



PROJECT REPORT

Machine Learning  
  
Student Placement Prediction

|  |  |  |  |
| --- | --- | --- | --- |
| **Created By:** | Anurag Pareek | **Approved By:** | Sarang Kakadia |
| **Created On:** | 11-07-2024 | **Approved On:** | 12-07-2024 |

Page left blank intentionally

**INDEX**

[**1** **PROJECT DETAILS** 2](#_Toc143445375)

[**2** **SUMMARY** 2](#_Toc143445376)

[**3** **INTRODUCTION** 2](#_Toc143445377)

[3.1 Background 2](#_Toc143445378)

[3.2 Stakeholders 2](#_Toc143445379)

[3.3 Objectives 2](#_Toc143445380)

[**4** **METHODOLOGY** 2](#_Toc143445381)

[4.1 Considerations & Assumption 3](#_Toc143445382)

[4.2 Approach 3](#_Toc143445383)

[4.3 Activities 3](#_Toc143445384)

[**5** **TARGETTED V/S ACHIEVED OUTPUT** 3](#_Toc143445385)

[**6** **CONCLUSION** 3](#_Toc143445386)

[**7** **APPENDICES** 4](#_Toc143445387)

[7.1 Appendix A – Title 4](#_Toc143445388)

**General Instructions for using the Live Project Report Template**

* This template and the subsequent document created using this template is a confidential document and is the intellectual property of Cloud Counselage Pvt. Ltd. Circulating it outside of the organisation without the consent of Cloud Counselage Pvt. Ltd. is the breach of company policies and will lead to legal actions
* This template is a guideline document to communicate the implementation of design ideas and the results of the work to the stakeholders.
* The **text between inequality (< >) is to be replaced** by relevant text
* Please **remove the yellow highlight on the Text** between the inequality (< >). This is done to help you notice the text to be changed/replaced
* The text in *italics* highlighted in grey is just for reference and should be removed after adding the relevant text

# **PROJECT DETAILS**

|  |  |  |  |
| --- | --- | --- | --- |
| **Project Name** | Student Placement Prediction | | |
| **Project Sponsor** | Sarang Kadakia | | |
| **Project Manager** | Harshada Topale | | |
| **Start Date** | 17-06-2024 | **Completion Date** | 12-07-2024 |

# **SUMMARY**

This project aims to develop a predictive model to forecast two key aspects of a student's academic journey: the year of graduation and placement status. Using a comprehensive dataset of student information, the model leverages machine learning techniques to provide valuable insights for both students and educational institutions.

# **INTRODUCTION**

## Background

*A lack of clear academic planning and career advice is the root cause of the issue with students not knowing when they will graduate and not receive a placement. Students may have trouble finishing their coursework on time if they don't receive timely information on their graduation status. Furthermore, individuals could not be well prepared for the job market*

*without a priority of the placement-specific skills which could make it harder for them to get employment. and harm their prospects for successful careers in general*

## Stakeholders

*The success of a student placement prediction model hinges on effectively engaging a diverse group of stakeholders, including students, educational institutions, employers, faculty, alumni, industry experts, policymakers, and data privacy officers.*

*This includes conducting surveys and questionnaires for quantitative data, facilitating focus group discussions and one-on-one interviews for in-depth perspectives, organizing workshops and advisory boards for collaborative input, implementing user testing for practical feedback, analyzing historical data, and fostering industry partnerships. This multi-faceted engagement process ensures that the model addresses real-world needs, enhances its effectiveness and adoption, manages stakeholder expectations, and builds trust across the entire student placement ecosystem.*

## Objectives

The solution involves developing predictive model for student placement and calculation for year of graduation which accurately forecasts whether students secure placement and when students will graduate.

The year of graduation calculation will function by utilizing essential features and historical data. A dataset containing data on student’s college details including college name, academic year and branch will be used.

Student placement prediction will function by utilizing essential features and historical data. A dataset containing data on

student’s academic records, course progress, extracurricular activities, and previous placement results will be used to train the machine learning system. In order for the model to accurately forecast whether students secure placement. It must learn pattern and correlations from this data.

# **METHODOLOGY**

These conventions are all about the positions of line breaks, how many characters should go on a line, and everything in between.

## Considerations & Assumption

Creating a student placement prediction model comes with several constraints and challenges. Data availability and quality are primary concerns, as limited historical data, incomplete records, and privacy issues can hinder model development. Selecting relevant features that balance academic and non-academic factors is crucial but complex. The model must strike a delicate balance between accuracy and interpretability while avoiding overfitting on limited data. The rapidly evolving job market, with changing requirements and emerging roles, adds another layer of complexity. Ensuring fairness and avoiding bias based on protected characteristics is essential, as is addressing multiple potential outcomes such as job roles, companies, and salary ranges. Time sensitivity, external economic factors, and individual student preferences must also be considered. Finally, the model should be interpretable, providing clear explanations for its predictions to maintain transparency and trust among students and administrators. These interconnected challenges require careful consideration and innovative approaches to develop an effective and reliable student placement prediction model.

## Approach

To address the complex challenge of student placement prediction, I adopted a structured approach that combines data-driven analysis with domain expertise. This approach begins with comprehensive data collection and preprocessing, ensuring a robust foundation for the model. Feature engineering and selection techniques are then applied to identify the most relevant predictors, incorporating both academic performance and soft skills.  
  
 The core of the approach involves employing a flexible machine learning framework, such as ensemble methods or neural networks, capable of capturing complex relationships in the data. This is complemented by regular model validation and updating processes to adapt to evolving market trends.  
  
 Importantly, the approach incorporates interpretability techniques to provide transparent insights into predictions. It also includes fairness constraints to mitigate potential biases and ensure equitable outcomes across different student groups.  
  
 This structured method was chosen because it balances predictive power with practical applicability. It addresses the key challenges of data quality, model complexity, and evolving job markets while maintaining transparency and fairness.

## Activities

To deliver the student placement prediction project, I performed a series of structured activities:  
  
 1. Requirement Gathering: I conducted surveys, interviews, and focus group discussions with key stakeholders including students, faculty, career services staff, and industry partners. This helped define the project scope and identify critical features for the prediction model.  
  
 2. Planning: Based on the gathered requirements, I developed a comprehensive project plan outlining timelines, milestones, resource allocation, and potential risks. This included creating a work breakdown structure and Gantt chart to visualize the project timeline.  
  
 3. Data Collection and Preprocessing: I worked on collecting historical student data, including academic records, extracurricular activities, and previous placement outcomes. This involved data cleaning, handling missing values, and standardizing formats to ensure data quality.  
  
 4. Feature Engineering and Selection: I analyzed the collected data to identify relevant features for the prediction model. This included creating new features from existing data and selecting the most impactful variables using statistical techniques and domain knowledge.  
  
 5. Model Development: I experimented with various machine learning algorithms, including decision trees, random forests, and neural networks, to develop the prediction model. This involved iterative training, testing, and validation processes to optimize model performance.  
  
 6. Interface Design: I created wireframes and prototypes for the user interface, focusing on making it intuitive for students and administrators. This involved multiple rounds of user feedback and refinement.  
  
 7. Integration Planning: I collaborated with the IT department to plan the integration of the prediction model with existing university systems, identifying potential challenges and solutions.  
  
 8. Testing and Validation: I conducted thorough testing of the model and interface, including unit tests, integration tests, and user acceptance testing with a pilot group of students and staff.  
  
 9. Documentation: Throughout the project, I maintained comprehensive documentation, including technical specifications, user guides, and a project report detailing the methodology, results, and recommendations for future improvements.  
  
 These activities, while not exhaustive, represent the core efforts undertaken to deliver the student placement prediction project, balancing technical development with stakeholder engagement and project management principles.

# **TARGETED V/S ACHIEVED OUTPUT**

The targeted output for this student placement prediction project was to develop a fully functional model with an accuracy rate of 85% in predicting student placements, along with a user-friendly interface for students and administrators. The plan also included integration with the university's existing systems and a comprehensive report on model performance and

insights.  
 Existing deviations provide valuable lessons for future projects. They highlight the importance of thorough data assessment in the early stages, more conservative time estimation for complex tasks, proactive stakeholder management, and early exploration of technical integration challenges. Despite not meeting all targets, the project has laid a solid foundation for further development and improvement of the student placement prediction system.

# **CONCLUSION**

This student placement prediction model will prove highly valuable for various stakeholders in the educational ecosystem. For students, it offers personalized insights into potential career paths, helping them make informed decisions about their academic focus and skill development. This guidance can lead to better alignment between student aspirations and market demands, potentially improving job satisfaction and career success-rates.  
  
Educational institutions can leverage the model to tailor their curriculum and career services, ensuring they're equipping students with the most relevant skills for the job market. This can enhance the institution's reputation and placement rates, attracting more students.  
Employers benefit from a more efficient recruitment process, as they can target their efforts towards students who are likely to be good fits for their organizations.

# **APPENDICES**

## Appendix A – Title