (SMOTE) and optimal binning	NB (optimal binning+SMOTE) Accuracy 75.28%	Small size of attributes that lead to high missclassification error

## VI. PROBLEM STATEMENT

In the current recruitment system, recruiters only have access to candidate information provided in their resumes. The proposed system aims to address this limitation by employing various algorithms to provide continuous values for student placement probability.

Utilizing machine learning, the system dynamically predicts candidates' placement probabilities based on parameters such as CGPA, HSC marks, and SSC marks.

Unlike the current system, which offers discrete values, this system offers a more accurate depiction of placement likelihood to candidates. Leveraging the Scikit learn module, the system utilizes algorithms like Random Forest Regressor to generate probabilities accurately, making it well-suited for handling large datasets.

Additionally, the system also caters to companies' preferences for candidates with specific skill sets, offering static predictions to aid in candidate selection based on skills and domains. Candidates can access a list of recommended jobs based on their skills.

## VII. PROPOSED METHODOLOGY

System Architecture

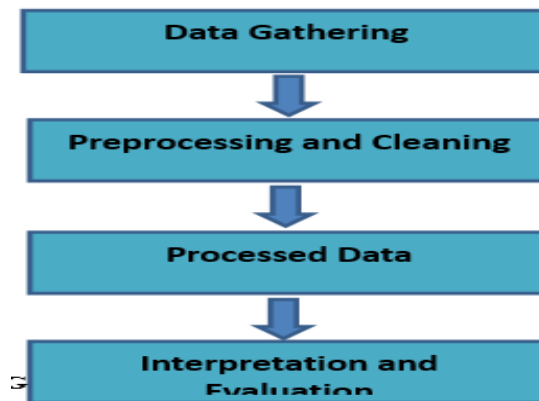


Figure 7.1: System Architecture

The First step is data gathering and we did the same as we collected the data in form of Google form responses of previous year students that is 2022 passed out students then we encoded the data that is preprocessing and cleaning of data then we trained the model on that dataset after one hot encoding, There is 12 features for evaluation.

VIII.BLOCK DIAGRAM

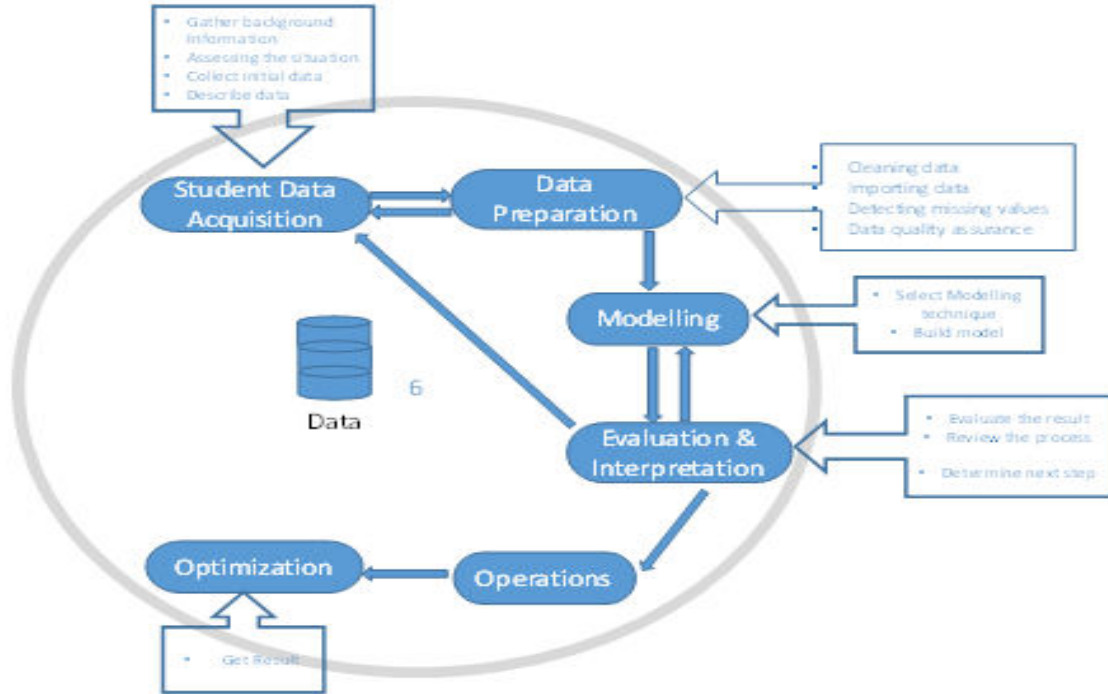


Figure 8.1: Design cycle of the project

Campus placement is a process where companies come to colleges and identify students who are talented and qualified before they finish their graduation. The proposed system determines the like hood of placement based on various attributes of the student’s profile. First, gather placement data of the previous year. Do the cleaning process on the data set. After cleaning detecting missing value, check data quality assurance. Next select modeling technique and build model. After this evaluate the result and get a result.

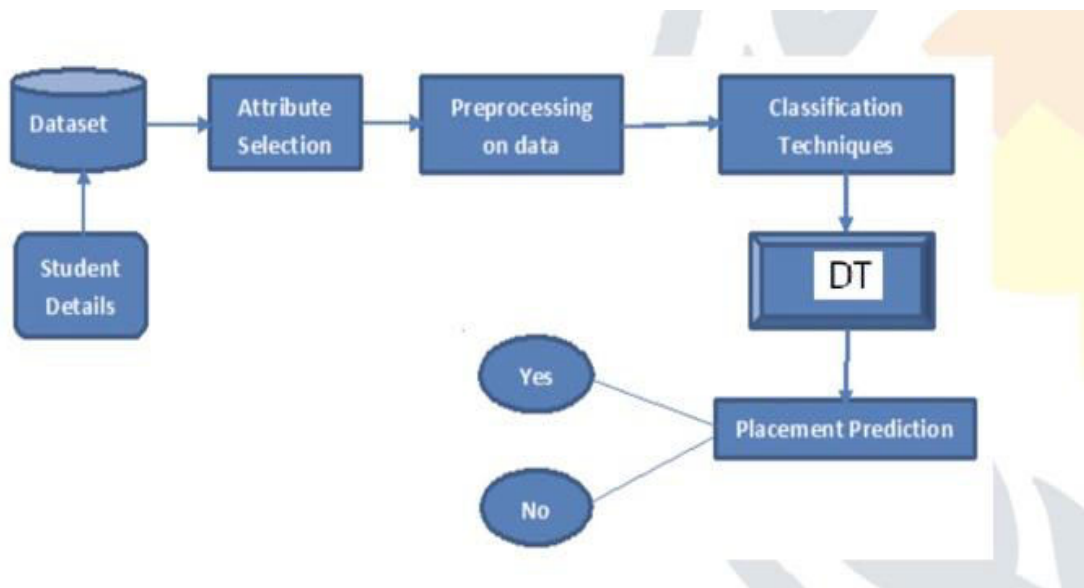


Figure 8.2: Block Diagram – Placement Prediction



Students fill up the details on form to respective college TPOs and then we encode the data for preprocessing and train the model on the same using Decision Tree (DT) and get 100% accuracy.

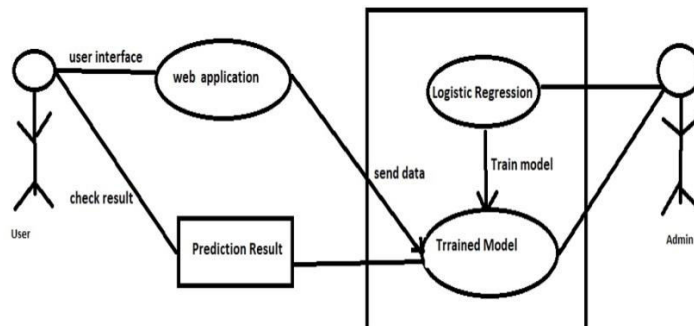


Fig 8.3: Use Case Diagram

### IX. PROPOSED METHODOLOGY

The proposed system, an Android application, serves as a recruitment tool for organizations to assess candidates' fitness for jobs effectively.

Colleges can leverage the system's Placement Prediction feature to gauge students' placement probabilities before placement drives.

The system recommends candidates based on various parameters such as SSC marks, HSC/Diploma marks, CGPA, skills, English fluency, and reading and writing proficiency.

Machine learning techniques, including Random Forest, Decision Tree Regressor, Logistic Regression, Linear Regression, KNN, and SVM algorithms, are employed for implementation.

The model is trained on previous placement drive datasets and used to predict candidates' placement probabilities. The internet and web technologies revolutionize information availability and user action, making the placement process manageable online.

Recruiters can use the system to shortlist candidates for further rounds based on obtained probabilities, maximizing placement probabilities and easing recruitment tasks.

### X. TECHNOLOGIES USED

- UI Development:
  - Android Studio for various layouts and XML for Android app
  - HTML CSS for web app
- Frontend Mobile Application Development:
  - Android Studio
- Frontend Web Application Development:
  - HTML, CSS, JavaScript
- Backend Development:
  - Python, Flask
  - MySQL, SQLite
  - ML, AI, NLP
- Database:
  - MySQL, SQLite
- Other Services:
  - Postman
  - APIs deployment on cloud (Heroku, PythonAnywhere)

- AI & ML, NLP
- ResumeParser Library

### XI.DESIGN, ANALYSIS & IMPLEMENTATION

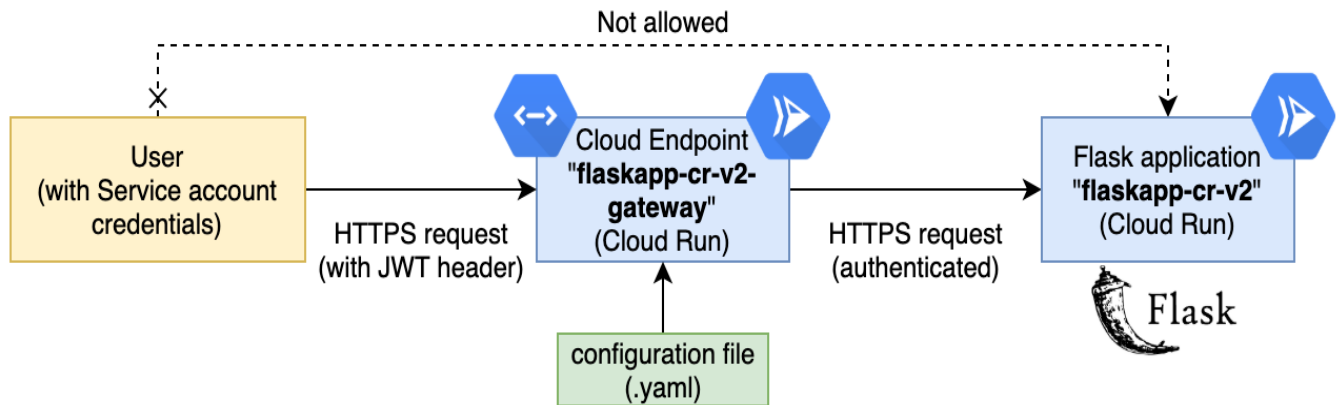


Figure 11.1: System Design

### XII.SOFTWARE ANALYSIS

#### WATERFALL MODEL / AGILE MODEL

Agile methodology denotes a software development approach characterized by iterative development. In Agile methods, tasks are segmented into smaller iterations, omitting direct long-term planning. Initially, project scope and requirements are delineated. A pre-defined plan outlines the number, duration, and scope of each iteration. This project adopts the Agile model for development.

In Agile, each iteration spans a short time frame, typically one to four weeks. Breaking down the project into smaller parts mitigates project risks and reduces delivery time requirements. Team members traverse the complete software development lifecycle in each iteration, encompassing planning, requirements analysis, design, coding, and testing.

### XIII.CONCLUSION

The primary objective of this system is to create a Placement and Salary Predictor with Job Recommendation app, integrating AIML technologies. Employing three machine learning algorithms per module for prediction, the system achieves a mean accuracy of 100%, surpassing previous work and enhancing reliability. The system's primary motivation lies in boosting students' motivation in studying, improving their skill sets, and providing insight into domain-specific skill requirements for early preparation. The development of mobile and web applications addresses the limitations of paper-based systems prevalent in some Indian companies, enhancing data accessibility and scalability.

### REFERENCES

1. Al-Barrak, M. A., & Al-Razgan, M. (2016). "Predicting students' final GPA using decision trees: a case study." *International journal of information and education technology*, 6(7), 528.
2. Abana et al. (2021). "Characteristics of data mining by classification educational dataset to improve students' evaluation." *Journal of Engineering Science and Technology*, 16(4), 2825-2844.
3. Ahmed, B., Selamat, A., & Krejcar, O. (2021). "Multiclass Prediction Model for Student Grade Prediction Using Machine Learning." *IEEE Access*, PP, 1-1.
4. Khan, I., Al Sadiri, A., Ahmad, A. R., & Jabeur, N. (2019). "Tracking Student Performance in Introductory Programming by Means of Machine Learning." *2019 4th MEC International Conference on Big Data and Smart City (ICBDSC)*, 1-6.
5. Anderson et al. (2017). "Applications of machine learning to student grade prediction in quantitative business

courses." *Global Journal of Business Pedagogy*, 1(3).

6. Iqbal, Z., Qadir, J., Mian, A., & Kamiran, F. (2017). "Machine Learning Based Student Grade Prediction: A Case Study."

7. Polyzou, A., & Karypis, G. (2016). "Grade Prediction with Course and Student Specific Models." *ArXiv*, abs/1906.00792.

8. Jishan, S. T., Rashu, R. I., Haque, N., et al. (2015). "Improving accuracy of students' final grade prediction model using optimal equal width binning and synthetic minority over-sampling technique." *Decision Analysis*, 2(1).

9. Thangavel, S. K., P, D. B., & Sankar, A. (2017). "Student Placement Analyzer: A Recommendation System Using Machine Learning." *2017 International Conference on Advanced Computing and Communication Systems (ICACCS - 2017)*, Jan. 06 – 07, 2017, Coimbatore, India.

10. Nichat, A. A., & Raut, A. B. (2017). "Predicting and Analysis of Student Performance Using Decision Tree Technique." *International Journal of Innovative Research in Computer and Communication Engineering*, 5(4).

11. Das, S., Barik, R., & Mukherjee, A. (2020). "Salary Prediction Using Regression Techniques." *SSRN Electronic Journal*.

12. Hung, C. C., & Lim, E. P. (2021). "On Aggregating Salaries of Occupations from Job Post and Review Data." *IEEE Access*, 9.

13. Kumar, N., et al. (2022). "Technical Job Recommendation System Using APIs and Web Crawling." *Computational Intelligence and Neuroscience*, 2022.

14. Shaikh, U., Siddiqui, I., Ali, Q., Buriro, S., & Babar, N. (2019). "NLP based service recommendation system."



INTERNATIONAL  
STANDARD  
SERIAL  
NUMBER  
INDIA



# INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN COMPUTER & COMMUNICATION ENGINEERING



9940 572 462



6381 907 438



ijircce@gmail.com



[www.ijircce.com](http://www.ijircce.com)

Scan to save the contact details