|  |  |
| --- | --- |
| Date | 19TH November |
| Team ID | PNT2022TMID11925 |
| Project Name | Corporate Employee Attrition Analysis |
| Industry mentor | Shanawaz Anwar |
| Faculty mentor | Saravanan S |

**Mmebers:**

**Surendran S**

**Sriram R**

**Suraj S**

**Suriya M**

**Project Report Format**

# 1.INTRODUCTION

1.1 Project Overview

Employee attrition is referred as reduction in number of employees in an organization. For IT industry, employee attrition has become a known challenge since last 2 decades. Employees leave the organization for various reasons. A few reasons are, demand of high salary, change in technology or role, professional challenges etc.

High attrition leads to expense over multiple attributes and functions in the company. Recruitment, Training and Development costs increases overall cost on the employees. The core reason of this attrition could be mismatch in expectations of organization and expectations of employees from each other. This project is used to analyze the attrition reasons as well as understand the expectation of employees from the organization.

1.2 Purpose

Attrition analysis contributes to the details generated by HR managers on employees leaving the company. The metrics offer accuracy in terms of the reasons given by employees themselves. Apart from this, a wider avenue for change and dynamism also emerges from analysis of attrition.

It brings to fore the cause of employee disengagement.

# 2.LITERATURE SURVEY

2.1 Existing problem

A quiet significant amount of works related to the Attrition of Employees using Machine Learning algorithms have been made. An efficient attrition rate prediction has been made by using various algorithms some of them include Logistic Regression, Decison Tree, Random Forest Classifier etc. It can be seen in results that each algorithm has its strength to register the defined objectives.

The model incorporating PAM had the ability to calculate based on various attributes such as age, sex, marital status, education level, work experience, distance from hometown, etc. and generates various levels of risk of attrition.It didnt use any one algorithm but depending on the organizational contexts, different models have to be tried and evaluated before making the final selection. But the accuracy that was obtained in such PDM model was far more less than the new upcoming model .

2.2 References

* Hardik P. K. (2016) , “a study on employee attrition: with special reference to kerala it industry”. IMPACT: International Journal of Research in Business Management. 75-82
* Bodjrenou Kossivi, Ming Xu, Bomboma Kalgora ( May 2016), Study on Determining Factors of

Employee Retention. Open Journal of Social Sciences, Vol.4 No.5, May 30, 2016

* Brijesh Kishore Goswami, Sushmita Jha (April 2012), “Attrition Issues and Retention Challenges of Employees” , International Journal of Scientific & Engineering Research Volume 3, Issue 4, April-2012 1 ISSN 2229-5518
* Vivek Sinha, (March 10, 2011) - Attrition is Indian firms’ new worry – Vivek Sinha, Hindustan Times, (March 10, 2011) Lucknow Edition
* Sabitha Niketh (March 2008 ), Attrition: A Global Problem, HRM Review, March 2008 Issue, Pg. no. 64-67, ICFAI University Press, Hyderabad

2.3. Problem Statement Definition o Over the past two years, this type of analytic practice has become indispensable. Global labour markets have swung dramatically due to the COVID-19 pandemic.

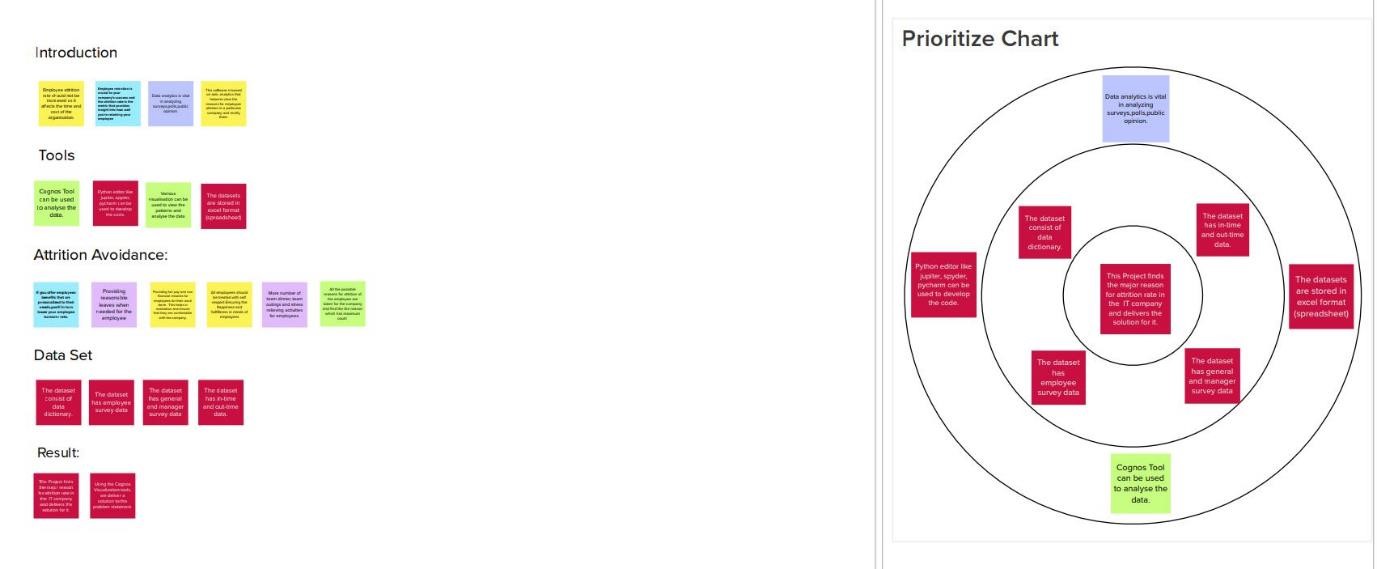
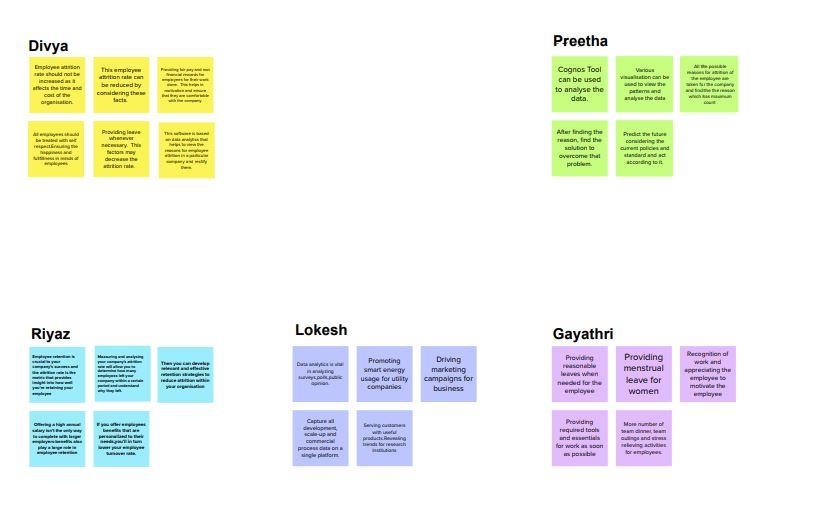
* In addressing the ongoing challenges of the pandemic and the rise of remote work, employee attrition analytics will remain important to organizations seeking to retain top talent.
* Predictive analytics capability enables the design of an employee retention model to keep these valuable employees engaged and on board.
* Employee attrition analytics is specifically focused on identifying why employees voluntarily leave, what might have prevented them from leaving, and how we can use data to predict attrition risk.
* There are actually two types of attrition problems: too little and too much.
* The more talented the worker, the greater the consequences of attrition: Replacing an individual employee typically cost one or two the worker’s annual salary. o Even if a good employee leaves as a result of “graduating” into a job with a client, if they become a great ambassador for the company, it can be a positive loss

# 3.IDEATION AND PROPOSED SOLUTION

3.1.Empathy Map



1. 2.Brainstorming and Ideation Process



* 1. Proposed solution

|  |  |  |
| --- | --- | --- |
| **S.No.** | **Parameter** | **Description** |

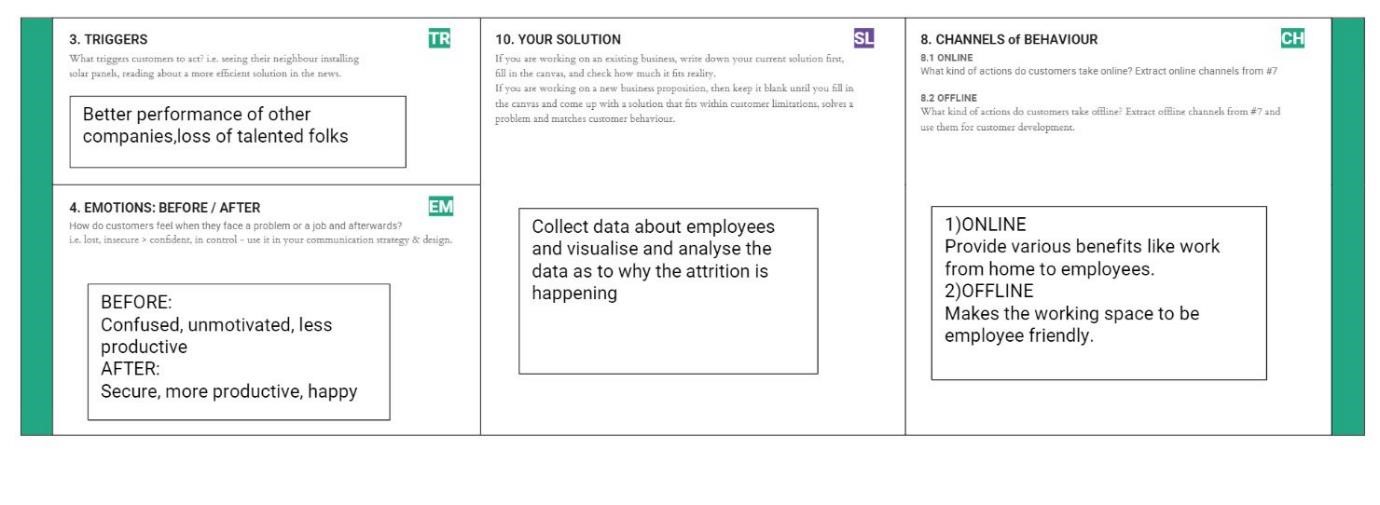
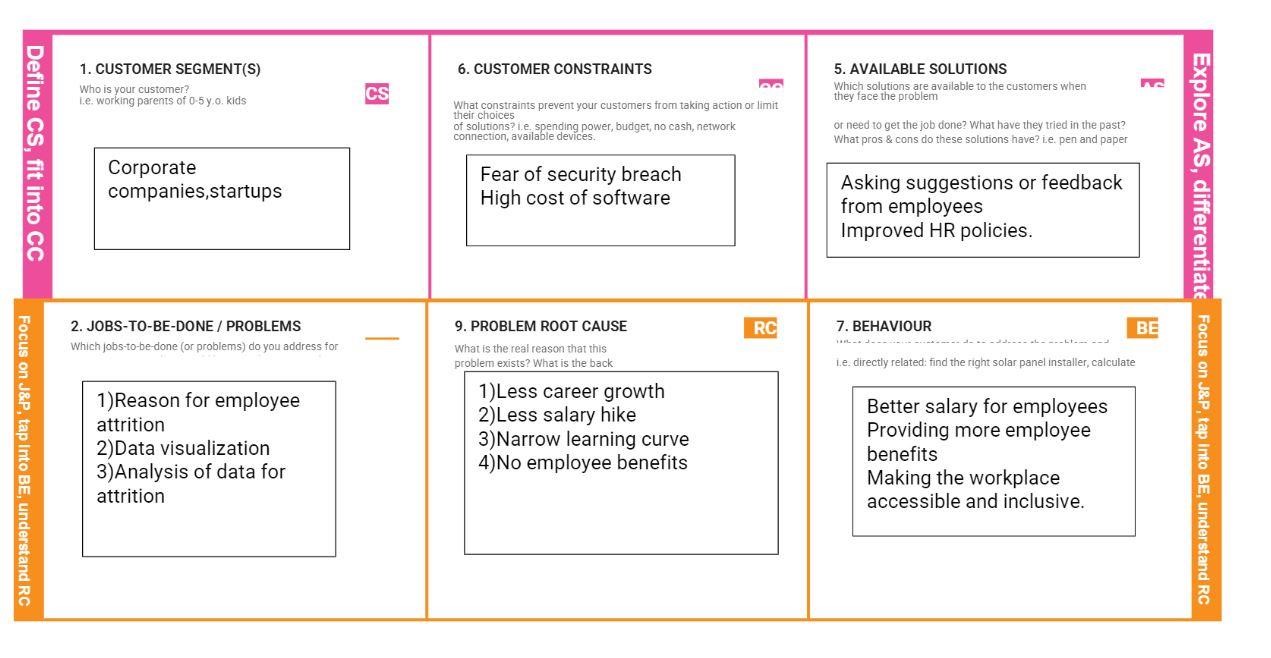
|  |  |  |
| --- | --- | --- |
| 1. | Problem Statement (Problem to be solved) | Over the past two years, this type of analytic practice has become indispensable. Global labour markets have swung dramatically due to the COVID-19 pandemic.  In addressing the ongoing challenges of the pandemic and the rise of remote work, employee attrition analytics will remain important to organizations seeking to retain top talent.  Employee attrition analytics is specifically focused on identifying why employees voluntarily leave, what might have prevented them from leaving, and how we can use data  to predict attrition risk. |

|  |  |  |
| --- | --- | --- |
| 2. | Idea / Solution description | Data analysis can be used to establish internal employee turnover benchmarks. Tracking these benchmarks over time can reveal how the employee experience is changing for better or worse, if the reasons employees are leaving have changed, or if the attrition pattern or time cycle is different. These benchmarks will illustrate whether the actions the organization is taking to reduce attrition are effective, alerting leaders and managers to make adjustments or take different targeted actions if needed. We can use K nearest algorithm to load, visualize, pre process the data .Using KNeighborsClassifier for finding the best number of neighbour with the help of misclassification error. |
| 3. | Novelty / Uniqueness | We consider three types of attritions here and try to solve the problems of overcoming it.    Voluntary- When an employee leaves the company for a better job opportunity or career growth or more pay, and leaves on his own.    Involuntary- If an employee is terminated from a job due to some ethical issue or lack or performance. Sometimes, a degrowing business also forces employees to quit the job, which leads to a higher rate of people leaving.    Retirement- Once an employee finishes his/her tenure at a company and retires. This is mostly a natural attrition that occurs and |
|  |  | companies are prepared with succession planning. |
| 4. | Social Impact / Customer Satisfaction | This helps the corporate in learning the reasons for attrition, understanding different types of attrition, trying to limit the attrition through various techniques. |
| 5. | Business Model (Revenue Model) | This project would be a profitable one for the corporate as Attrition is something which every company faces especially in this post covid period |
| 6. | Scalability of the Solution | Initially this model is focused on a small number of companies in the development phase. Once its successful ,the number of users increases so we can use cloud for higher storage of the large datasets of each company |

* 1. Problem Solution fit

The Problem-Solution Fit simply means that we have found a problem with our customer and that the solution we have realized for it actually solves the customer’s problem. It helps entrepreneurs, marketers and corporate innovators identify behavioural patterns and recognize what would work and why. The purpose is to solve complex problems in a way that fits

the state of your customers and succeed fasterand increase your solution adoptionby tapping int o existing mediumsand channels of behaviour



# 4. REQUIREMENT ANALYSIS

Functional requirement

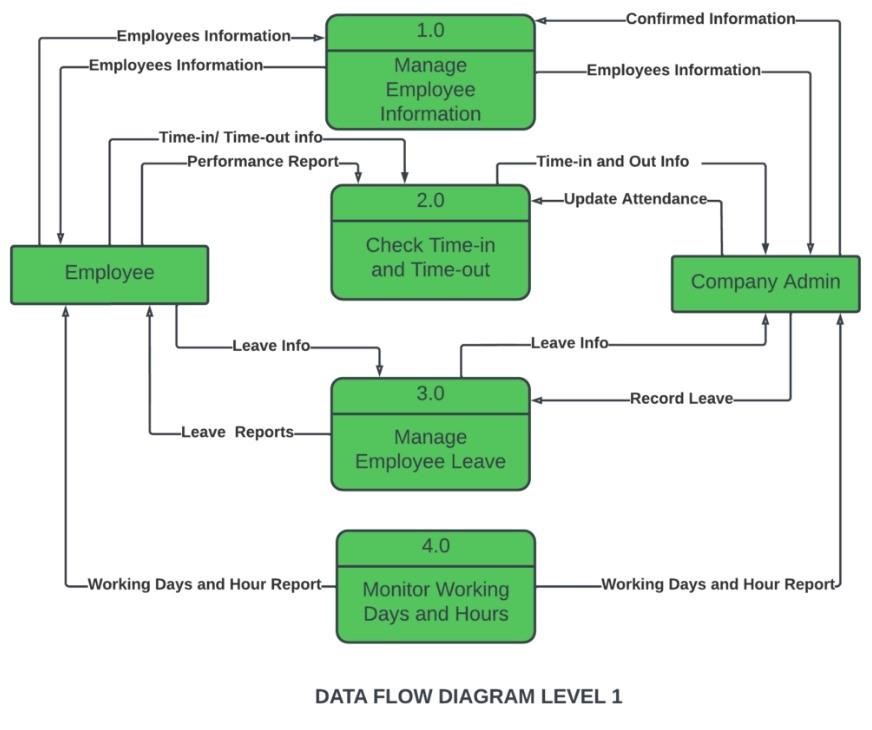
|  |  |  |
| --- | --- | --- |
| **FR No.** | **Functional Requirement (Epic)** | **Sub Requirement (Story / Sub-Task)** |
| FR-1 | User Registration | Registration through Gmail |
| FR-2 | User Authentication and Confirmation | Authenticate the user trying to login using database Confirmation via Email Confirmation via OTP |
| FR-3 | Attrition analysis | Employee attrition analysis by biannual or quarterly performance appraisal, Identifying the team where the major resignation happens  (three types of attrition-voluntary, involuntary, retirement) |
| FR-4 | Employee management | Validating and managing the registered employee details. |
| FR-5 | Progress management | Maintaining the progress of each employee for Appraisals |
| FR-6 | Link | It is used for predicting the likely attrition factors. We’ll be taking the following route for this analysis - Getting our input and storing it, Select the necessary attributes forthe  Prediction, Creating Dashboard, Report  & Stories ,Predicting our results, Showcase the results with the help of dashboard, Report & Stories |

Non-Functional requirements

|  |  |  |
| --- | --- | --- |
| **FR No.** | **Non-Functional Requirement** | **Description** |
| NFR-1 | **Usability** | This Data Visualization shall be easy to use for all users with minimal instructions. 100% of the languages on the graphical user interface (GUI) shall be intuitive and understandable by non-technical users. |
| NFR-2 | **Security** | The user of the system should be provided the surety that their account details are secure. |
| NFR-3 | **Reliability** | The Link shall be operable in all conditions. The system must be less prone to errors. |
| NFR-4 | **Performance** | The performance of the system must assist the system’s quality. |
| NFR-5 | **Portability** | The link shall be portable to all  operating platforms. Therefore, this link  should not depend on the different operating systems. |

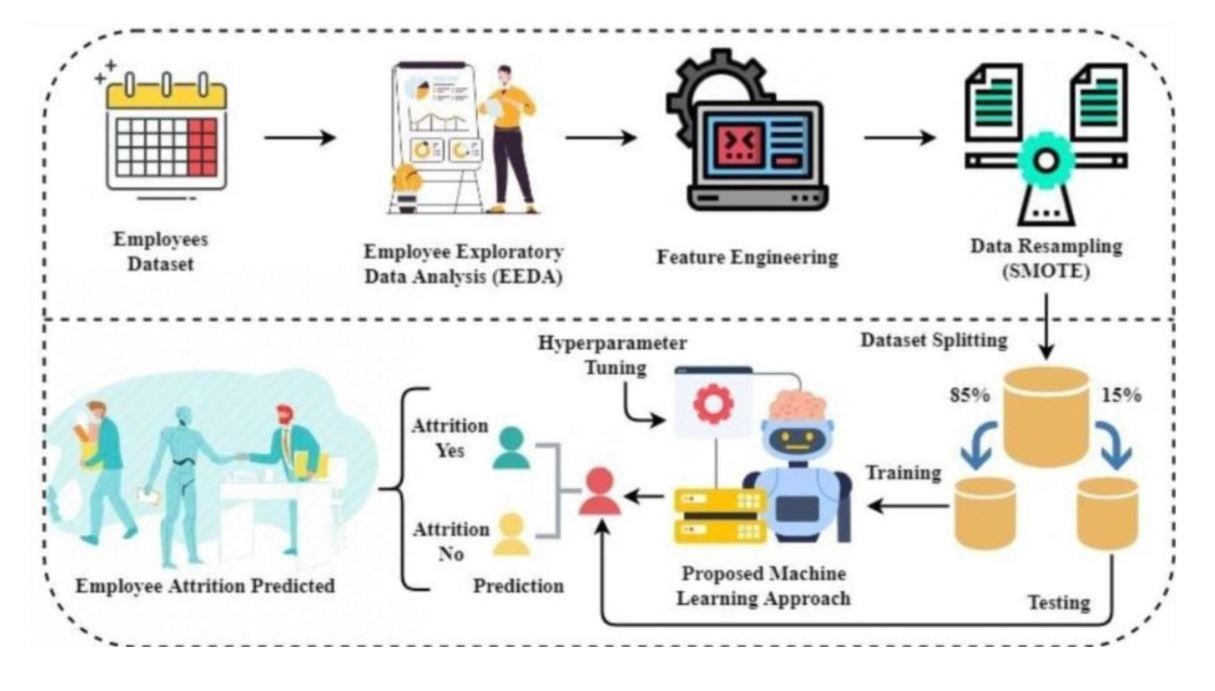
# 5.PROJECT DESIGN

5.1 Data Flow Diagrams



5.2 Solution & Technical Architecture

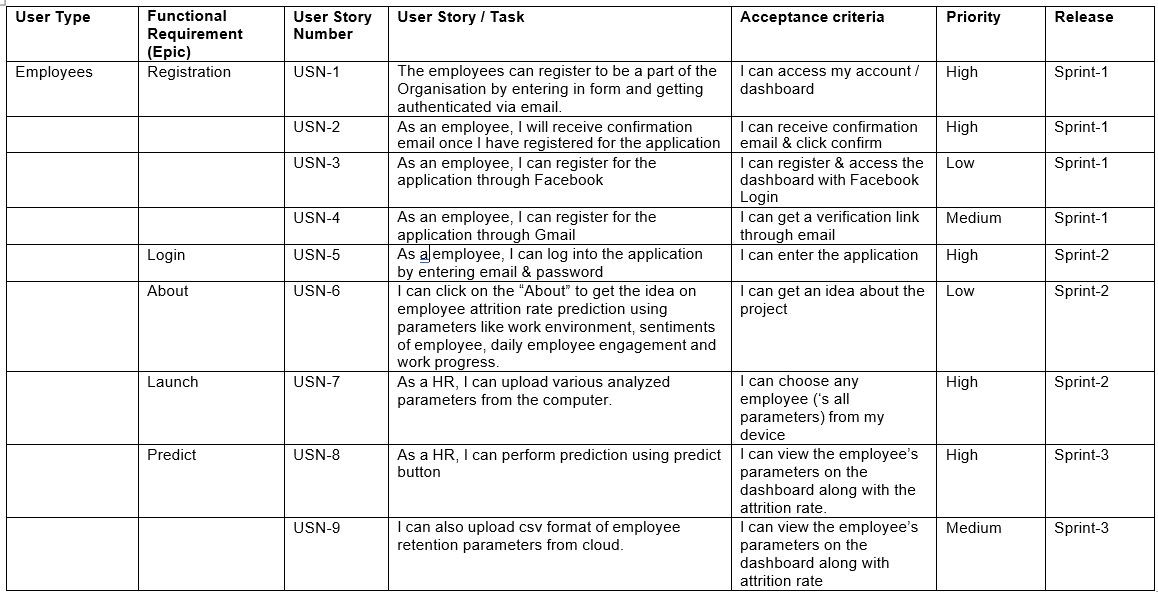
Solution Architecture



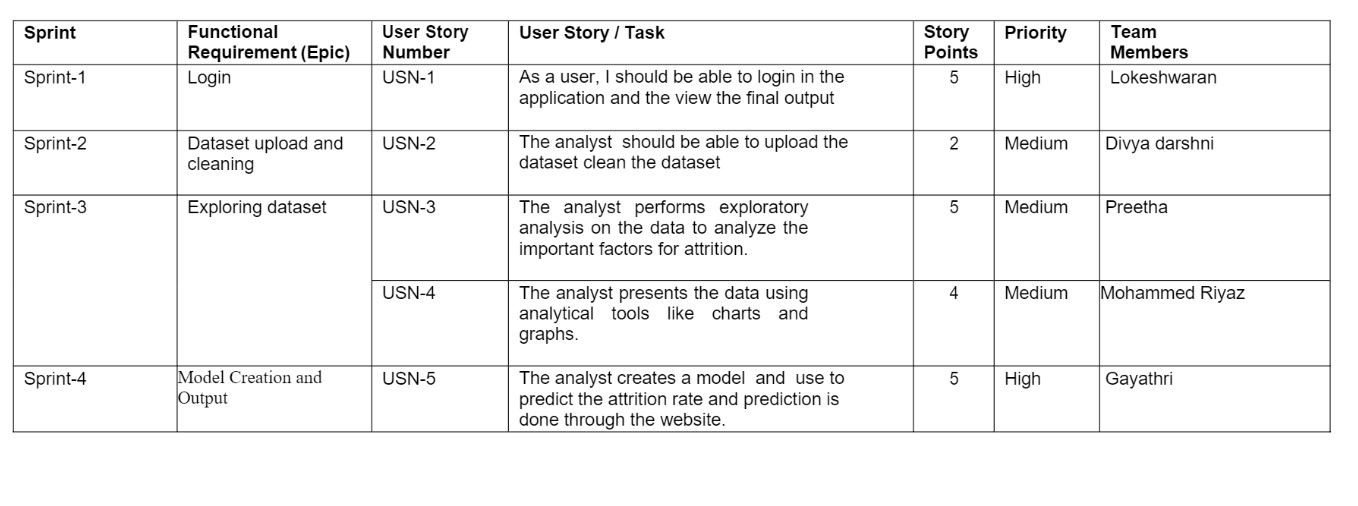
Technical Architecture

|  |  |  |  |
| --- | --- | --- | --- |
| **S.**  **N**  **o** | **Component** | **Description** | **Technology** |
| 1  . | User Interface | How user interacts with application e.g.W eb UI, MobileApp, Chatbot etc. | HTML, CSS, JavaScript / Angul ar Js / ReactJs etc. |
| 2  . | Database | Data Type,Configurations etc. | MySQL, NoSQL,etc. |
| 3  . | Cloud Database | Database Service on Cloud and Storing the datasets uploaded | IBM DB2, IBM Cloudant etc. |
| 4  . | Machine Learning Model | Purpose of MachineLearning Model | Training the model |
| 5  . | Infrastructure (Ser ver / Cloud) | Application Deployment on LocalSystem / CloudLoc al Server Configuration: | Local, Cloud Foundry, Kubern etes, etc. |

5.3 User Stories



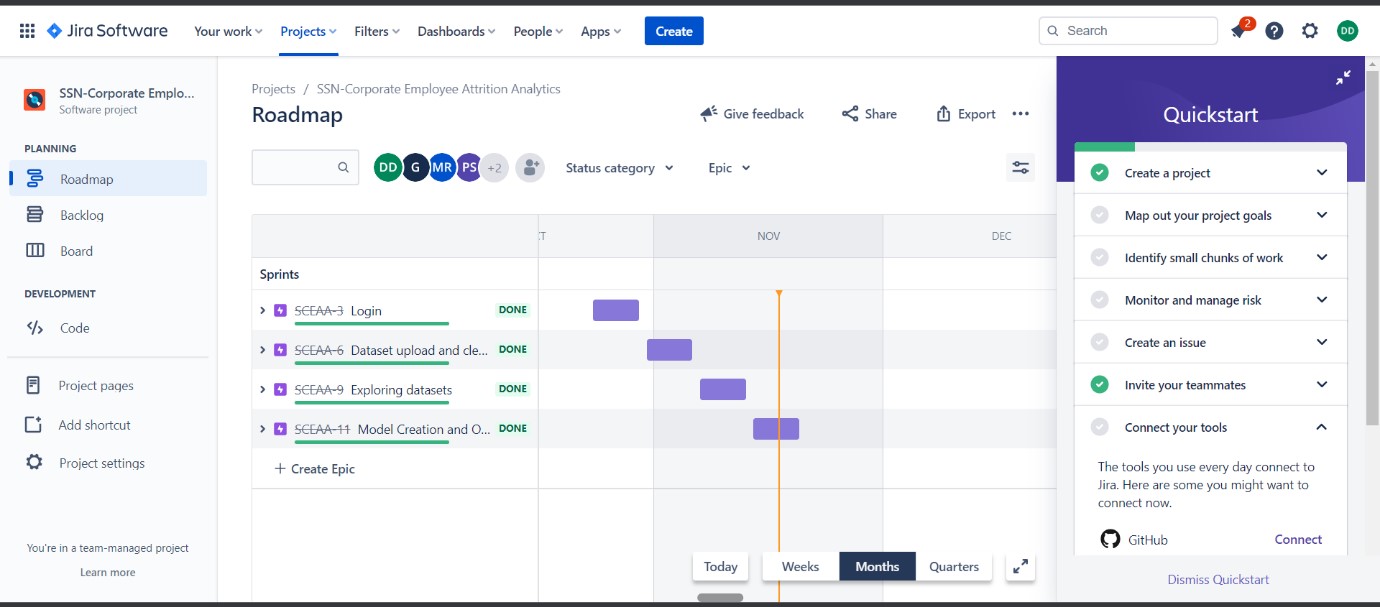
1. **PROJECT PLANNING & SCHEDULING**
   1. Sprint Planning & Estimation



* 1. Sprint Delivery Schedule

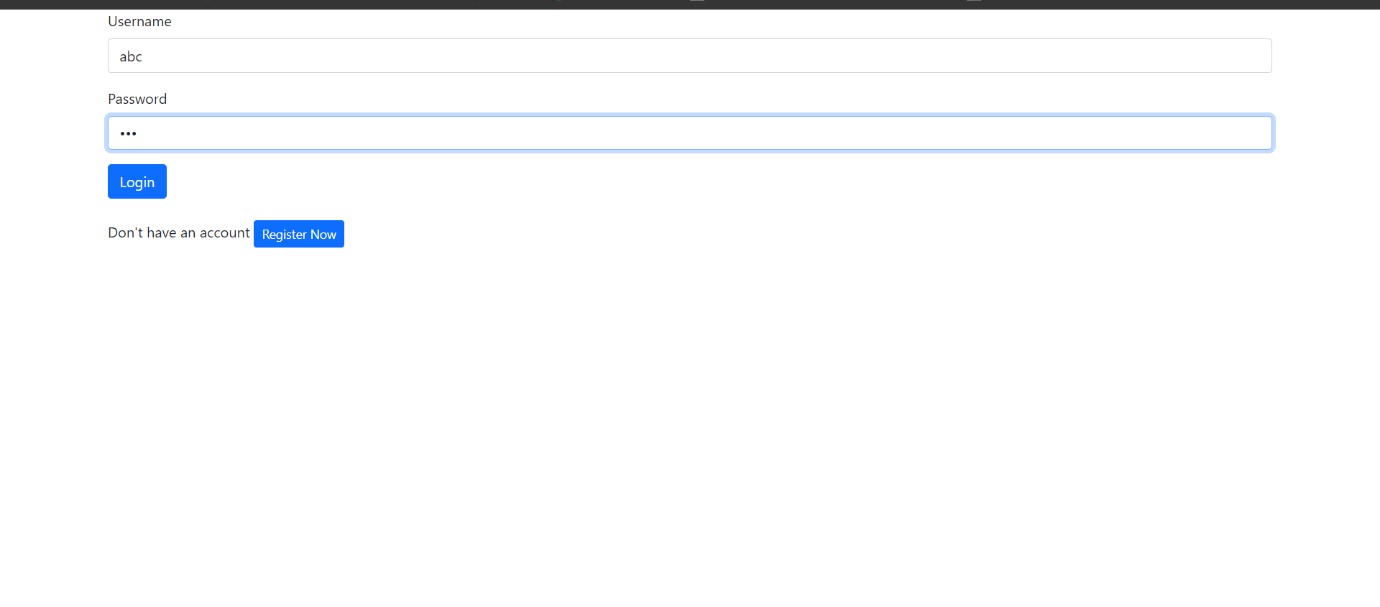


* 1. Reports from JIRA



1. **CODING & SOLUTIONING** Feature 1 - Login

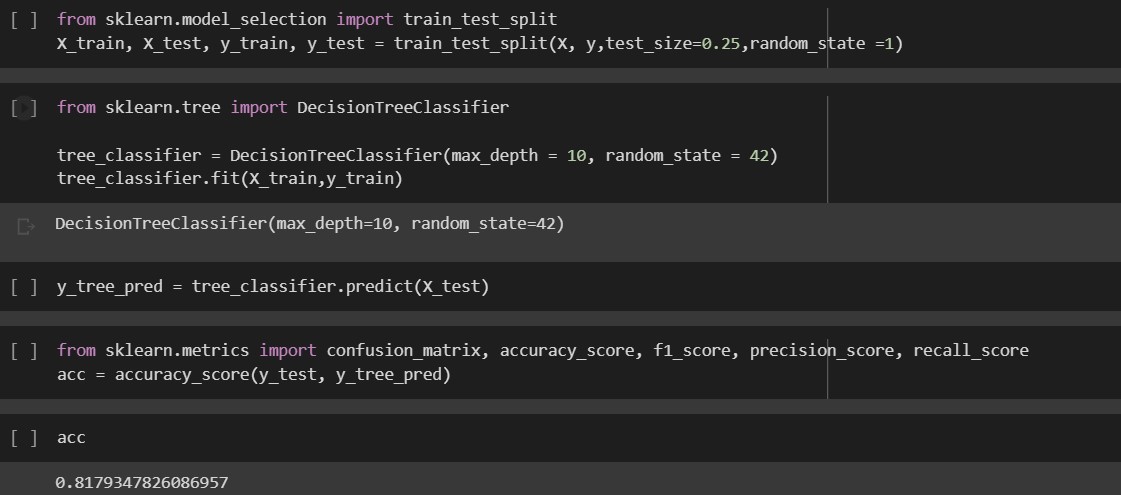
LOGIN



Feature 2 - Accuracy rate (Attriton)

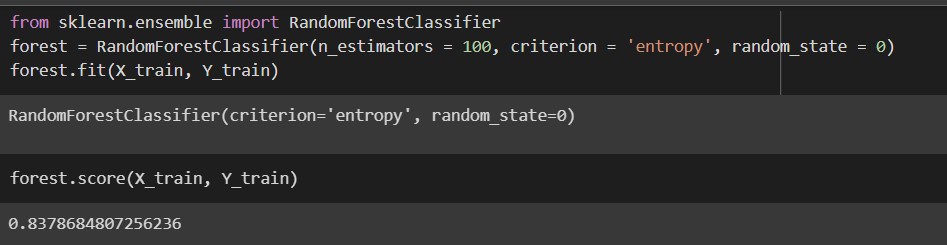
DECISION TREE -

Accuracy :81%



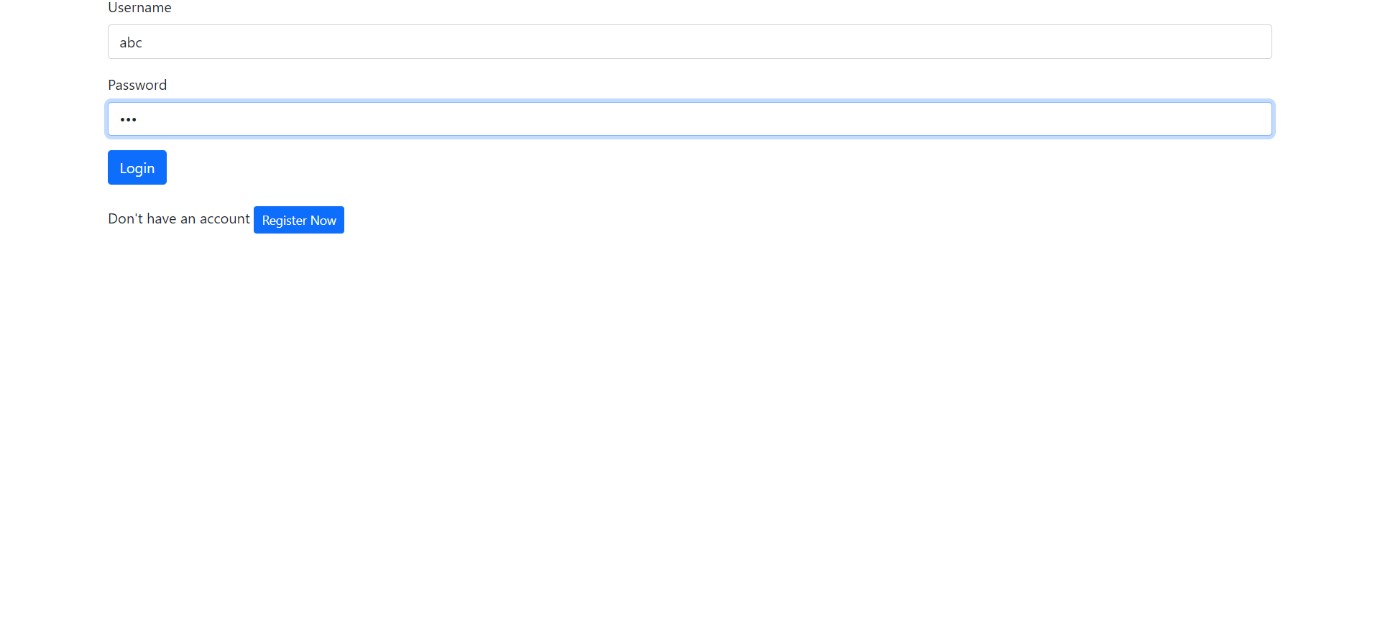
RANDOM FOREST -

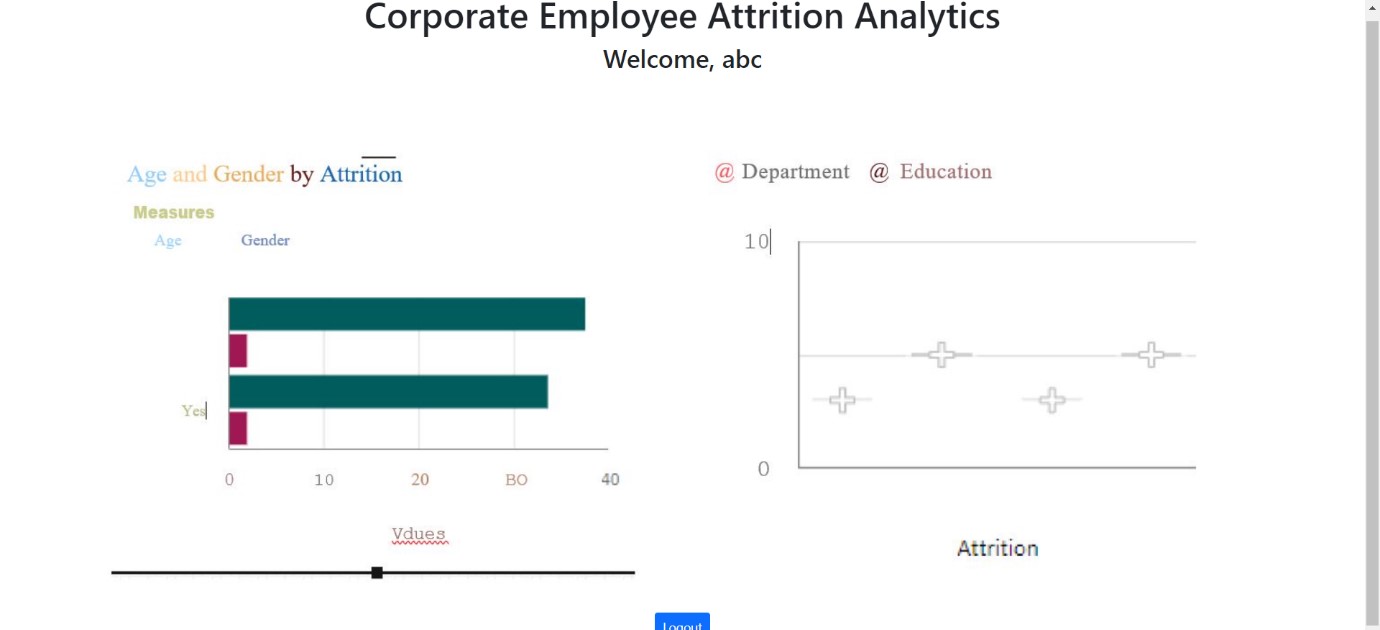
Accuracy :83%



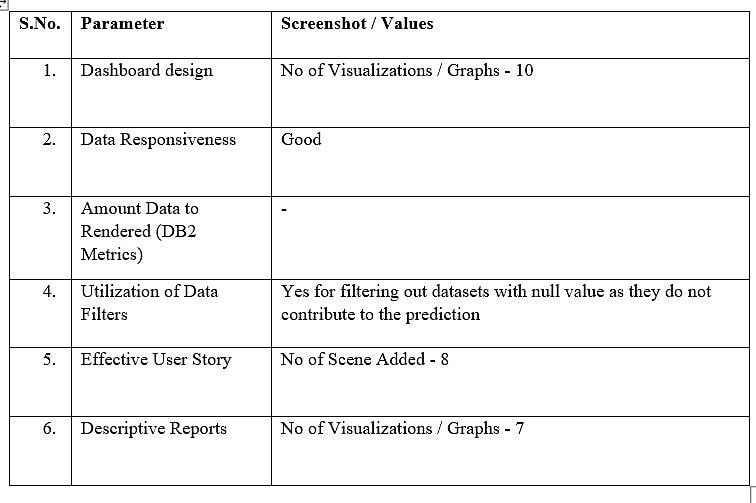
1. **TESTING**
   1. Test Cases

Sucessfully Login will make the website go to next page-Home page





Model performance testing



* 1. User Acceptance Testing

**Purpose of Document**

The purpose of this document is to briefly explain the test coverage and open issues of the project at the time of the release to User Acceptance Testing (UAT).

**Defect Analysis**

This report shows the number of resolved or closed bugs at each severity level, and how they were resolved

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Resolution** | **Severity 1** | **Severity 2** | **Severity 3** | **Severity 4** | **Subtotal** |
| By Design | 10 | 4 | 2 | 3 | 20 |
| Duplicate | 1 | 0 | 3 | 0 | 4 |
| External | 2 | 3 | 0 | 1 | 6 |
| Fixed | 11 | 2 | 4 | 20 | 37 |
| Not Reproduced | 0 | 0 | 1 | 0 | 1 |
| Skipped | 0 | 0 | 1 | 1 | 2 |
| Won't Fix | 0 | 5 | 2 | 1 | 8 |
| Totals | 24 | 14 | 13 | 26 | 77 |

**Test Case Analysis**

This report shows the number of test cases that have passed, failed, and untested

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Section** | **Total Cases** | **Not Tested** | **Fail** | **Pass** |
| Print Engine | 7 | 0 | 0 | 7 |
| Client Application | 51 | 0 | 0 | 51 |
| Security | 2 | 0 | 0 | 2 |
| Outsource Shipping | 3 | 0 | 0 | 3 |

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Exception Reporting | 9 | 0 | 0 | 9 |
| Final Report Output | 4 | 0 | 0 | 4 |
| Version Control | 2 | 0 | 0 | 2 |

# 9. RESULTS

9.1 Performance Metrics

➤ Hours worked: 50 hours

➤ Stick to Timelines: 100%

➤ Stay within budget: 100%

➤ Consistency of the product: 85%

➤ Efficiency of the product: 85%

➤ Quality of the product: 85%

# 10. ADVANTAGES & DISADVANTAGES

Advantages:

Identifying attrition can really help the company in identifying where they are going wrong and correcting it This project has -

Smooth User Interface and Accuracy is achieved quickly

Disadvantages:

Random forest can be used for both classification and regression tasks,but it is no more suitable for Regression tasks.

This analysis is only based on the dataset or data provided,so it has to be perfectly correct.

# 11. CONCLUSION

Research findings suggest that attrition reasons in IT organizations primarily revolve around professional growth and challenges in the organization. Although economic factors happen to the most influential factor, professionals may settle for second best criteria of their preference that is career growth and supportive work policies in the organization.

On the other hand, candidates who aspire to have a better job than the one in hand are more interested in securing the next job. Young talent wants to work on latest technology and functional domain. IT professionals who are young career makers are less influenced by Brand name or geographical area. Most of the IT professionals look for challenging role and position in the organization. Candidates as well as senior professionals believe that challenging work motivate them to maintain the interest in the work life.

This overview of the project conveys the idea that numerous methods have been investigated for detecting the attrition rate. Big data, machine learning, and data mining can be used to great success to analyse the prediction model with the highest degree of accuracy. The primary goal of this project is calculate employee attrition in an orgamization .

# 12. FUTURE SCOPE

A future update shall comprise of section for upload datasets even which have null values and still get the attrition rate. The obtained output can be further processed and sent to smart devices to provide necessary information. Constant monitoring can provide necessary data to recommend to use it incase an emergency.Also it can developed as an app and hosted so that anywhere any organization can use it.

# 13. APPENDIX

Source Code:[https://github.com/IBM-EPBL/IBM-Project-17040-](https://github.com/IBM-EPBL/IBM-Project-17040-1659627084/tree/main/Project%20Development%20Phase/Sprint%201/Login%20Source%20codes)

[1659627084/tree/main/Project%20Development%20Phase/Sprint%201/Login%20Source%20codes](https://github.com/IBM-EPBL/IBM-Project-17040-1659627084/tree/main/Project%20Development%20Phase/Sprint%201/Login%20Source%20codes)

GitHub &Project Demo Link : <https://github.com/IBM-EPBL/IBM-Project-17040-1659627084>