

# AI-Enabled HR Analytics

Submitted By

**Malhar Bhavesh Shinde**

**20BCE148**



**DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING**

**INSTITUTE OF TECHNOLOGY**

**NIRMA UNIVERSITY**

**AHMEDABAD-382481**

**May 2024**

# AI-Enabled HR Analytics

## Project Report

Submitted in partial fulfillment of the requirements

for the degree of

**Bachelor of Technology in Computer Science and Engineering**

By

**Malhar Bhavesh Shinde**

**(20BCE148)**

Guided By

**Dr. Sudeep Tanwar**

**DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING**



**DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING**

**INSTITUTE OF TECHNOLOGY**

**NIRMA UNIVERSITY**

**AHMEDABAD-382481**

**May 2024**



## Certificate

This is to certify that the major project entitled “**AI-Enabled HR Analytics**” submitted by **Malhar Bhavesh Shinde (20BCE148)**, towards the partial fulfillment of the requirements for the award of the degree of Bachelor of Technology in Computer Science and Engineering, Nirma University, Ahmedabad, is the record of work carried out by him/her under my supervision and guidance. In my opinion, the submitted work has reached the level required for being accepted for examination.

Dr Sudeep Tanwar

Professor

Department of Computer Science & Eng.

Institute of Technology,

Nirma University, Ahmedabad.

Dr Madhuri Bhavsar

Professor and Head,

Department of Computer Science & Eng.

Institute of Technology,

Nirma University, Ahmedabad

## Statement of Originality

---

I, **Malhar Bhavesh Shinde**, Roll. No. **20BCE148**, give an undertaking that the major project entitled “**AI-Enabled HR Analytics**” submitted by me, towards the partial fulfilment of the requirements for the degree of Bachelor of Technology in **Computer Science and Engineering**, Nirma University, Ahmedabad, contains no material that has been awarded for any degree or diploma in any university or school in any territory to the best of my knowledge. It is the original work carried out by me and I give assurance that no attempt of plagiarism has been made. It contains no material that is previously published or written, except where reference has been made. I understand that in the event of any similarity found subsequently with any published work or any dissertation work elsewhere; it will result in severe disciplinary action.

---

Signature of Student

Date:

Place: Ahmedabad

Endorsed by  
Dr Sudeep Tanwar

# Acknowledgements

I want to extend my heartfelt appreciation to all those who played a pivotal role in bringing the project to fruition. Without the unwavering assistance and guidance from a select few, this endeavor would not have reached its successful conclusion.

I would like to express my gratitude to Komal Lamba, my mentor during the internship period. Her exceptional guidance, patience, and profound expertise has significantly enhanced the quality and relevance of the project. Her practical insights from the industry have been particularly enlightening.

I am deeply grateful to Dr Sudeep Tanwar, my esteemed mentor throughout my college journey. His invaluable insights, unwavering support, and adept guidance have been instrumental in shaping my understanding and motivating me to strive for excellence.

Furthermore, I extend my thanks to Bhoomi Mehta, Manas Jani, and Keyur Shah for their valuable contributions. Your involvement in this enriching experience and your significant contributions to the development of this project have been truly indispensable.

Thank you all for your participation in this rewarding journey. Your dedication has been instrumental in achieving our collective success.

**- Malhar Bhavesh Shinde**

**20BCE148**

# Abstract

Human Resources (HR) departments are increasingly relying on artificial intelligence (AI) to strengthen their analytical capabilities in the dynamic world of modern business. This abstract explores the revolutionary possibilities of AI-driven HR analytics, shedding light on how companies use cutting-edge algorithms to optimize HR procedures and guide strategic decision-making.

AI-powered HR analytics use machine learning techniques to extract insights from large and diverse datasets that may be put to use. Through examining employee data from training, performance reviews, recruitment, and retention metrics, businesses have a better understanding of the dynamics of their staff. With the use of these information, HR managers may identify trends, predict future developments, and tailor actions to improve worker satisfaction, productivity, and engagement.

Moreover, AI-powered predictive models enable HR directors to foresee talent needs and proactively address problems like skill shortages and attrition. Organizations can better connect their human capital strategies with overarching business goals and encourage adaptation and resilience in response to change by anticipating workforce trends and evaluating the possible impact of changes.

However, there are unique difficulties in incorporating AI into HR analytics. Ethics-related issues, such as algorithmic bias, transparency, and data privacy, need to be carefully taken into account. Organizations must carefully negotiate these complexities to guarantee the ethical and responsible application of AI technologies. Developing fair models, implementing strong data governance, and being transparent in algorithmic decision-making are all essential to building confidence and controlling the dangers involved in the deployment of AI.

AI-driven HR analytics uses techniques like machine learning, sentiment analysis, and natural language processing (NLP) for tasks like resume screening, scheduling interviews, and employee satisfaction analysis. Its goals are centered around hiring and employee retention.

Simplified hiring procedures, enhanced candidate experiences, thorough evaluations of employee happiness, preventative steps to reduce attrition, and targeted talent management programs are all examples of the intended results. Organizations can achieve

these goals—promoting operational efficiency, developing a pleasant workplace culture, and gaining a competitive edge in the dynamic field of HR management—by utilizing AI technologies.

# Abbreviations

<b>AI</b>	Artificial Intelligence
<b>HR</b>	Human Resources
<b>KPI</b>	Key Performance Indicator
<b>TF-IDF</b>	Term Frequency-Inverse Document Frequency
<b>PDF</b>	Portable Document Format
<b>URL</b>	Uniform Resource Locator
<b>CSV</b>	Comma Separated Values
<b>AVP</b>	Associate Vice President
<b>HTML</b>	Hypertext Markup Language
<b>CSS</b>	Cascading Style Sheets
<b>PK</b>	Primary Key
<b>FK</b>	Foreign Key
<b>NLP</b>	Natural Language Processing
<b>RVSGDC</b>	Random Vector-Support Gradient Descent Classifier
<b>ROI</b>	Return on Investment
<b>APIs</b>	Application Program Interfaces
<b>ML</b>	Machine Learning
<b>AWS</b>	Amazon Web Services
<b>UI/UX</b>	User Interface/User Experience
<b>FAQ</b>	Frequently Asked Questions
<b>CPU</b>	Central Processing Unit
<b>OS</b>	Operating System

---



# Contents

<b>Certificate</b>	<b>iii</b>
<b>Statement of Originality</b>	<b>iv</b>
<b>Acknowledgements</b>	<b>v</b>
<b>Abstract</b>	<b>vi</b>
<b>Abbreviations</b>	<b>viii</b>
<b>List of Tables</b>	<b>xi</b>
<b>List of Figures</b>	<b>xii</b>
<b>1 Introduction</b>	<b>1</b>
1.1 About the Company . . . . .	1
1.1.1 Introduction of the Company . . . . .	1
1.2 The System . . . . .	2
1.2.1 Definition of System . . . . .	2
1.2.2 Purpose and Objectives . . . . .	2
1.2.3 About the Present System . . . . .	3
1.2.4 Proposed System . . . . .	4
1.3 Project Profile . . . . .	5
1.3.1 Project Title . . . . .	5
1.3.2 Scope of Project . . . . .	5
1.3.3 Project Team . . . . .	6
1.3.4 Hardware/Software environment . . . . .	6
<b>2 System Analysis</b>	<b>7</b>
2.1 Admin . . . . .	7
2.2 Candidate . . . . .	9
2.3 Employee . . . . .	10
2.4 AVP . . . . .	10
2.5 HR . . . . .	12
2.6 Feasibility Study . . . . .	15
2.6.1 Technical Feasibility . . . . .	15
2.6.2 Operational Feasibility . . . . .	15
2.6.3 Financial Feasibility . . . . .	15
2.6.4 Economic Feasibility . . . . .	16

<b>3</b>	<b>System Design</b>	<b>17</b>
3.1	System Flow . . . . .	17
3.2	Entity-Relationship Diagram . . . . .	20
3.3	Data Dictionary . . . . .	22
<b>4</b>	<b>Result and Discussion</b>	<b>32</b>
4.1	Results . . . . .	32
4.1.1	Resume Screening and Cover Letter . . . . .	32
4.1.2	Chatbot . . . . .	33
4.1.3	Employee Review . . . . .	34
4.1.4	Employee Promotion . . . . .	36
4.1.5	Front-End . . . . .	40
<b>5</b>	<b>Conclusion and Future Work</b>	<b>50</b>
5.1	Conclusion . . . . .	50
5.2	Future Scope . . . . .	51
	<b>Appendices</b>	<b>52</b>
	<b>Bibliography</b>	<b>54</b>

# List of Tables

1.1	Company Information . . . . .	2
2.1	Collections used by Admin . . . . .	8
2.2	Collection used by Candidate . . . . .	10
2.3	Collection used by Employee . . . . .	11
2.4	Collections for AVP . . . . .	12
2.5	Collections for HR . . . . .	14

# List of Figures

2.1	Flow of Admin . . . . .	8
2.2	Flow of Candidate . . . . .	9
2.3	Flow of Employee and AVP . . . . .	13
2.4	Flow of HR . . . . .	14
3.1	Entity-Relationship Diagram . . . . .	20
3.2	Entity-Relationship Diagram . . . . .	21
4.1	Chat-bot . . . . .	33
4.2	Chat-bot Demo . . . . .	34
4.3	Employee Review Graphs . . . . .	35
4.4	KPIs met and Awards won . . . . .	37
4.5	Score Graph . . . . .	37
4.6	Age Category . . . . .	38
4.7	Service Category . . . . .	38
4.8	Department Category . . . . .	39
4.9	Previous Year Rating Category . . . . .	39
4.10	Login Page . . . . .	40
4.11	HR Dashboard . . . . .	41
4.12	Candidate Dashboard . . . . .	41
4.13	Application Status . . . . .	42
4.14	Meeting Scheduled . . . . .	42
4.15	Result Announced . . . . .	43
4.16	HR-Schedule Meet . . . . .	43
4.17	Application Status . . . . .	44
4.18	Admin Dashboard . . . . .	44
4.19	Add Employee . . . . .	45
4.20	Delete Employee . . . . .	45
4.21	Edit Employee . . . . .	46
4.22	Add Candidate . . . . .	46
4.23	AVP Dashboard . . . . .	47
4.24	Feedback Statistics . . . . .	47
4.25	Feedback Statistics . . . . .	48
4.26	Feedback Statistics . . . . .	48
4.27	Promotion Statistics . . . . .	49
4.28	Promotion Statistics . . . . .	49

# Chapter 1

## Introduction

### 1.1 About the Company

#### 1.1.1 Introduction of the Company

Cygnet Digital, a trusted partner in the digital era, is here to lead the way. The company offers-centric digital services, streamlined tax solutions, and finance empowerment. With expertise in Digital Engineering, Quality, Enterprise Apps, Data, AI, and Digital Commerce, we amplify your success.

It provides the following services namely :-

- Digital, Applications & Product Engineering Cloud Engineering.
- Enterprise Applications AWS, Sales Force, Microsoft Cloud.
- Data, Analytics, AI, IoT & Automation Data Virtualization, Data Visualization, Data Modernization.
- Digital Commerce & Experience UI/ UX, Adobe, Unified Commerce.

One of the most reputable brands in the IT industry, Cygnet Infotech provides technological solutions in 35 nations. Cygnet was founded with the goal of becoming a software development firm where high standards, inventiveness, and customized services are prioritized over inexpensive, improvised solutions. The company helps its clients digitize, grow, and become high-performing enterprises.

<b>Company Name</b>	Cygnnet One
<b>Website</b>	<a href="https://www.cygnnet.one/">https://www.cygnnet.one/</a>
<b>Email</b>	info@cygnnet-digital.com
<b>Address</b>	2nd Floor, The Textile Association, Dinesh Hall, Navrangpura, Ahmedabad-380009

Table 1.1: Company Information

Cygnnet offers extensive knowledge of company processes and industries, access to resources across the globe, and a track record of creating innovative technological solutions. Cygnnet can help to increase corporate performance by mobilizing the appropriate people, skills, and technologies.

## 1.2 The System

### 1.2.1 Definition of System

The AI Enabled HR Analytics system is a comprehensive solution designed to streamline and optimize various aspects of human resources management within an organization. Comprising multiple modules, including Recruitment, Employee Promotion, and Sentiment Analysis, the system harnesses the power of artificial intelligence (AI) and machine learning (ML) techniques to facilitate data-driven decision-making and enhance organizational efficiency.

### 1.2.2 Purpose and Objectives

The purpose of the AI Enabled HR Analytics project is to revolutionize the human resources management process within organizations by integrating advanced artificial intelligence (AI) and machine learning (ML) techniques. This project aims to address key challenges faced by HR departments, such as streamlining recruitment processes, identifying potential candidates for promotion, gauging employee sentiment, and enhancing candidate engagement.

By leveraging cutting-edge technologies, the project seeks to automate repetitive tasks, improve decision-making accuracy, and provide actionable insights to HR personnel. Ultimately, the goal is to create a more efficient, transparent, and data-driven HR ecosystem that aligns with organizational objectives and fosters employee satisfaction and organizational success.

The project aims to enhance the efficiency and effectiveness of recruitment processes by leveraging AI algorithms for resume parsing, skill assessment, and cultural fit evaluation. This objective seeks to reduce the time and effort involved in candidate evaluation while ensuring the selection of top-quality candidates who align with the organization's values.

Through machine learning models, the project aims to predict eligible candidates for promotion based on performance metrics and criteria established by the organization. This objective aims to streamline the promotion decision-making process, foster transparency, and ensure that deserving employees are recognized and rewarded for their contributions.

By conducting sentiment analysis on employee surveys, the project seeks to gain insights into employee sentiment, satisfaction, and concerns. This objective enables HR personnel to identify trends, address issues proactively, and implement targeted interventions to improve employee engagement and retention.

The integration of a rule-based chatbot module into the candidate dashboard aims to enhance candidate engagement and experience. This objective provides candidates with a platform to gain insights into the organization's culture, values, and working environment, fostering a positive candidate experience and improving the employer brand.

Overall, the project aims to promote a culture of data-driven decision-making within the HR department and the organization as a whole. By providing HR personnel with comprehensive data analytics and insights, the project empowers them to make informed decisions that align with organizational goals, drive efficiency, and support strategic initiatives.

In summary, the AI Enabled HR Analytics project seeks to leverage advanced technologies to transform HR management processes, drive operational excellence, and create a workplace environment conducive to employee success and organizational growth.

### **1.2.3 About the Present System**

Conventional HR analytics is a methodical approach to using data to obtain understanding of many facets of an organization's personnel. Comprehensive data gathering from several sources, including payroll systems, attendance records, performance reviews, and personnel files, is the first step in the process. After that, to ensure confidentiality and

accessibility, the data is kept in a centralized HR database or management system.

Data integration and cleaning must be completed before analysis can start. This stage entails finding and fixing any errors, duplicate entries, or inconsistencies in the datasets that have been gathered. The data creates a solid basis for analysis once it is standardized and integrated.

HR specialists use statistical and analytical methods to find patterns, trends, and correlations in the workforce data during the analysis phase. This approach provides insights on a range of HR measures, such as training efficacy, performance drivers, employee turnover rates, and other relevant elements.

Charts, graphs, and reports are examples of visualization tools used to successfully communicate these findings. Human resources workers and other corporate stakeholders can better understand and retrieve complex data with the use of visual representations.

Making educated decisions is one of the main benefits of HR analytics. HR teams and organizational executives can make strategic decisions about hiring, talent management, training and development programs, pay and benefits plans, and other areas of workforce management by utilizing the insights gained from data analysis.

#### **1.2.4 Proposed System**

The Recruitment module serves as the initial phase of the system, enabling candidates to submit their resumes and cover letters for evaluation. Advanced parsing algorithms extract crucial information from resumes, while cultural fit assessments are performed based on the content of cover letters. Candidate profiles, along with cultural fit scores, are presented on the HR dashboard for review and decision-making by HR personnel.

In the Employee Promotion module, a machine learning model is utilized to predict eligible employees for promotion based on various performance metrics and criteria established by the organization. The list of eligible employees is presented to HR personnel and their respective reporting managers for review and consideration.

The system conducts sentiment analysis on employee surveys conducted biannually, leveraging natural language processing (NLP) techniques. The sentiment analysis results are presented to HR personnel, categorized by department, enabling them to gain insights into employee satisfaction, engagement, and concerns. These insights aid HR in making informed decisions and implementing targeted interventions to improve employee



experience and retention.

Integrated into the candidate dashboard, the Rule-Based Chatbot module enables candidates to gain insights into the organization's culture, values, and working environment. This interactive feature serves as a resource for candidates to acquire information and clarifications, enhancing their understanding of the company and facilitating a positive candidate experience.

## **1.3 Project Profile**

### **1.3.1 Project Title**

AI-Enabled HR Analytics

### **1.3.2 Scope of Project**

The two main areas of concentration in the field of HR analytics are hiring and employee turnover. My methodology, which makes use of sophisticated machine learning techniques like sentiment analysis and natural language processing (NLP), is designed to improve HR procedures and effectively inform decision-making.

Recruitment Process:

- **Resume and Cover Letter Screening:** I carefully go over resumes, extracting pertinent information and determining a candidate's appropriateness for different roles using sentiment analysis and natural language processing techniques. I guarantee a more detailed assessment of candidates by probing into sentiment, which expedites the recruiting process. I also use sentiment analysis to examine cover letters in order to assess cultural alignment and identify any red flags. Prescriptive analysis, which is based on performance indicators and historical data, optimizes recruitment channels and improves selection criteria to support these efforts.
- **Chatbot:** Integrating a chatbot enhances the experience of the user and creates a good impression of the company. Some FAQs about the company can be answered by the chat bot.

Turnover of Employees:

- Employee Feedback and assessments: NLP is essential for deciphering themes and emotions in employee feedback and assessments. It provides insightful information about common problems, areas of satisfaction, and areas where the organization may improve.
- Employee Turnover and Leadership Development: By utilizing machine learning techniques, I am able to achieve a thorough understanding of employee and participant satisfaction and satisfy engagement levels in a holistic manner. Prescriptive models also support activities for career development, reduce attrition, and identify people with high potential for leadership positions.

Through the strategic application of these methodologies and approaches, my objective is to elevate HR practices, refine recruitment processes, and foster organizational development. Ultimately, I aim to contribute significantly to the company's overall success and sustained growth.

### **1.3.3 Project Team**

- Malhar Bhavesh Shinde
- Kahan Jash
- Parth Parikh

The project was monitored and guided by my industry guide Komal Lamba.

### **1.3.4 Hardware/Software environment**

- Machine: Lenovo Thinkpad
- CPU: Quad 2.40GHz 11th Gen Intel Core i5-1135G7 (Hyper-Threaded)
- Memory: 16040 MB
- OS Version: Windows 10

# Chapter 2

## System Analysis

### 2.1 Admin

The first step has the user entering their login information on the login page, which usually consists of their Employee ID and password. Access to the admin dashboard is provided upon a successful confirmation of the admin's credentials. This dashboard shows a complete roster of all of the organization's current employees. The administrator's primary responsibility in this project is data entry.

The administrative functions comprise a range of operations, such as adding new candidates and employees, modifying current employee records, removing employee records, and adjusting the feedback form that is used for sentiment analysis. The administrator must enter the employee's name, email address, department, position, and phone number when adding them. On the other hand, the method for adding candidates functions through automatic generation. The administrator needs to supply the matching employee ID that needs to be changed in order to amend an employee record. When a former employee returns to the organization, the system uses email ID verification to locate the former employee's Emp ID in the database of prior collections. This procedure gives the employee their previous emp\_id, which guarantees continuity.

It is noteworthy that the administrator has the power to alter the feedback form that staff members use to rate candidates. This form is essential to the sentiment analysis process since it makes it easier to evaluate how employees feel about candidates.

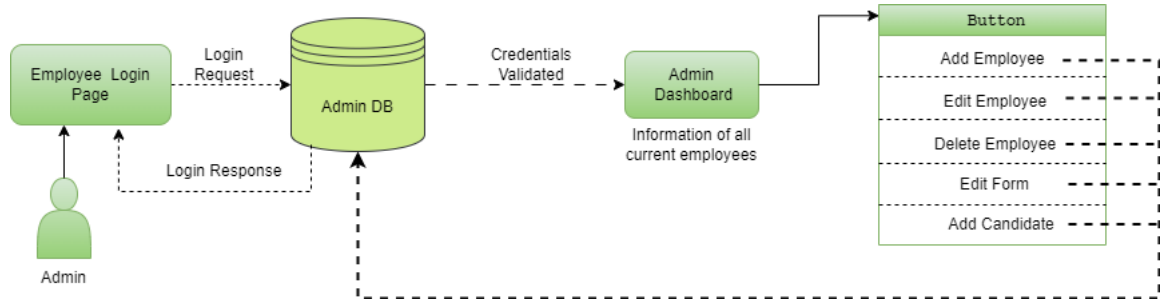


Figure 2.1: Flow of Admin

Table 2.1: Collections used by Admin

Name of Collection	Role of Collection
Login_Details	Stores all the details of the employees like emp_id, name, department, designation, email ID, phone and password
candidate	It stores the details of the candidate namely cand_id, password, contact number, email ID, name and the upload date (of the resume and the cover letter)
past_employees	It stores the details of the employees who have left the company which include emp_id, name, department, designation, email ID and phone
Feedback	It stores all the questions which the admin has edited and they will be displayed on the employee dashboard so as to fill the form

## 2.2 Candidate

Candidates must choose their candidate status on the login screen in order to log in. After being identified as a candidate, users are required to provide their candidate ID, which begins with the letter "C," as well as the password that the HR department emailed them. After these credentials are successfully validated, access to the candidate dashboard is made available.

Relevant information about the application's status is available within the candidate dashboard. This information includes the link, date, and time of meetings that are arranged for later rounds of interviews, which are handled by the HR department. Furthermore, a specific area has been set aside for the mandatory upload of necessary files, specifically the CV and cover letter. Candidates cannot resubmit these documents once they have been submitted, protecting the integrity of the data.

In addition, applicants are provided with a chatbot interface that is intended to respond to questions about the organization, improving communication effectiveness and encouraging the applicant to participate in the hiring process. The purpose of this integrated chatbot function is to expedite communication channels and give candidates prompt answers to their questions.

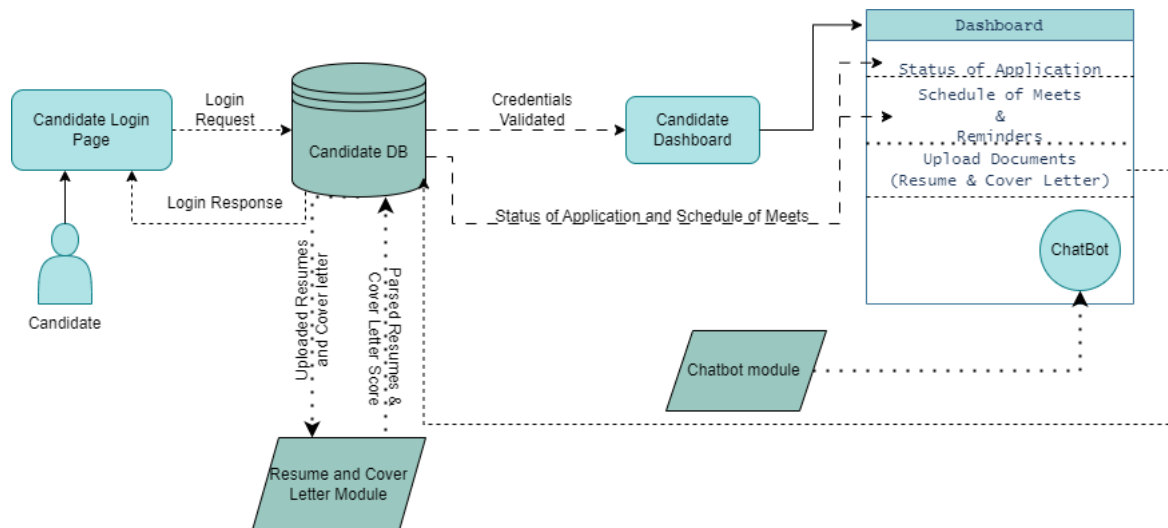


Figure 2.2: Flow of Candidate

Table 2.2: Collection used by Candidate

Name of Collection	Role of Collection
candidate	It stores the details of the candidate namely cand_id, password, contact number, email ID, name and the upload date (of the resume and the cover letter).
resume	It stores the extracted details of the candidate like name, contact info, email, domain,experience_category,skills,cultural_fit score, and if the HR has posted a decision then the status of the result.
rejected_candidate	If the candidate is rejected then the cand_id along with the reason of rejection will be shown this collection.

## 2.3 Employee

Authorized employees are required to provide their password and unique Employee ID in order to gain access to the HR Analytics system's login page. The system takes the authenticated employee to their assigned employee page after successful authentication. The system has a scheduled task that runs every six months and is controlled by a cronjob. The logged-in employee sees a feedback form after completing this automated task. This periodic feedback mechanism's main goal is to determine how satisfied the organization's workforce is.

Employee feedback forms are essential platforms for employees to voice their opinions and offer insightful information about their experiences and opinions about different elements of their work at the company. HR staff members and higher authorities can obtain a thorough grasp of employee satisfaction levels and pinpoint possible areas for improvement within the company by carefully gathering and evaluating this input.

## 2.4 AVP

An essential interface in the HR Analytics system, the AVP Dashboard provides the Associate Vice President (AVP) with a thorough overview of a range of topics pertaining to employee promotion and feedback analysis.

The AVP is required to provide their Employee ID and password in order to have access to the AVP Dashboard. After a successful authentication process, the AVP is granted

Table 2.3: Collection used by Employee

Name of Collection	Role of Collection
Login_Details	It stores all the details of the employees like emp_id, name, department, designation, email ID, phone and password
Feedback	It stores all the questions which the admin has edited and they will be displayed on the employee dashboard so as to fill the form
Feedback_answers	It stores the answers and ratings given by each employee in the feedback form and all of it would be shown as anonymous

access to a structured table that presents relevant details about recommended employee promotions. For every proposed promotion, this table provides important information such the employee ID, department, and matching accept and reject buttons.

A Python program that makes use of a Gradient Boosting Classifier that has been trained on a wide range of parameters generates recommendations for employee promotions. These characteristics include important information including ratings from the prior year, duration of service, met Key Performance Indicators (KPIs), honors received, and average training scores. The AVP is responsible for assessing and selecting which employees to promote based on data from a CSV file that the AVP uploads that has the previously specified information in it.

In addition, the AVP is given informative data in the form of graphical depictions, which makes it easier for them to comprehend the dynamics of employee promotion. These graphs cover a wide range of characteristics, including departmental distribution, awards won, KPIs reached, length of service classification, overall score combining training measures, and service tenure categories.

The AVP receives extensive feedback analysis in addition to promotion statistics. This analysis is generated by a Python tool that uses the RVSGDC method. This method is trained on a feedback dataset and gathers inputs from feedback answers. It combines Support Vector Machine, Stochastic Gradient Descent, and Logistic Regression with hard voting support. The feedback statistics are visually shown, providing information on departmental and overall ratings in a number of different categories, such as senior management perception, work-life balance, career opportunities, cultural values, and overall rating. The feedback replies are also subjected to sentiment analysis, which

Table 2.4: Collections for AVP

Name of Collection	Role of Collection
Login_Details	It stores all the details of the employees like emp_id, name, department, designation, email ID, phone and password
Feedback_answers	It stores the answers and ratings given by each employee in the feedback form and all of it would be shown as anonymous
Employee_Promotion	It stores the data from the CSV file which is uploaded by the AVP
promotion_upload_date	It stores the upload date of the CSV file in the form of timestamp
Predicted_Promotion	It stores the employee ID and the predicted promotion status that is promoted and not promoted
Plots_Promotion	It stores the plots images in the form of binary format along with the plot ID
Plots_Review	It stores the plots images in the form of binary format along with the plot ID.

separates the sentiments into positive, neutral, and negative categories.

## 2.5 HR

An essential interface for the HR Analytics system is the HR Dashboard, which provides HR staff with a full arsenal for expedited resume screening and feedback analysis.

HR staff members are required to provide their Employee ID and password in order to access the HR Dashboard. HR staff have access to a structured table that contains relevant data about resumes that have been examined and the corresponding cover letter cultural fit scores after successfully authenticating.

The table presents crucial information gleaned from resumes: the candidate's name, contact details (phone number and email address), domain mastery as assessed by a Naive Bayes model trained on a sample dataset, work experience classified as Fresher, Beginner, Mid-level, and Experienced, skills mapped using Natural Language Processing (NLP) techniques using a skills dataset, and a link to the cover letter in PDF format for download.

Sentence transformers and cosine similarity metrics are used to create the cover letter's



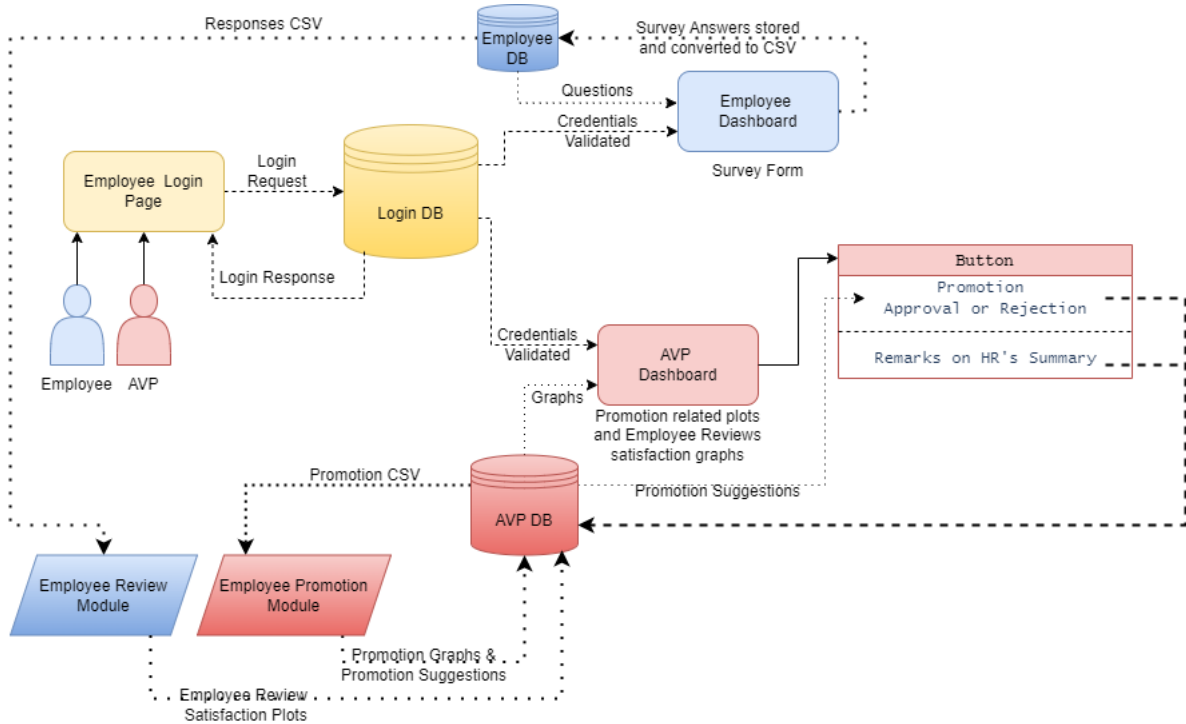


Figure 2.3: Flow of Employee and AVP

cultural fit score, which compares the letter's content to the company's mission and vision to determine alignment and cultural compatibility. Additionally, the HR Dashboard has user-friendly features designed to improve the hiring process. HR staff members have buttons that allow them to approve or reject any candidate shown in the table with ease, expediting the decision-making process.

Additionally, HR staff members have the authority to arrange appointments with candidates straight from the dashboard. Each candidate entry has a specific "Schedule Meeting" button, which enables HR staff to effectively handle the logistics of interviews. When this button is clicked, a new tab opens and HR staff is presented with a form to fill in with the necessary meeting information. This form asks HR staff to upload the meeting link, indicate the time and date of the meeting, and include a brief agenda summary.

Apart from resume screening features, the HR Dashboard provides HR staff with valuable feedback data obtained by a Python code that utilizes the RVSGDC algorithm. This algorithm is trained on a feedback dataset and requests inputs from feedback answers. It is a combination of Logistic Regression, Support Vector Machine, and Stochastic Gradient Descent, supplemented by hard voting.

The feedback statistics are displayed graphically in the form of graphs, providing HR staff with information on department- and overall-level ratings in a variety of categories,

Table 2.5: Collections for HR

Name of Collection	Role of Collection
Login_Details	It stores all the details of the employees like emp_id, name, department, designation, email ID, phone and password
Feedback_answers	It stores the answers and ratings given by each employee in the feedback form and all of it would be shown as anonymous
Plots_Review	It stores the plots images in the form of binary format along with the plot ID.
resume	It stores the extracted details of the candidate like name, contact info, email, domain, experience_category, skills, cultural_fit score, scheduled meeting details, and if the HR has posted a decision then the status of the result
rejected_candidate	If the candidate is rejected then the cand_id along with the reason of rejection will be shown this collection

including work-life balance, cultural values, career opportunities, company benefits, and senior management perceptions. In addition, the feedback replies undergo sentiment analysis, which divides the sentiments into three categories: positive, neutral, and negative.

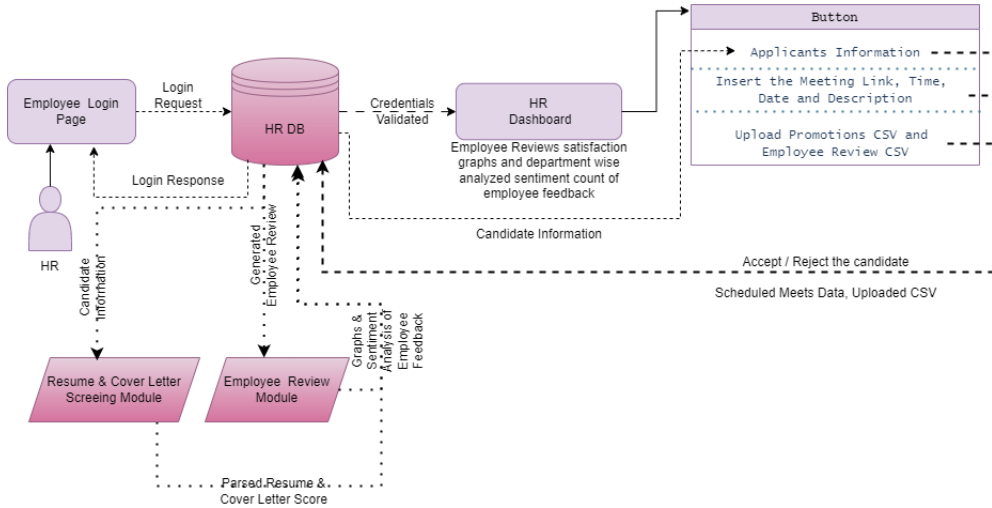


Figure 2.4: Flow of HR

## **2.6 Feasibility Study**

### **2.6.1 Technical Feasibility**

In terms of technical feasibility, the proposed project stands on a solid foundation. Leveraging advanced machine learning techniques such as sentiment analysis and natural language processing (NLP) to sift through resumes, assess cover letters, and interpret employee feedback is not just within reach but increasingly commonplace in modern HR practices. The wealth of data typically available within HR departments provides a fertile ground for these techniques to flourish. Integrating a chatbot to streamline candidate communication further underscores the project's technical viability. While integration with existing HR systems may pose some challenges, the availability of APIs and middleware solutions offers practical pathways to overcome such obstacles. Overall, the project's technical feasibility is robust, buoyed by the maturity of the underlying technologies and their proven effectiveness in similar applications.

### **2.6.2 Operational Feasibility**

Operationally, the project promises to be a game-changer for HR processes. By automating routine tasks like resume screening and interview scheduling through the introduction of a chatbot, HR professionals can redirect their focus towards more strategic endeavors. The insights gleaned from sentiment analysis and NLP empower HR teams to make data-driven decisions, whether it's in candidate selection, identifying areas for organizational improvement, or nurturing leadership talent. While there may be hurdles in terms of change management and training, the tangible benefits in terms of efficiency gains and improved decision-making justify the operational investment. Ultimately, the project's operational feasibility hinges on its ability to streamline workflows and empower HR teams to work smarter, not harder.

### **2.6.3 Financial Feasibility**

From a financial standpoint, the project presents a compelling case for investment. While there may be initial costs associated with implementing machine learning technologies and developing the chatbot infrastructure, the potential for long-term cost savings is significant. By optimizing recruitment processes, reducing turnover rates, and identifying high-

potential employees, the project can yield substantial returns over time. Additionally, the operational efficiencies gained through automation and data-driven decision-making contribute to overall cost reduction. The integration of a chatbot further enhances the project's financial viability by streamlining candidate communication and reducing administrative overhead. In essence, the financial feasibility of the project is underscored by its potential to deliver tangible ROI and drive bottom-line results.

#### **2.6.4 Economic Feasibility**

On the economic front, the project aligns closely with the organization's strategic objectives and growth aspirations. By elevating HR practices, refining recruitment processes, and nurturing talent development, the project lays the groundwork for sustainable organizational growth. The potential for improved employee satisfaction, reduced turnover rates, and enhanced leadership capabilities not only bolsters the company's competitive position but also fosters a positive work culture conducive to long-term success. Moreover, by embracing cutting-edge technologies like machine learning and NLP, the project positions the organization as a forward-thinking leader in HR analytics, paving the way for continued innovation and differentiation in the marketplace. In summary, the economic feasibility of the project is rooted in its ability to drive strategic value and foster a culture of excellence and growth within the organization.

# Chapter 3

## System Design

### 3.1 System Flow

The system consists of the following end users:

- HR
- AVP
- Employee
- Admin
- Candidate

The following is the system flow for both end users and the hiring process:

- Admin: Using their login credentials, the admin accesses the system. The admin dashboard is displayed to them after they log in. The admin is then taken to the add candidate page by clicking the "Add Candidate" button. Here, the candidate is registered by the admin, who also gives the candidate their login information.
- Candidate: The administrator has already given the candidate their login information. The login page is where the candidate interacts with the system for the first time. After the candidate inputs their login information, they are verified. They are sent to the candidate dashboard if the credentials are legitimate. There are three parts on the candidate dashboard. The candidate can view their application status in the first part. They are informed about their scheduled meetings in the

next part. The applicant is prompted to upload their cover letter and résumé in the last part. Additionally, they are asked to enter their years of experience and domain. These particulars are kept in the MongoDB resume collection within the Project database.

- **Resume Parsing Module:** Next, the resume parsing module retrieves the cover letters and resumes that the candidate uploaded. This module retrieves the candidate's name, phone number, and email address. Additionally, it retrieves the candidate's skills and saves them in the MongoDB resume collection within the Project database.
- **HR:** After logging in, the HR is taken to the HR dashboard. The employee review graphs produced by the employee review module are displayed on the HR dashboard. Next, under the menu bar, the HR selects the Applicants button. This sends HR to the applicants page, where the Resume Parsing module results are displayed. After careful consideration and consultation with the recruitment team, HR is then able to decide whether or not to accept the candidate. The choice to accept or reject is entered into the database and subsequently shown on the candidate dashboard's application status page. HR is required to compose a message explaining the reason for rejection if the candidate is not accepted.

The following is the system flow for the end users and the promotion and review module:

- **HR:** After logging into the system, the HR selects the promotion option. After clicking this, the HR is taken to a different page where he must upload a CSV file including all the information needed to assess an employee's potential for promotion. All of the company's reporting managers prepare this CSV file, which is subsequently sent to HR. After then, the file is kept in the database's promotion collection.
- **Admin:** After logging in, the admin selects the edit form button. The admin can access the edit form page by clicking this button. The administrator can amend the survey questions on this page. These queries are then kept in the database's survey collection.

- Employee: Only after the survey is launched are the employees able to log in. Every six months, the survey is floated. Then the staff members can respond to the inquiries. The responses are categorized by department and anonymously recorded. The department-specific reviews are kept in the database's survey collection.
- Employee Review Module: This module receives department-specific reviews that are entered into the database's survey collection. Next, the algorithm forecasts the reviews categorized by department. Positive, negative, and neutral department-wise reviews are produced by the model. After that, this information is kept in the database's survey collection.
- Employee Promotion Module: The promotion csv file is the input for the employee promotion module. Next, the algorithm forecasts every employee who qualifies for a promotion. This information is also included in the database's promotion collection.
- AVP HR: After logging into the system, the AVP HR is taken to the dashboard. The graphs produced by the employee promotion module are shown on the dashboard. The dashboard additionally displays the employee satisfaction rates broken down by department. The navigation bar of AVP HR contains two buttons. When users click the promotion button, a promotion page with a list of all the employees from the database's promotion collection who are qualified for the collection is displayed. After that, they have the option to accept or reject any employee's promotion. These particulars are then kept in the database's promotion collection.

The feedback page is displayed to the AVP HR upon clicking the feedback button. The AVP is asked to provide input for each department on this page. The survey collection then contains this feedback. The HR department will then see this feedback.

## 3.2 Entity-Relationship Diagram

### ER diagram for hiring module

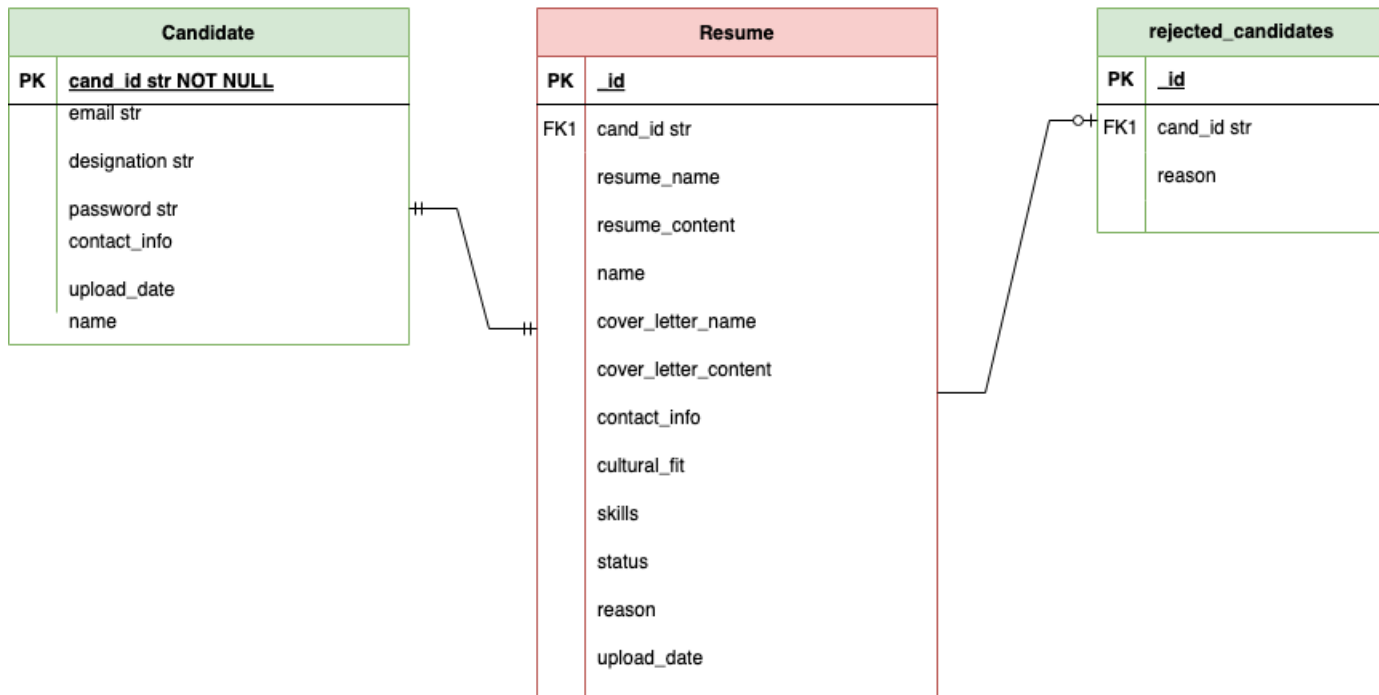


Figure 3.1: Entity-Relationship Diagram



## ER diagram for Turnover and Feedback

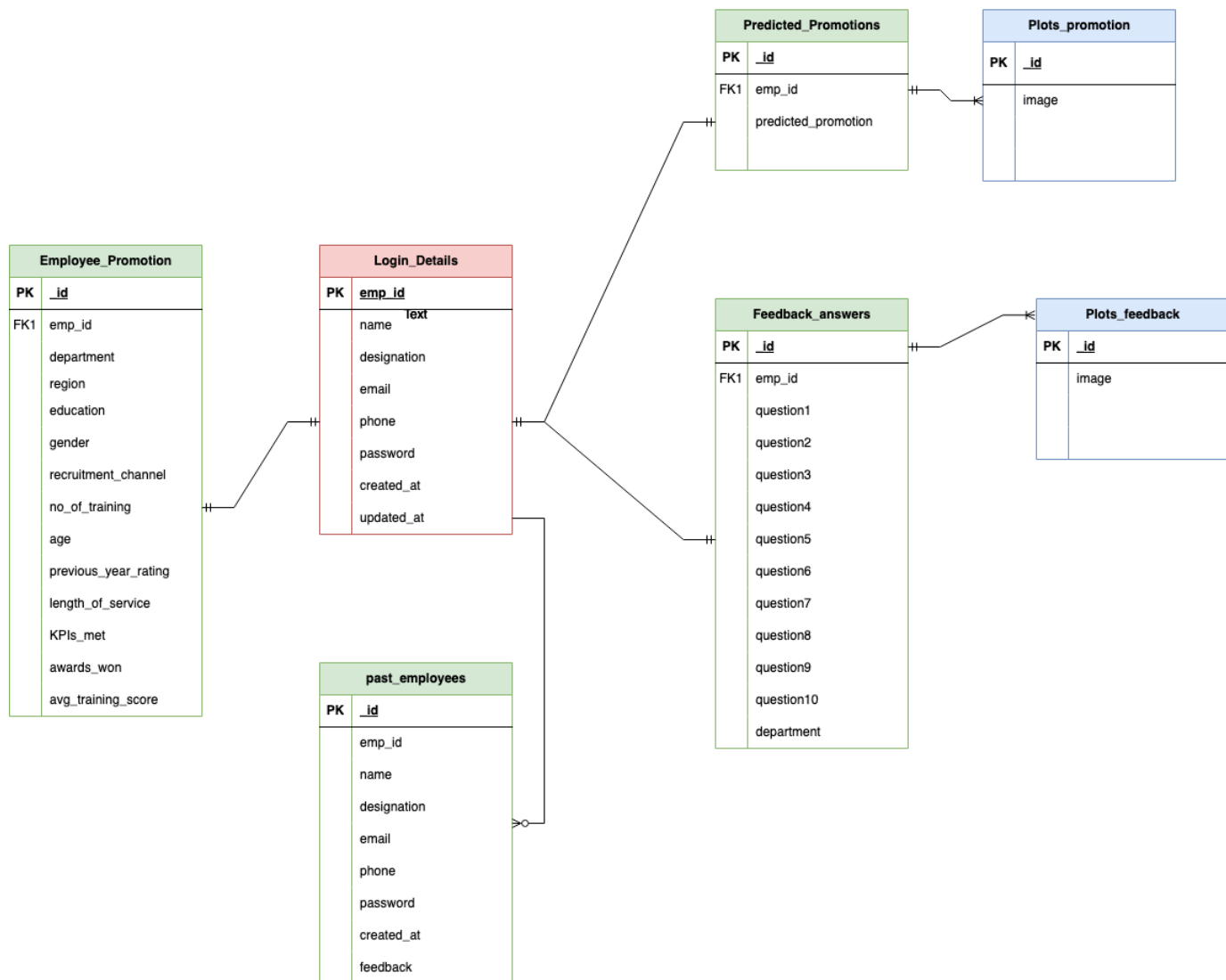


Figure 3.2: Entity-Relationship Diagram

### 3.3 Data Dictionary

#### Collection: Login\_Details

PK: emp\_id

- **name:**
  - **Description:** Full name of the employee.
  - **Data Type:** String.
  - **Constraints:** None.
- **designation:**
  - **Description:** Job title or position of the employee.
  - **Data Type:** String.
  - **Constraints:** None.
- **email:**
  - **Description:** Email address of the employee.
  - **Data Type:** String.
  - **Constraints:** Unique.
- **phone:**
  - **Description:** Phone number of the employee.
  - **Data Type:** String.
  - **Constraints:** None.
- **password:**
  - **Description:** Encrypted password for employee login.
  - **Data Type:** String (Encrypted).
  - **Constraints:** None.
- **created\_at:**

- **Description:** Timestamp indicating when the employee record was created.
- **Data Type:** Date/Time.
- **Constraints:** None.
- **updated\_at:**
  - **Description:** Timestamp indicating when the employee record was last updated.
  - **Data Type:** Date/Time.
  - **Constraints:** None.

### Usage Guidelines:

- The *emp\_id* field serves as the primary key and uniquely identifies each employee in the system.
- The *email* field must be unique for each employee and is used as a unique identifier for login purposes.
- Passwords stored in the *password* field should be securely encrypted to ensure data security.
- Timestamp fields (*created\_at* and *updated\_at*) are automatically generated and updated by the system to track record creation and modification times.

## Collection: Employee\_Promotion

**PK:** *\_id*

**FK:** *emp\_id*

- **department:**
  - **Description:** Department to which the employee belongs.
  - **Data Type:** String.
  - **Constraints:** None.
- **region:**

- **Description:** Region where the employee is located.
- **Data Type:** String.
- **Constraints:** None.
- **education:**
  - **Description:** Highest education qualification of the employee.
  - **Data Type:** String.
  - **Constraints:** None.
- **gender:**
  - **Description:** Gender of the employee.
  - **Data Type:** String.
  - **Constraints:** None.
- **recruitment\_channel:**
  - **Description:** Channel through which the employee was recruited.
  - **Data Type:** String.
  - **Constraints:** None.
- **no\_of\_training:**
  - **Description:** Number of training programs attended by the employee.
  - **Data Type:** Integer.
  - **Constraints:** None.
- **age:**
  - **Description:** Age of the employee.
  - **Data Type:** Integer.
  - **Constraints:** None.
- **previous\_year\_rating:**

- **Description:** Rating received by the employee in the previous year’s performance review.
- **Data Type:** Integer.
- **Constraints:** None.
- **length\_of\_service:**
  - **Description:** Length of service (in years) of the employee.
  - **Data Type:** Integer.
  - **Constraints:** None.
- **KPIs\_met:**
  - **Description:** Indicates whether the employee met the Key Performance Indicators (KPIs) in the past year.
  - **Data Type:** Boolean (True/False).
  - **Constraints:** None.
- **awards\_won:**
  - **Description:** Indicates whether the employee won any awards in the past year.
  - **Data Type:** Boolean (True/False).
  - **Constraints:** None.
- **avg\_training\_score:**
  - **Description:** Average score obtained by the employee in training programs.
  - **Data Type:** Integer.
  - **Constraints:** None.

### Usage Guidelines:

- The *\_id* field serves as the primary key for the collection.
- The *emp\_id* field serves as a foreign key referencing the employee’s unique identifier in another collection.

- The collection contains various attributes related to employee demographics, performance metrics, and training history, which are used for promotion evaluation purposes.

## Collection: Past\_Employees

**PK:** \_id

**FK:** emp\_id

- **name:**
  - **Description:** Full name of the past employee.
  - **Data Type:** String.
  - **Constraints:** None.
- **designation:**
  - **Description:** Last job title or position of the past employee.
  - **Data Type:** String.
  - **Constraints:** None.
- **email:**
  - **Description:** Email address of the past employee.
  - **Data Type:** String.
  - **Constraints:** None.
- **phone:**
  - **Description:** Phone number of the past employee.
  - **Data Type:** String.
  - **Constraints:** None.
- **password:**
  - **Description:** Encrypted password for past employee login.
  - **Data Type:** String (Encrypted).

- **Constraints:** None.
- **created\_at:**
  - **Description:** Timestamp indicating when the past employee record was created.
  - **Data Type:** Date/Time.
  - **Constraints:** None.
- **feedback:**
  - **Description:** Feedback provided by the past employee.
  - **Data Type:** String.
  - **Constraints:** None.

#### Usage Guidelines:

- The *\_id* field serves as the primary key for the collection.
- The *emp\_id* field serves as a foreign key referencing the past employee's unique identifier in another collection.
- The collection contains attributes related to past employees' information and feedback provided by them upon exit.

## Collection: Predicted\_Promotions

**PK:** *\_id*

**FK:** *emp\_id*

- **predicted\_promotion:**
  - **Description:** Prediction of promotion for the employee.
  - **Data Type:** String.
  - **Constraints:** None.

#### Usage Guidelines:

- The *\_id* field serves as the primary key for the collection.

- The *emp\_id* field serves as a foreign key referencing the employee's unique identifier in another collection.
- The collection contains the predicted promotion status (e.g., "Promotion Predicted" or "No Promotion Predicted") for each employee.

## Collection: Feedback\_answers

**PK:** *\_id*

**FK:** *emp\_id*

- **question\_1:**
  - **Description:** Answer to question 1 in the feedback survey.
  - **Data Type:** String.
  - **Constraints:** None.
- **question\_2:**
  - **Description:** Answer to question 2 in the feedback survey.
  - **Data Type:** String.
  - **Constraints:** None.
- **question\_3:**
  - **Description:** Answer to question 3 in the feedback survey.
  - **Data Type:** String.
  - **Constraints:** None.
- **department:**
  - **Description:** Department to which the employee belongs.
  - **Data Type:** String.
  - **Constraints:** None.

## Usage Guidelines:

- The *\_id* field serves as the primary key for the collection.



- The *emp\_id* field serves as a foreign key referencing the employee's unique identifier in another collection.
- The collection contains answers to a set of feedback questions provided by employees, along with their department information.

## Collection: Plots\_Promotion

PK: *\_id*

- **image:**
  - **Description:** Image representing promotion plots or graphs.
  - **Data Type:** Binary.
  - **Constraints:** None.

### Usage Guidelines:

- The *\_id* field serves as the primary key for the collection.
- The collection stores images depicting plots or graphs related to promotions, which may include visualizations of promotion trends, analysis results, etc.

## Collection: Plots\_feedback

PK: *\_id*

- **image:**
  - **Description:** Image representing promotion plots or graphs.
  - **Data Type:** Binary.
  - **Constraints:** None.

### Usage Guidelines:

- The *\_id* field serves as the primary key for the collection.
- The collection stores images depicting plots or graphs related to promotions, which may include visualizations of promotion trends, analysis results, etc.

## Collection: Candidate

**PK:** cand\_id

- **email:**
  - **Description:** Email address of the candidate.
  - **Data Type:** String.
  - **Constraints:** None.
- **designation:**
  - **Description:** Job designation or position of the candidate.
  - **Data Type:** String.
  - **Constraints:** None.
- **password:**
  - **Description:** Encrypted password for candidate login.
  - **Data Type:** String (Encrypted).
  - **Constraints:** None.
- **contact\_info:**
  - **Description:** Contact information of the candidate.
  - **Data Type:** String.
  - **Constraints:** None.
- **upload\_date:**
  - **Description:** Date when the candidate's information was uploaded.
  - **Data Type:** Date/Time.
  - **Constraints:** None.
- **name:**
  - **Description:** Full name of the candidate.

- **Data Type:** String.
- **Constraints:** None.

#### Usage Guidelines:

- The *cand\_id* field serves as the primary key for the collection.
- The collection stores information related to candidates, including their email, designation, password, contact information, upload date, and name.

### Collection: Candidate

**PK:** *\_id*

**FK:** *cand\_id*

- **reason:**
  - **Description:** Reason associated with the candidate.
  - **Data Type:** String.
  - **Constraints:** None.

#### Usage Guidelines:

- The *\_id* field serves as the primary key for the collection.
- The *cand\_id* field serves as a foreign key referencing the candidate's unique identifier in another collection.
- The collection stores information related to candidates, including a reason associated with them.

# Chapter 4

## Result and Discussion

### 4.1 Results

#### 4.1.1 Resume Screening and Cover Letter

The first step in the preparation procedure is to convert all of the text in the skills dataset to lowercase in order to standardize it across entries. Then, a function to extract skills is carefully designed, taking into account bigrams and trigrams in order to capture multi-word skill sets. Developing a function to identify potential names using a combination of advanced named entity recognition techniques and pattern matching techniques is another essential component. Furthermore, a crucial feature is developed using strong regular expressions to reliably extract contact details, including phone numbers and email addresses. These methodical techniques produce a carefully produced dataset that is ideal for applications such as resume screening, enabling more efficient talent acquisition procedures.

In order to prepare the extracted resume text for domain extraction, we lowercase it, eliminate blank spaces, eliminate links, eliminate alpha-numeric values, eliminate stop words, perform lemmatization (which allows us to extract only the word stem), and tokenize the extractions.

By specifying the TF-IDF vectorization pipelines, standard scaling (if necessary), and the model name, several models can be trained. Five models are compared: Random Forest, Support Vector Machine, Neural Network, Multinomial Naïve Bayes, and Logistic Regression. To get each model's ideal parameter, I had to utilize autoML [1]. Given that the Naïve Bayes model outperforms all others, we save the model using joblib, read the

resume using the Fitz library, and provide the domain forecast. We must preprocess the collected text by deleting the phone numbers in order to extract the work experience. I predicted the experience category by binning the extracted years.

In order to calculate the score for the cover letter screening, I will first extract the text from the PDF and then define the keywords that will be matched with the document. In addition, I will lowercase the text, remove stopwords, remove alpha-numeric data, remove handles, remove punctuation, stem, and tokenize the extracted text as part of my preprocessing steps. Additionally, after utilizing sentence transformers, I receive the result. In order to obtain the dimensions of the matrix's orientation and the cover letter score on the similarity index, I lastly used cosine similarity and torch. If the similarity score is **higher than 0.4**, the candidate is considered to be a strong cultural fit; if it is less than 0.4, the candidate is not.

### 4.1.2 Chatbot

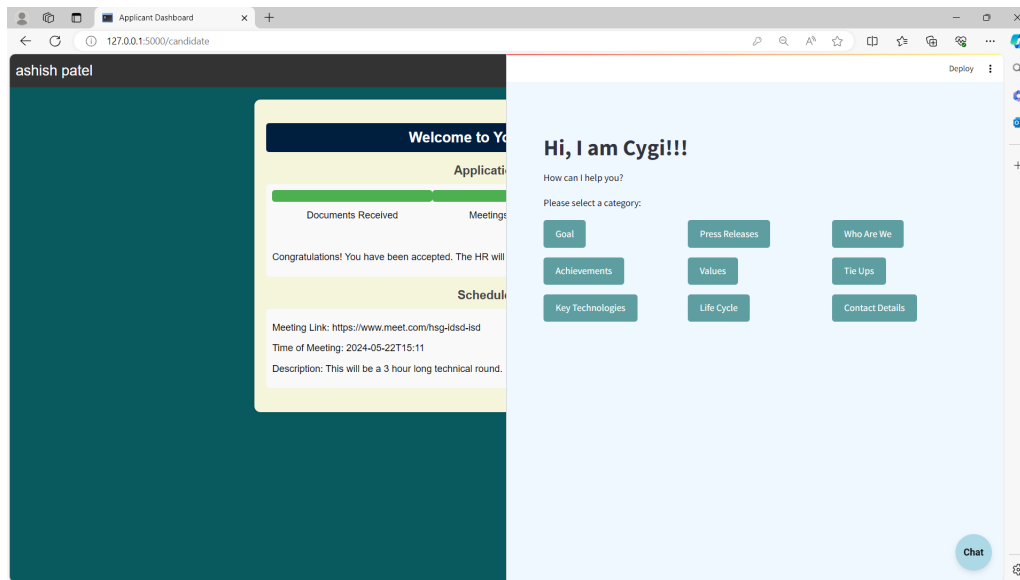


Figure 4.1: Chat-bot

Conditional statements are used in the definition of the "get response" function, which classifies queries and answers. The purpose of creating categories and subcategories is to effectively arrange the content on the chatbot page. This methodology optimizes user interactions by furnishing pertinent data according to their queries. By means of methodical classification, the chatbot optimizes user experience and promotes smooth communication.

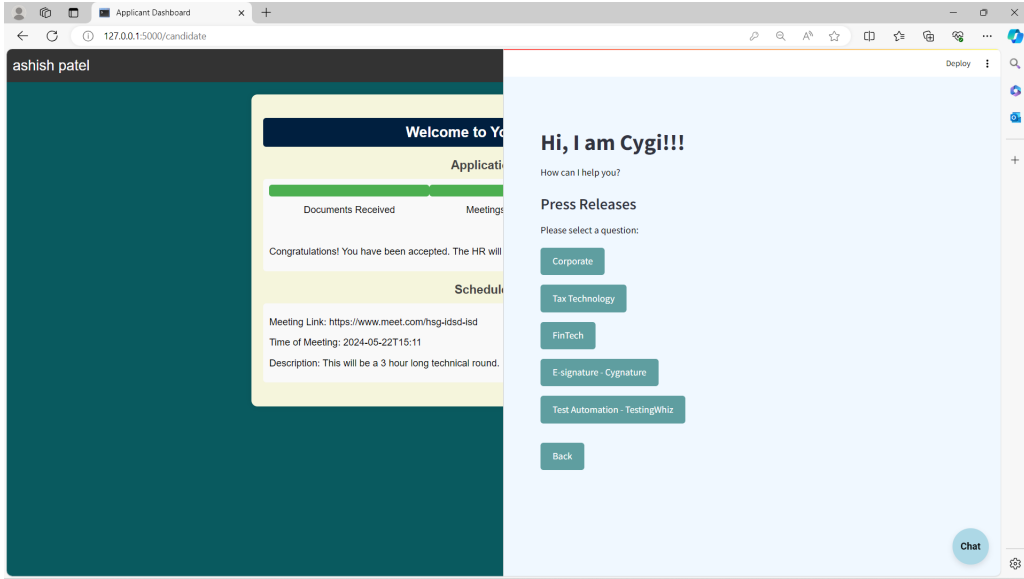


Figure 4.2: Chat-bot Demo

Users can choose from predefined categories and subcategories on a Streamlit page that is designed to display them. The matching answer is obtained after selection. Furthermore, the page's aesthetics can be customized, and its visual appeal and usability are improved through the use of HTML and CSS.

### 4.1.3 Employee Review

Counting the employee ratings in the following categories: Overall Ratings, Work Balance Stars, Culture Value Stars, Career Opportunity Stars, Company Benefit Stars, and Senior Management Stars is done by reading the CSV file and creating visual graphs.

I combined textual columns (pros, negatives, management suggestions, and overall review) to form a new column called "Employee Summary," which summarizes the most important findings. To improve data clarity, the initial columns are subsequently removed. Moreover, preprocessing methods such as lowercasing, tokenization, lemmatization, stopwords, handles, punctuation, and alphanumeric characters are used to polish the text before analysis, guaranteeing precise and effective natural language processing processes.

The sentiments of the combined text from "employee\_summary" are evaluated using TextBlob and polarity scores in order to label the unlabeled data for sentiment analysis. Insights regarding sentiment orientation are provided by polarity scores, which help classify the data into positive, negative, and neutral feelings. Sentiment analysis models are

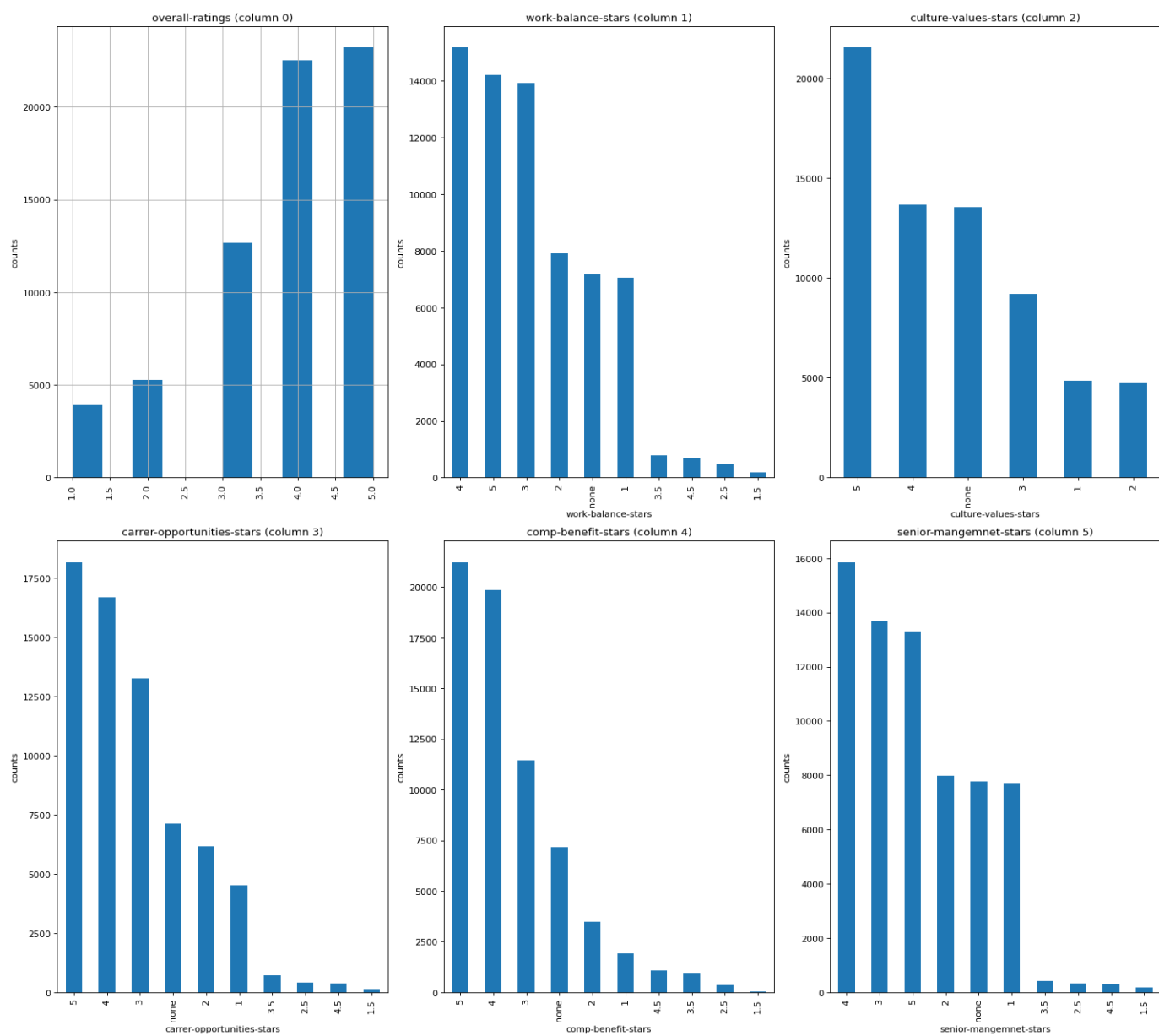


Figure 4.3: Employee Review Graphs

developed using these tagged sentiments as training data.

An imbalance in the distribution of sentiments was found during data analysis, with a preponderance of positive sentiments. Undersampling was used during model training to lessen this imbalance. To produce a more balanced distribution, this included first randomly rearranging the data and then lowering the number of occurrences in the majority class (positive feelings). To guarantee that the model is trained successfully across all sentiment categories, undersampling is used.

Using TF-IDF vectorization, the text inputs were converted into numerical representations to make text analysis by the machine learning model easier. A train-test split with a ratio of 75:25 was then established. The optimal parameters for the defined model, called RV-SGDC [2] (a combination of Stochastic Gradient Descent, Support Vector Machine, and Logistic Regression), were chosen by rigorous voting. It was also trained using the training set of data. Test accuracy, precision, and recall were among the evaluation metrics that were calculated to evaluate the model's performance on unobserved data. Finally, I used joblib to save the trained model and used the model to predict a random string.

#### **4.1.4 Employee Promotion**

After reading the CSV file and performing a null value analysis, it was discovered that the "Education" and "Previous\_Year\_Ratings" fields contained missing data. They were kept on because null values in "Previous\_Year\_Ratings" might be related to recent hires. The mode of the domain to which the employee belongs was used to replace any null entries in the "Education" column, though, because education information is crucial and cannot be null. After that, the CSV file's preprocessing was finished, guaranteeing the accuracy of the data and its suitability for additional analysis.

An examination of the data was done to find out what influences employee promotions. In order to train the model, features that greatly influence promotion were identified, with a specific emphasis on these important variables.

Employee promotion is significantly influenced by both awards received and KPIs met, according to the analysis in Figure 4.1. Both promoted and non-promoted employees' performance is evaluated using a combination of these metrics.

The average training score \* number of trainings yields the total training score, which is



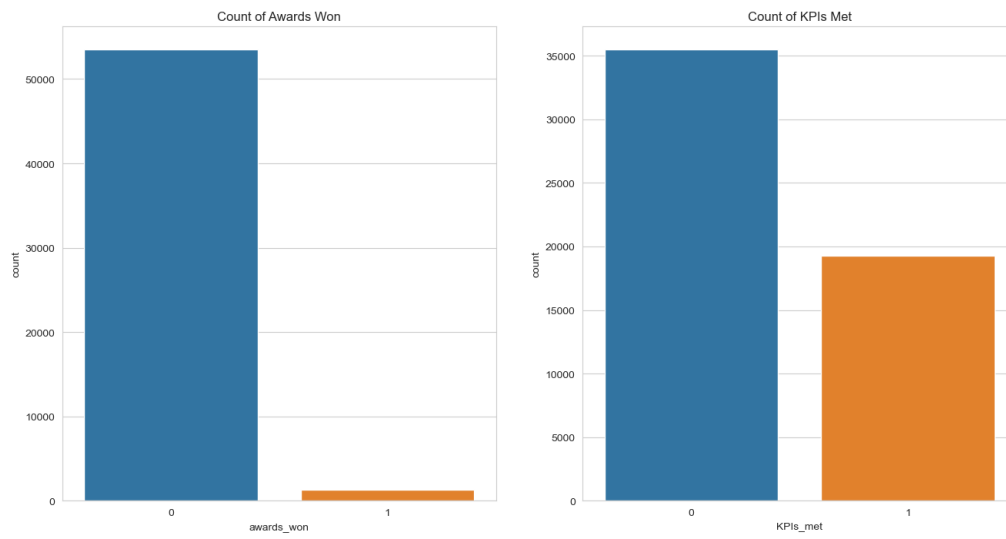


Figure 4.4: KPIs met and Awards won

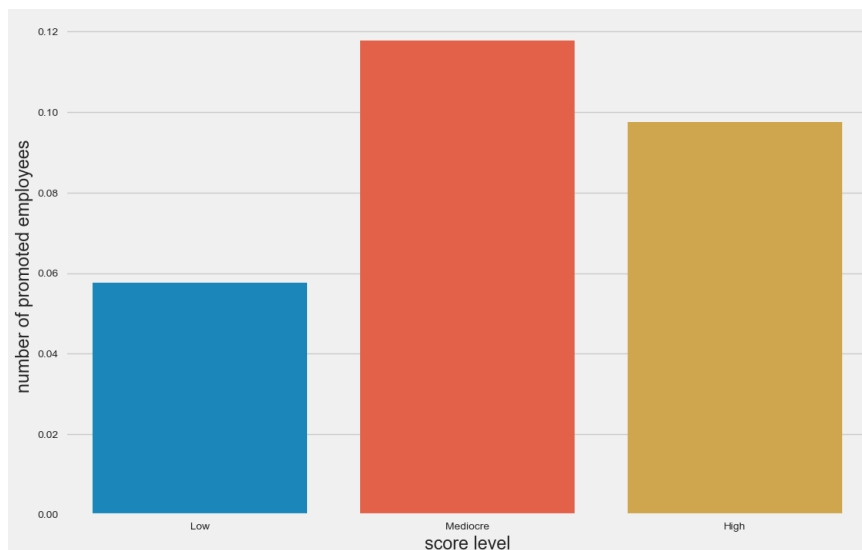


Figure 4.5: Score Graph

shown in Figure 4.2 as a critical element for promotion. The mediocre score group shows the highest rates of promotion.

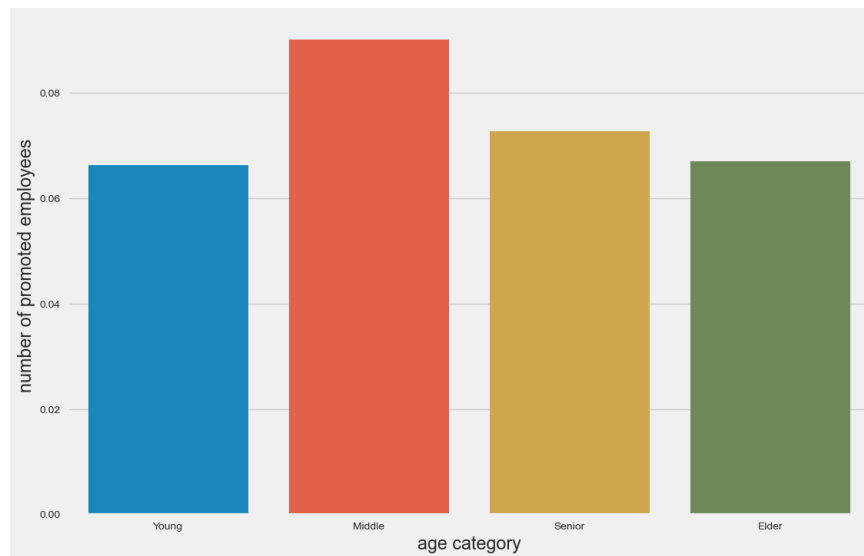


Figure 4.6: Age Category

Age is a substantial but biased determinant for promotion, according to Figure 4.3 study, and should be disregarded.

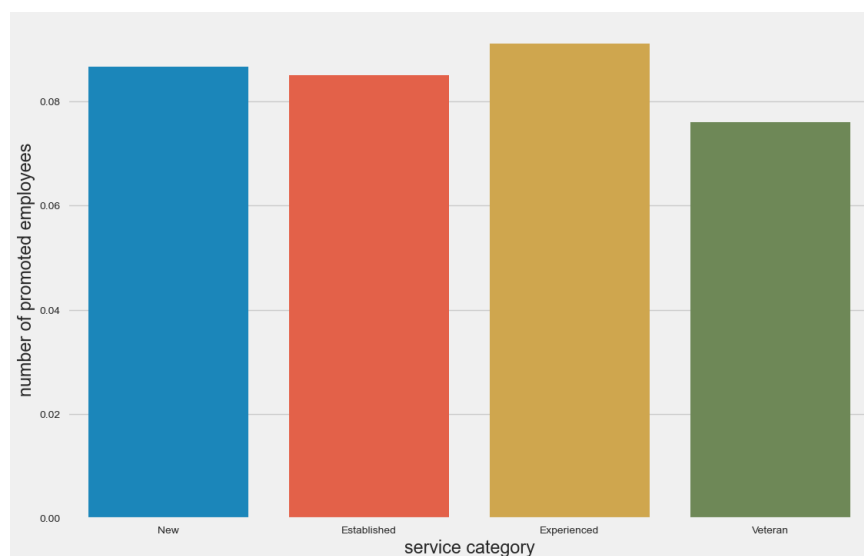


Figure 4.7: Service Category

Figure 4.4 shows length of service is deemed insignificant for promotion.

Additional information is shown in Figure 4.5, which shows that although the Sales & Marketing department has the highest promotion rate, this is dependent more on department size than on internal departmental traits. Promotion possibility is therefore not solely influenced by departmental membership.

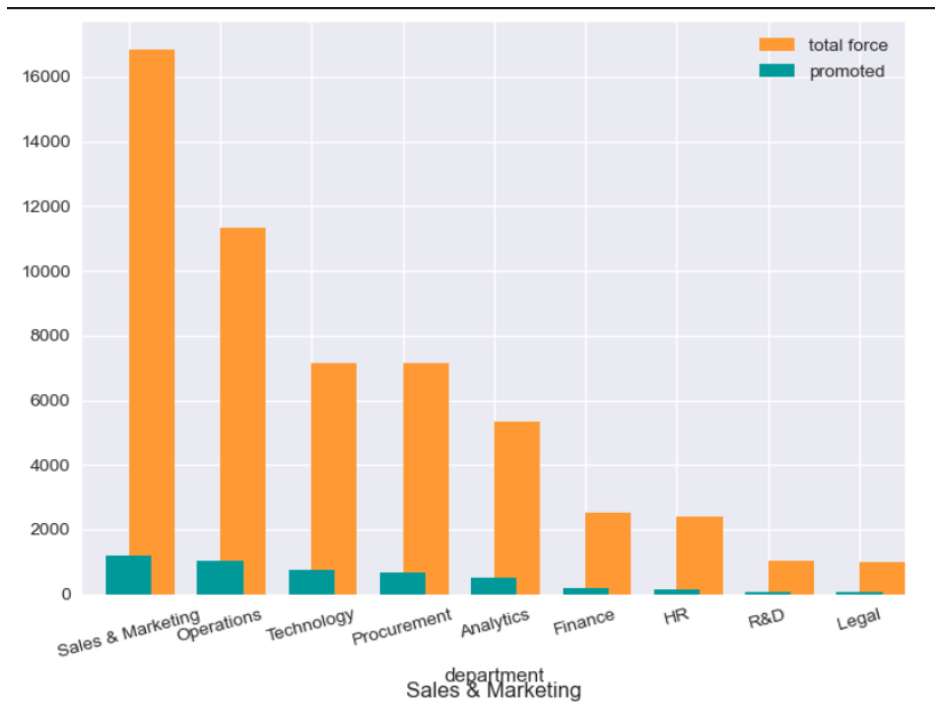


Figure 4.8: Department Category

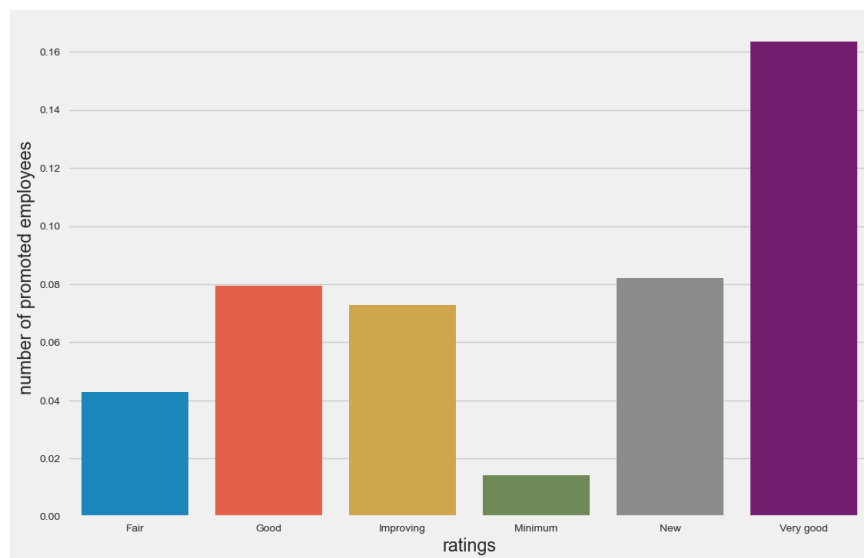


Figure 4.9: Previous Year Rating Category

The importance of prior year ratings in employee advancement is highlighted by Figure 4.6 analysis, which shows a clear preference for employees with a "very good" rating.

Unnecessary columns were eliminated from the dataset prior to model evaluation. Department, region, education, gender, recruitment\_channel, service category, and age

label are among the columns that have been removed. I used random oversampling and Min Max scaling to lessen the data imbalance. This strategy guarantees equity for both promoted and non-promoted staff members. In order to protect data integrity and stop information loss, oversampling was chosen.

Gradient boosting classifiers with an 80:20 train-test split were used to train the model. The train and test accuracies were then calculated, and a classification report was produced. The prediction performance of the model is rigorously evaluated and validated thanks to this scientific approach.

### 4.1.5 Front-End

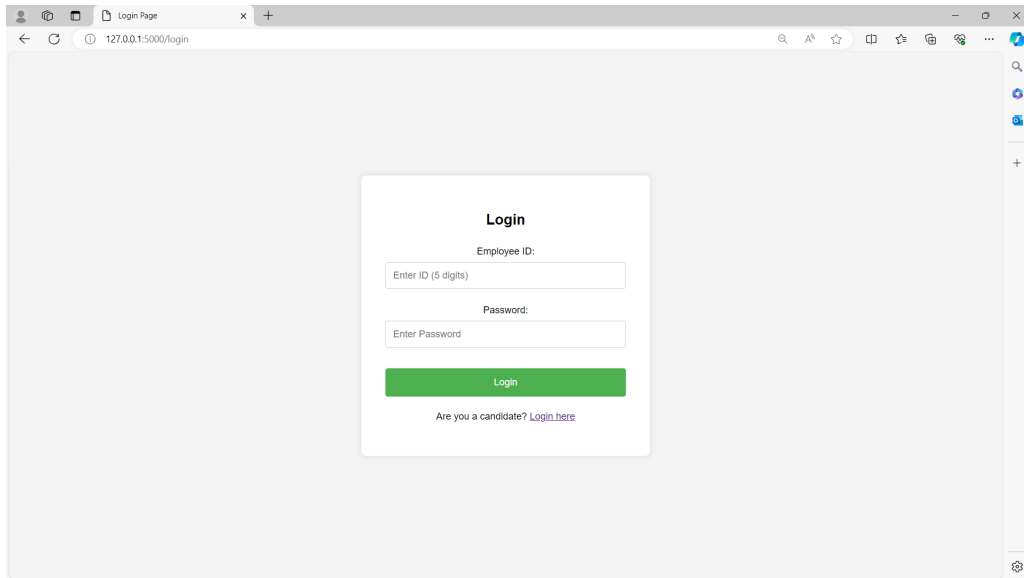


Figure 4.10: Login Page

The AVP Dashboard shows a roster of employees eligible for promotion based on uploaded CSV data that includes metrics like KPIs met, awards received, average training attendance, training scores, tenure, and age. The promotion prediction module makes the prediction. The table, which lists personnel IDs and department names, gives the AVP the final say over whether to approve or reject promotions. With the use of thorough data analysis, this methodical technique guarantees effective administration of promotion decisions and promotes strategic people developments within the company.

The buttons to change your password and log out are located in the rightmost area of the top bar in each module, including candidate, admin, AVP, and HR.

Resume File Name	Cover Letter File Name	Name	Contact Info	Email	Domain	Work Experience	Skills	Cultural Fit	Upload Date	Result	Meets
Ashish Patel Data Scientist 7.pdf	braintree.pdf	ashish patel	8460238402	ashishpatel.ce.2011@gmail.com			<ul style="list-style-type: none"> <li>• matplotlib</li> <li>• python</li> <li>• machine learning</li> <li>• pandas</li> <li>• data science</li> <li>• postgresql</li> <li>• big data</li> <li>• keras</li> <li>• requirement gathering</li> <li>• php</li> <li>• mongodb</li> <li>• data mining</li> <li>• pytorch</li> <li>• opencv</li> <li>• deep learning</li> <li>• scipy</li> <li>• data analysis</li> <li>• programming</li> <li>• problem-solving</li> <li>• hadoop</li> <li>• training</li> <li>• tensorflow</li> <li>• numpy</li> <li>• r</li> </ul>	0.33	2024-04-15 11:18:09.599000	<div>Accept</div> <div>Reject</div>	<div>Schedule</div>

Figure 4.11: HR Dashboard

Welcome to Your Dashboard

Application Status

Your application status will be displayed here.

Scheduled Meets

Your meeting schedule will be displayed here.

Upload Documents

Please upload your resume and cover letter

The format of the files must be pdf

Keep the filename as your fullname.pdf (firstname\_lastname.pdf)

Upload Resume (PDF only):  
[Choose File] No file chosen

Upload Cover Letter (PDF only):  
[Choose File] No file chosen

Submit

Chat

Figure 4.12: Candidate Dashboard

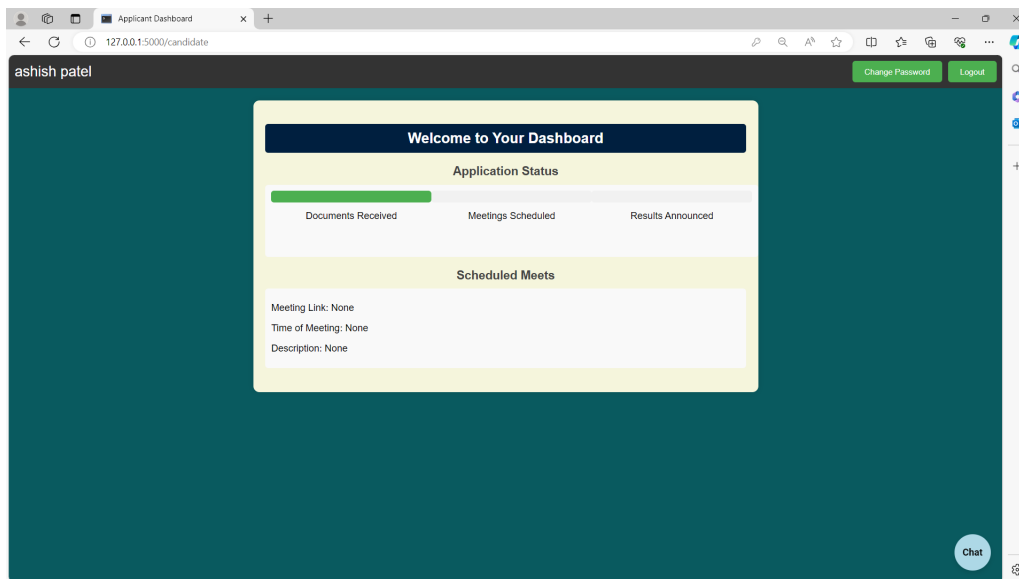


Figure 4.13: Application Status

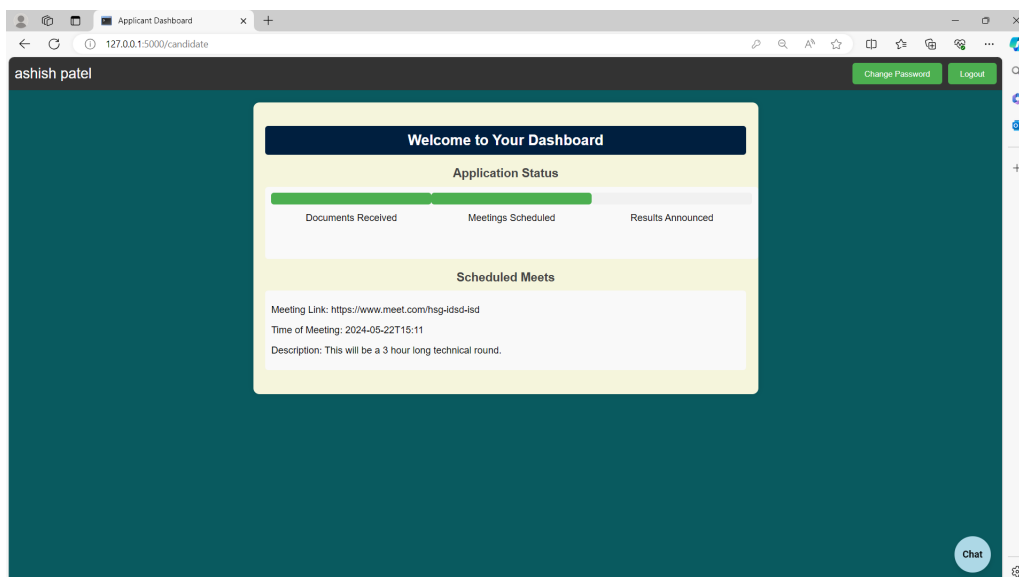


Figure 4.14: Meeting Scheduled

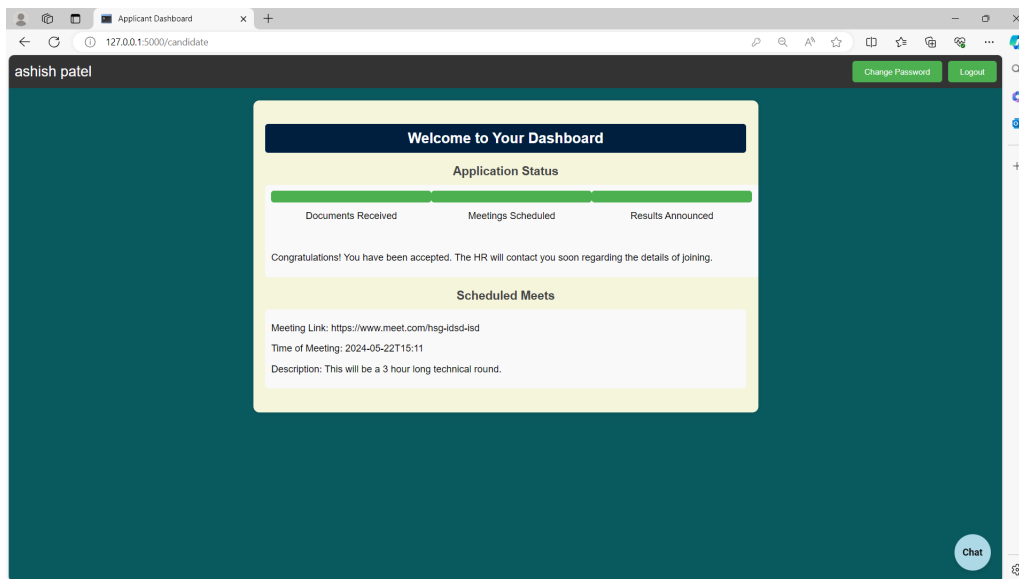


Figure 4.15: Result Announced

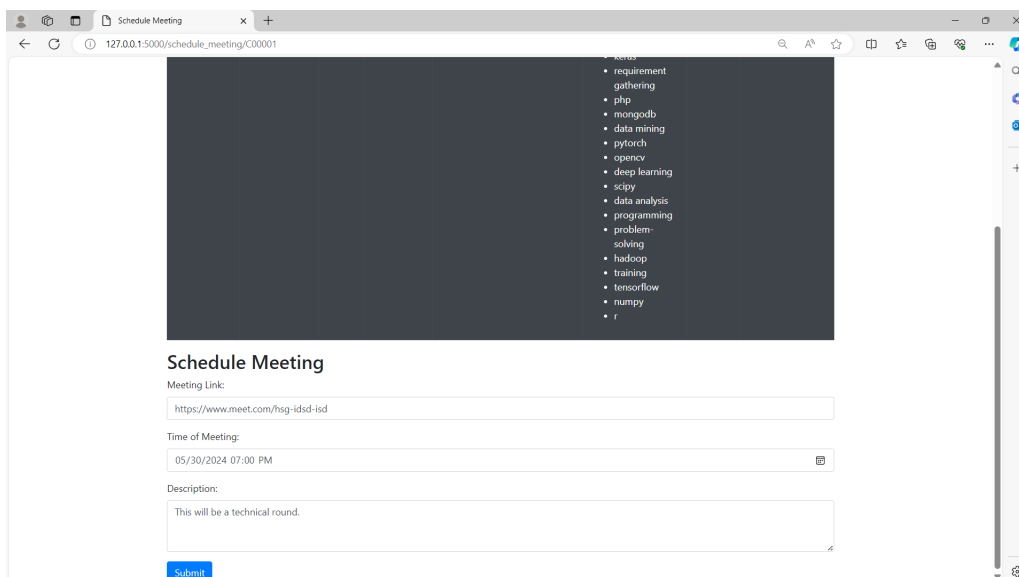


Figure 4.16: HR-Schedule Meet

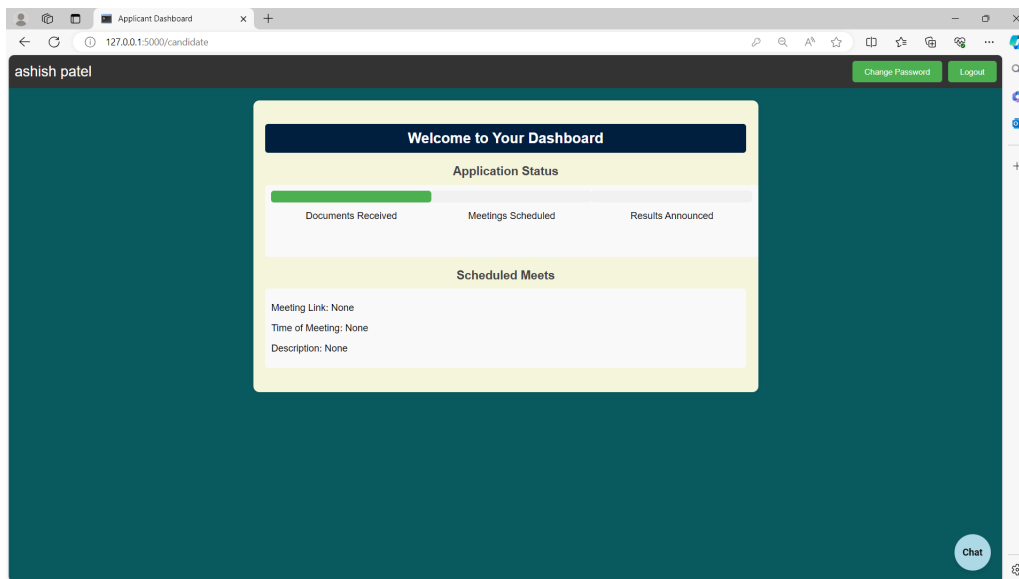


Figure 4.17: Application Status

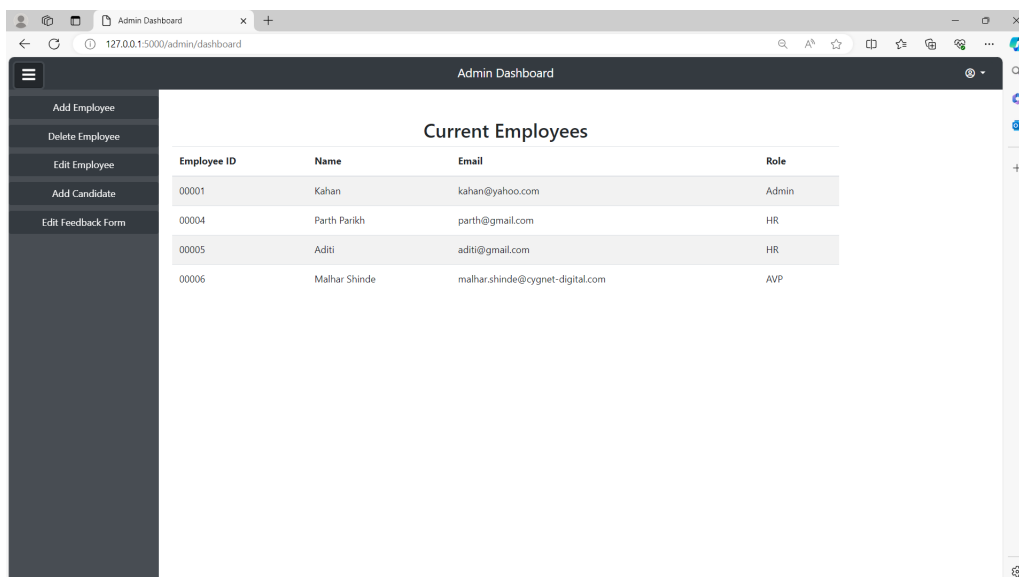


Figure 4.18: Admin Dashboard



**Add Employee**

Employee ID: 00007

Name: Malhar Shinde

Department: Sales & Marketing

Designation: AVP

Email: malhar.shinde@cygnet-digital.com

Phone Number: 08141044073

Password: .....

Add Employee

Figure 4.19: Add Employee

**Delete Employee**

Employee ID	Name	Designation	Email	Delete
00001	Kahan	Admin	kahan@yahoo.com	Delete
00004	Parth Parikh	HR	parth@gmail.com	Delete
00005	Aditi	HR	aditi@gmail.com	Delete
00006	Malhar Shinde	AVP	malhar.shinde@cygnet-digital.com	Delete

Figure 4.20: Delete Employee

The screenshot shows a web browser window with the address bar displaying '127.0.0.1:5000/admin/dashboard/edit\_employee'. The page has a dark header bar with a 'Back' button. The main content area is titled 'Edit Employee'. It contains a form with the following fields: 'Employee ID' (empty), 'Get Details' (button), 'Employee ID' (00001), 'Name' (Kahan), 'Designation' (Admin), 'Email' (kahan@yahoo.com), and 'Phone' (9999999990). A 'Save Changes' button is at the bottom of the form.

Back

### Edit Employee

Employee ID

Get Details

Employee ID

00001

Name

Kahan

Designation

Admin

Email

kahan@yahoo.com

Phone

9999999990

Save Changes

Figure 4.21: Edit Employee

The screenshot shows a web browser window with the address bar displaying '127.0.0.1:5000/admin/dashboard/addcand'. The page has a dark header bar with a 'Back' button. The main content area is titled 'Add Candidate'. It contains a form with the following fields: 'Candidate ID:' (C00007), 'Designation:' (Candidate), and 'Password:' (Admin123@). An 'Add Candidate' button is at the bottom of the form.

Back

### Add Candidate

Candidate ID:

C00007

Designation:

Candidate

Password:

Admin123@

Add Candidate

Figure 4.22: Add Candidate

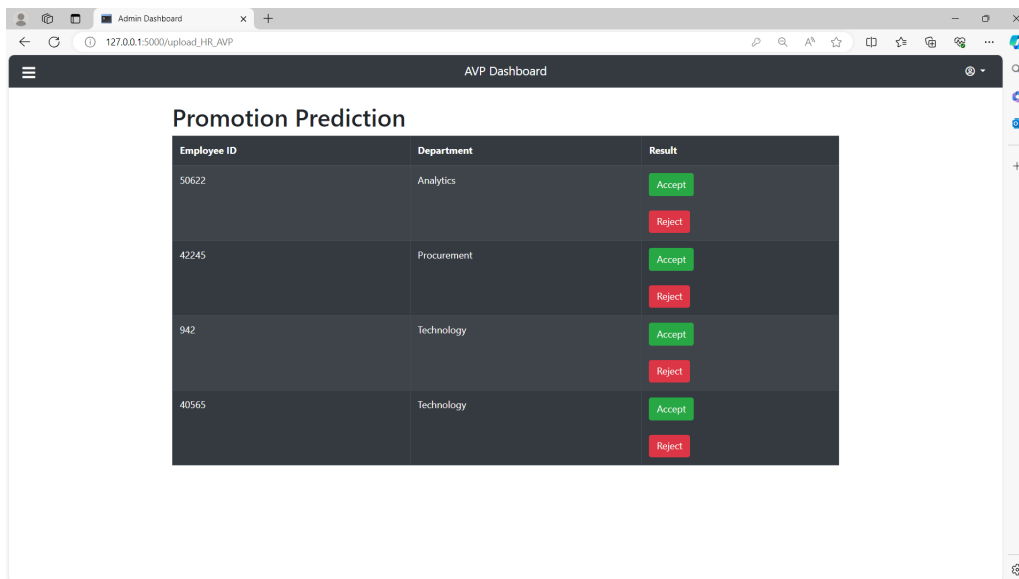


Figure 4.23: AVP Dashboard

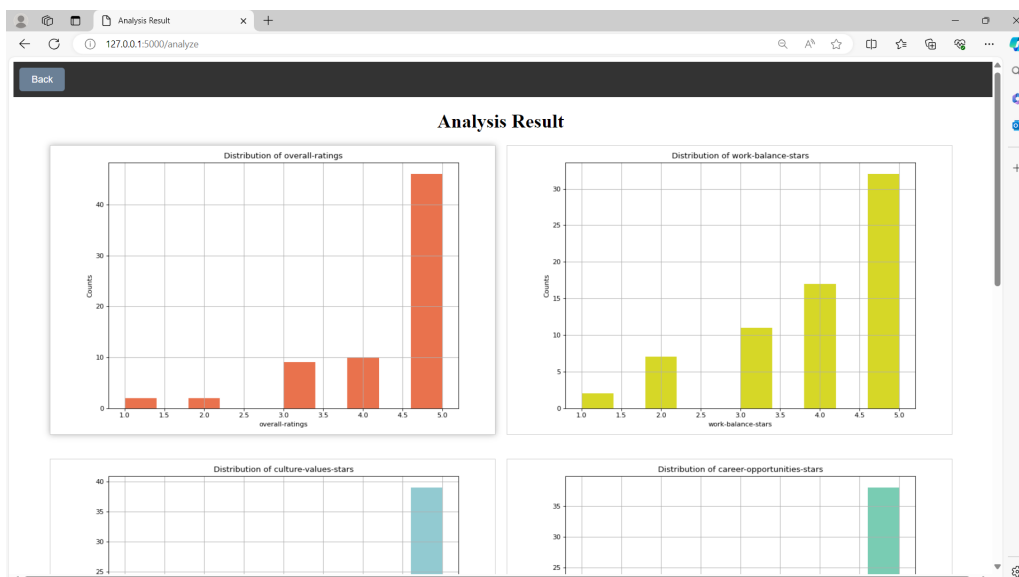


Figure 4.24: Feedback Statistics

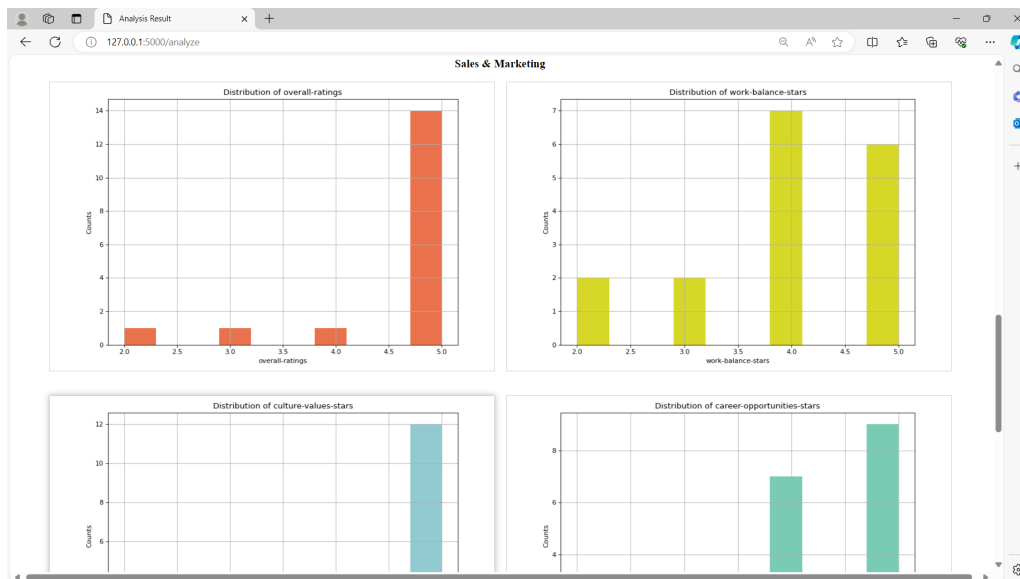


Figure 4.25: Feedback Statistics

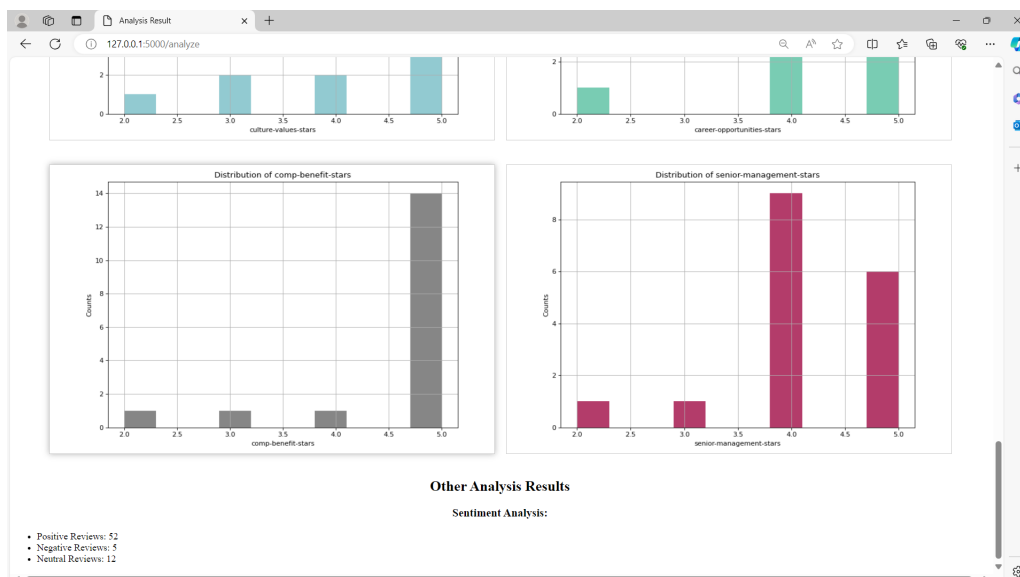


Figure 4.26: Feedback Statistics

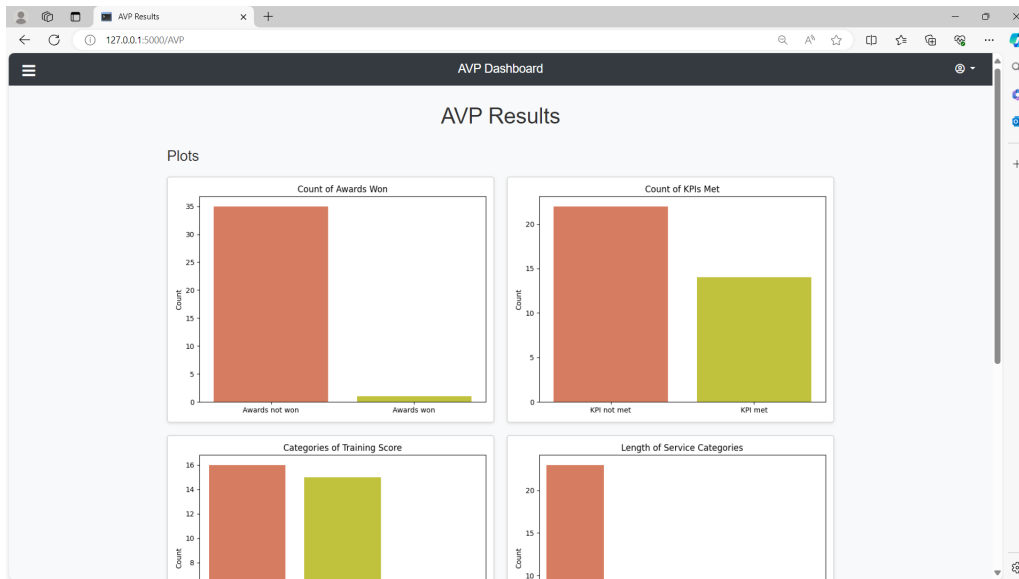


Figure 4.27: Promotion Statistics

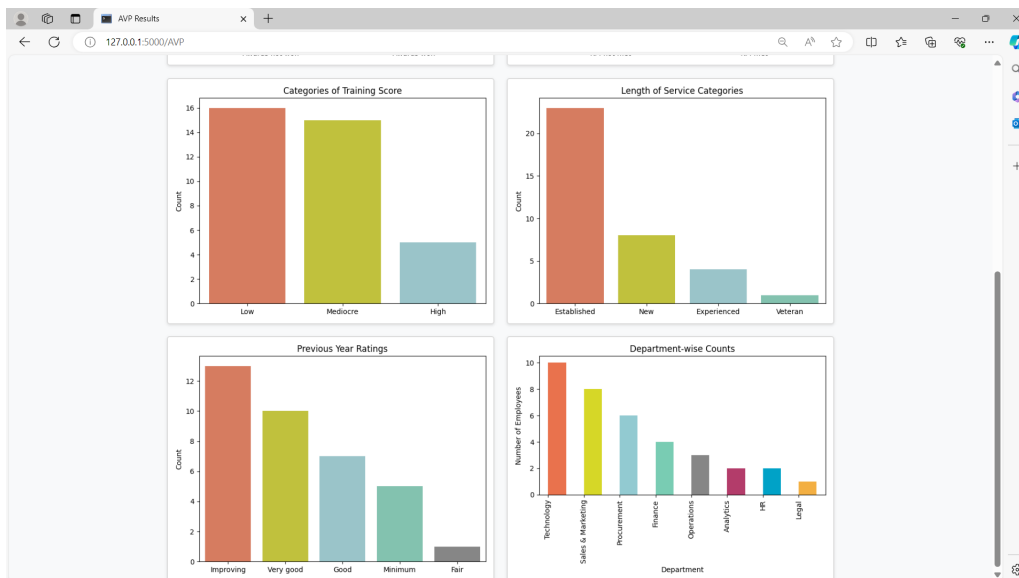


Figure 4.28: Promotion Statistics

# Chapter 5

## Conclusion and Future Work

### 5.1 Conclusion

The incorporation of AI-driven HR analytics is a revolutionary force in the field of contemporary human resources, changing traditional methods and bringing in a new era of workforce management. By utilizing advanced technologies like sentiment analysis, natural language processing, and predictive analytics, businesses may improve and expedite a range of HR procedures.

Much time and money can be saved by automating processes like resume screening and arranging interviews, all the while maintaining impartiality and fairness in the candidate selection process. This speeds up the hiring process and improves the candidate experience in general, paving the way for future interactions between the candidate and the company. Additionally, HR specialists may proactively address retention issues and promote a more engaged and devoted workforce by using their skills to monitor employee attitude and forecast attrition risks. Organizations can lower employee turnover and increase employee satisfaction by implementing targeted interventions and efforts before potential issues worsen. Furthermore, AI-enabled HR analytics' real-time data visualization capabilities provide priceless insights into workforce dynamics and trends. HR workers are able to make well-informed decisions and plan strategically by gaining a thorough understanding of critical data through interactive dashboards and visual representations.

The advantages of AI in HR analytics are evident, even in the face of obstacles with data integration and quality. Organizations can get a competitive advantage in talent recruit-

ing, retention, and management by adopting these cutting-edge technology. Adoption of AI-driven HR analytics becomes important for sustaining development and profitability as firms traverse an ever-more complicated and dynamic landscape. Organizations may fully utilize their human resources by utilizing AI, which will drive innovation, productivity, and organizational excellence in the rapidly changing digital era.

## 5.2 Future Scope

The AI-enabled HR analytics project has a bright future ahead of it, full with possibilities. First, as AI and machine learning algorithms continue to progress, there will be chances to improve the precision and effectiveness of HR procedures even more. While advances in predictive analytics can enable more precise detection of retention concerns and talent development possibilities, improvements in natural language processing models can allow for deeper insights from candidate resumes and cover letters.

Additionally, the use of cutting-edge technology like computer vision and sentiment analysis from multimedia sources opens up new possibilities for more thorough employee feedback analysis and candidate assessments. HR analytics may improve decision-making processes by delving deeper into candidate personalities, cultural fit, and employee sentiments by combining visual data from social media platforms and video interviews.

Furthermore, the project's future scope will cover more HR-related tasks like performance management, learning and development, and workforce planning in addition to recruitment and retention. Personalized learning routes, skill gaps, and future labor demands based on changing company requirements and industry trends can all be supported by AI-driven solutions.

Collaboration with other organizational divisions, such operations and finance, to use HR analytics for cross-functional insights and strategic alignment, is another possibility. Organizations may optimize resource allocation, reduce risks, and promote overall organizational performance by integrating HR data with KPIs and business measures.

Finally, to address issues with algorithmic bias, data security, and regulatory compliance, continued developments in data privacy and ethics frameworks will be essential. Organizations may cultivate trust and confidence in the application of AI-enabled HR analytics and pave the path for future sustainable growth and innovation by placing a high priority on openness, accountability, and ethical AI practices.

# Appendices

1. <https://www.anaconda.com/download/>
2. <https://www.coursera.org/specializations/natural-language-processing>
3. Live NLP Playlist - YouTube
4. Convolutional Neural Networks — Coursera
5. Sequence Models — Coursera
6. What is Scrum? [+ How to Start] — Atlassian
7. What is HR Analytics? — Definition, Examples, Metrics (valamis.com)
8. AI in HR Analytics: Benefits, Challenges & Ways to Integrate (pockethrms.com)
9. MongoDB Tutorial (w3schools.com)
10. <https://www.kaggle.com/datasets/gauravduttakiit/resume-dataset>
11. ShashiVish/cover-letter-dataset · Datasets at Hugging Face
12. Pradeep Kumar Roy, Sarabjeet Singh Chowdhary, Rocky Bhatia - A Machine Learning approach for automation of Resume Recommendation system
13. Python RegEx (w3schools.com)
14. <https://www.pragnakalp.com/case-study/nlp-resume-parser-bert-python/>
15. <https://medium.com/@2020.chetaniya.bajaj/building-a-resume-parser-using-nlp-dd36>
16. <https://huggingface.co/sentence-transformers>
17. <https://www.geeksforgeeks.org/cosine-similarity/>



18. We are Cygnet, We are One
19. Sangita Pokhrel, Swathi Ganesan, Tasnim Akther, Lakmali Karunaratne - Building Customized Chatbots for Document Summarization and Question Answering using Large Language Models using a Framework with OpenAI, Lang chain, and Streamlit
20. <https://www.youtube.com/watch?v=qM6aqeXGkWQ>
21. Generative AI with Large Language Models — Coursera
22. Employees-Reviews dataset (kaggle.com)
23. Babacar Gaye, Dezheng Zhang, and Aziguli Wulamu - Sentiment classification for employees reviews using regression vector- stochastic gradient descent classifier (RV-SGDC)
24. <https://towardsdatascience.com/predicting-sentiment-of-employee-reviews-ec0c0c83>
25. <https://raw.githubusercontent.com/rajtulluri/Employee-Promotion-Prediction/master/employeePromotion.csv>
26. <https://www.youtube.com/watch?v=FX6QtI1R6bI>
27. HTML Tutorial (w3schools.com)
28. <https://www.w3schools.com/css/default.asp>

# Bibliography

- [1] P. K. Roy, S. S. Chowdhary, and R. Bhatia, “A machine learning approach for automation of resume recommendation system,” *Procedia Computer Science*, vol. 167, p. 2318–2327, 2020.
- [2] B. Gaye, D. Zhang, and A. Wulamu, “Sentiment classification for employees reviews using regression vector- stochastic gradient descent classifier (rv-sgdc),” *PeerJ Computer Science*, vol. 7, p. e712, Sept. 2021.