

Corporate Employee Attrition Analysis

1. INTRODUCTION

- 1.1 Project Overview
- 1.2 Purpose

2. LITERATURE SURVEY

- 2.1 Existing problem
- 2.2 References
- 2.3 Problem Statement Definition

3. IDEATION & PROPOSED SOLUTION

- 3.1 Empathy Map Canvas
- 3.2 Ideation & Brainstorming
- 3.3 Proposed Solution
- 3.4 Problem Solution fit

4. REQUIREMENT ANALYSIS

- 4.1 Functional requirement
- 4.2 Non-Functional requirements

5. PROJECT DESIGN

- 5.1 Data Flow Diagrams
- 5.2 Solution & Technical Architecture
- 5.3 User Stories

6. PROJECT PLANNING & SCHEDULING

- 6.1 Sprint Planning & Estimation
- 6.2 Sprint Delivery Schedule
- 6.3 Reports from JIRA

7. CODING & SOLUTIONING

- 7.1 Feature 1
- 7.2 Feature 2
- 7.3 Database Schema

8. TESTING

- 8.1 Test Cases
- 8.2 User Acceptance Testing

9. RESULTS

- 9.1 Performance Metrics

10. ADVANTAGES & DISADVANTAGES

11. CONCLUSION

12. FUTURE SCOPE

13. APPENDIX

- Source Code
- GitHub & Project Demo Link

1. INTRODUCTION

1.1 Project Overview:

Employee attrition is referred to as the regular, but unexpected and uncontrollable, decline in the workforce caused by retirement, illness, or death. The amount of employees who leave a firm without being replaced is measured by the employee attrition rate.

A low attrition rate indicates that open standards have been met quickly, and while a high attrition rate means that more people are departing than are being engaged.

Higher attrition rates can ultimately be caused by excessive employee turnover because it becomes increasingly challenging to fill open positions when there are more of them. This measure is typically used to plan new promotional strategies and to keep tabs on positions available.

Attrition contrasts from attrition in that the employer does not fill a vacancy that had emerged in a position after it has already been filled. The underlying cause of employee attrition are explored in this guide, along with suggestions on how employers might use attrition to their favor.

Employee Rate is that it is defined as the pace at which professionals depart a firm is known as its attrition rate, sometimes known as a "churn rate." It may be computed by dividing the total number of outgoing employees by the average number of employees during a period of time.

Sometimes, a proportion (%) is used to express it. For instance, when a firm has a history of high turnover, recruiters frequently report that it is difficult to find new employees.

1.2 Purpose:

Organizations must frequently explore for ways to increase retention due to the significant costs of turnover. Increasing employee engagement, or the amount of dedication and intention to stay evidenced by an employee given their past, present, and anticipated future achievements at their firm, is one approach to reducing turnover. HRBPs should strive to increase workers' rational commitment, or the degree to which employees think that their managers, teams, or companies are in their consciousness, be it personally, professionally, or otherwise.

This will help to boost retention via engagement. This may be accomplished by determining the aspects of employees' experiences that employees respect and reject, as well as the factors that might make them quit or stay in the future.

Exit interviews are important because you could get frank criticism from people departing the company, and you can leverage their information to direct activities aimed at improving the employee experience. But they fall short due to the fact that they do not acknowledge disengaged workers before they quit. Qualitative methods provide you the chance to hear what existing workers think and shape forthcoming development software, but they may be time-consuming, tough to oversee, and assessment and resources from line staff as well as HR. It is that it can minimize labour costs without taking employee departures into account, attrition is vital for employers to understand. The business has the opportunity to halt hiring when employees depart.

It generally means that when workers begin to retire, the business does not replace them. Should they need to, it permits the company to manage employee expenditures. The company's attrition rate may have been a crucial strategic metric to examine due to its importance significant for an organization.

2. LITERATURE SURVEY

2.1 Existing problem:

Title: Predicting Employee Attrition using Machine Learning

Author: Sarah S. Alduayj ; Kashif Rajpoot

The growing interest in machine learning among business leaders and decision makers demands that researchers explore its use within business organizations. One of the major issues facing business leaders within companies is the loss of talented employees. This research studies employee attrition using machine learning models.

Title: Employee Attrition Using Machine Learning And Depression Analysis

Author: Richard Joseph; Shreyas Udupa; Sanket Jangale; Kunal_Kotkar;

PartheshPawar

Amongst the significant issues that corporate leaders have to deal with within an organization is the decline in proficient employees. This decline is primarily attributed to extreme work pressure, dissatisfaction at work, and ignored mental health issues such as depression, anxiety, etc. A survey containing attrition-related questions helped us gather the required data for analysis.

Title: Prediction of Employee Attrition using Data mining

Author: R. Shiva Shankar; J. Rajanikanth; V.V. Sivaramaraju; K.V.S.S.R. Murthy

This is a issue for the organizations specially when trained, technical and key employees leave for opportunity from the organization. This results in financial loss to replace a trained employee. We use the current and past employee data to analyze the common reasons for employee attrition or attrition.

2.2 References:

Title: Predicting Employee Attrition using Machine Learning

Author: Sarah S. Alduayj ; Kashif Rajpoot

[1] **S. Kaur and R. Vijay**, "Job Satisfaction – A Major Factor Behind Attrition or Retention in Retail Industry," Imperial Journal of Interdisciplinary Research, vol. 2, no. 8, 2016.

[2] **D. G. Gardner, L. V. Dyne and J. L. Pierce**, "The effects of pay level on organization-based self-esteem and performance: a field study," Journal of Occupational and Organizational Psychology, vol. 77, no. 3, pp. 307-322, 2004.

Title: Employee Attrition Using Machine Learning And Depression Analysis

Author: Richard Joseph; Shreyas Udupa; Sanket Jangale; Kunal_Kotkar;
Parthesh Pawar

[1] **K Sunanda (2017)**, An Empirical Study On Employee Attrition In It Industries- With Specific Reference To Wipro Technologies .

[2] **Talapatra, Pradip & Rungta, Saket & Anne, Jagadeesh. (2016)**. Employee Attrition And Strategic Retention Challenges In Indian Manufacturing Industries: A Case Study. VSRD International Journal of Business and Management Research. VI. 251-262.

Title: Prediction of Employee Attrition using Datamining

Author: R. Shiva Shankar; J. Rajanikanth; V.V. Sivaramaraju; K.V.S.S.R. Murthy

[1]. **W. Verbeke, K. Dejaeger, D. Martens, J. Hur, and B. Vaesens**, "New insights into a churn prediction in the telecommunication sector. An profit driven datamining approach," European journal of operational research, vol. 218, no. 1, pp. 211-229, 2012.

[2]. **K.Coussement and D. VandenPoel**, "Integrating the voice of customers through call center emails into a decision support system for attrition prediction," Information & Management, vol. 45, no. 3, pp. 164–174, 2008.

2.3 Problem Statement Definition:

Title: Predicting Employee Attrition using Machine Learning

Author: Sarah S. Alduayj ; Kashif Rajpoot

Employee attrition can be defined as the loss of employees due to any of the following reasons: personal reasons, low job satisfaction, low salary, and a bad business environment. Employee attrition can be categorized into two categories: voluntary and involuntary attrition. Involuntary attrition occurs when employees are terminated by their employer for different reasons, such as low employee performance or business requirements. In voluntary attrition, on the other hand, high-performing employees decide to leave the company of their own volition despite the company's attempt to retain them. Voluntary attrition can result from early retirement or job offers from other firms, for example. Although companies that realise the importance of their employees usually invest in their workforce by providing substantial training and a great working environment, they too suffer from voluntary attrition and the loss of talented employees. Another issue, hiring replacements, imposes high costs on the company, including the cost of interviewing, hiring and training. Predicting employees attrition at a company will help management act faster by enhancing their internal policies and strategies. Where talented employees with a risk of leaving can be offered several propositions, such as a salary increase or proper training, to reduce their likelihood of leaving. Using machine learning models can help companies predict employees attrition.

Title: Employee Attrition Using Machine Learning And Depression Analysis

Author: Richard Joseph; Shreyas Udupa; Sanket Jangale; Kunal_Kotkar;
PartheshPawar

Employee turnover can be described as a constant decline in the workforce due to retirement, death, or resignation. Every organization needs to have a certain percentage of attrition to ensure the growth of the organization. Positive attrition is considered beneficial as it generally results in incapable and less productive employees quitting the organization. Meagre attrition rates result in the stagnation of ideas in the workplace. They do not promote intellectual growth caused by exposure to new fresh recruits' new ideas. High attrition rates prove to be exorbitant for the corporation as the corporation invests time, money, and assets to train employees to make them prepared for the job in a particular corporation. In the case where employees quit the job, it causes considerable losses to the corporation. Companies have an uphill task as they must manage recruiting and training recruits and talent loss due to industry attrition trends. Negative attrition implies a larger, more severe problem inside an organization when high-performing employees quit the company searching for better avenues. The losses incurred when an efficient employee quits are not limited to advanced product beliefs, admirable project administration or links with the customers. This can have a detrimental effect on companies as their productivity decreases considerably, which hampers the organization's morale.

Title: Prediction of Employee Attrition using Data mining

Author: R. Shiva Shankar; J. Rajanikanth; V.V. Sivaramaraju; K.V.S.S.R. Murthy

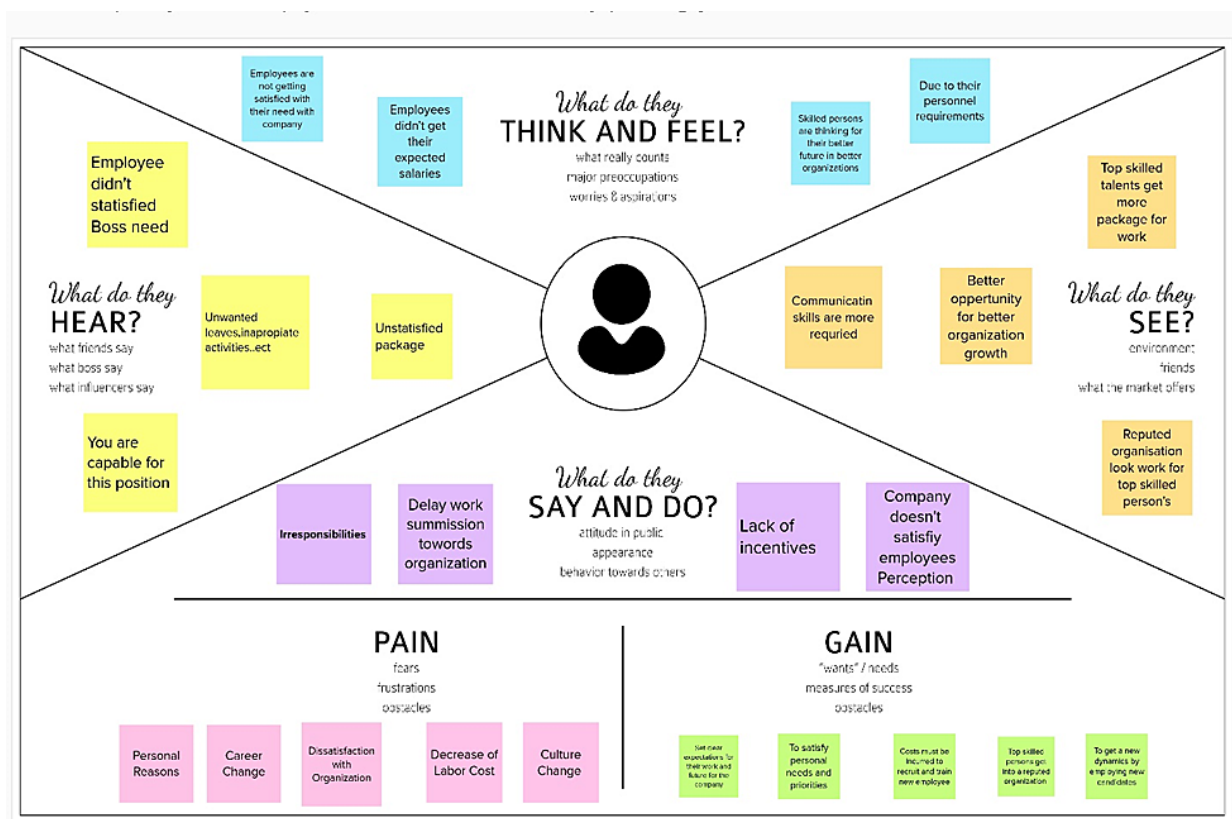
An employee would choose to join or depart an organization depending on many causes i.e. work environment, work place, gender equity, pay equity and many other. The rest of the employees may think about personal reasons for instance relocation due to family, maternity, health, issues with the managers or co-workers in a team. Employee attrition is a major problem for the organizations particularly when trained, technical and key employees leave for best opportunities from the organizations. This finally results into monetary loss to substitute a trained employee. Consequently, we utilize the present and past employee data to assess the familiar issues for employee attrition. The employee attrition identification helps in predicting and resolving the issues of attrition. We can use this data to stop the attrition rate of the employees. For this working we use some methodologies of data classification. For the prevention of employee attrition, we applied a well known classification methods, that is, Decision tree, Logistic Regression, SVM, KNN, Random Forest, Naive bayes methods on the human resource data. For this we implement feature selection method on the data and analysis the results to prevent employee attrition. This is helpful to companies to predict employee attrition, and also helpful to their economic growth by reducing their human resource cost. By improving internal policies and methods, management will be able to respond more quickly when it may anticipate staff attrition. Where talented workers who run the risk of leaving might be offered a number of incentives, such a pay raise or appropriate training, to lessen their likelihood of leaving. Businesses that use machine learning algorithms can forecast staff attrition.

3. IDEATION & PROPOSED SOLUTION

3.1 Empathy Map Canvas:

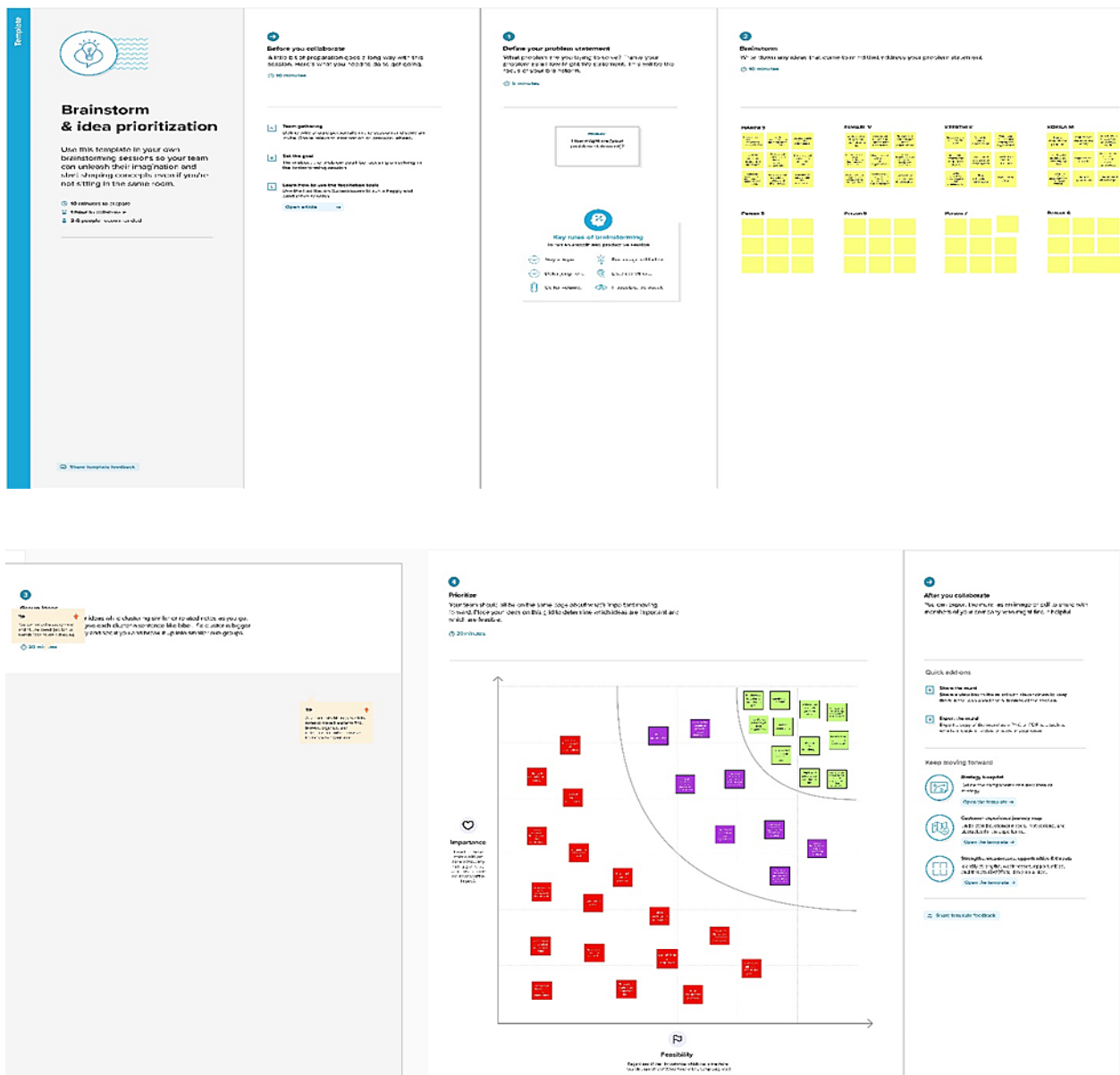
An empathy map is a simple, easy-to-digest visual that captures knowledge about a user's behaviours and attitudes. It is a useful tool to help teams better understand their users. It is a collaborative tool teams can use to gain a deeper insight into their customers. Much like a user persona, an empathy map can represent a group of users, such as a customer segment. The empathy map was originally created by Dave Gray and has gained much popularity within the agile community.

Creating an effective solution requires understanding the true problem and the person who is experiencing it. The exercise of creating the map helps participants consider things from the user's perspective along with his or her goals and challenges.



3.2 Ideation & Brainstorming:

Brainstorming provides a free and open environment that encourages everyone within a team to participate in the creative thinking process that leads to problem solving. Prioritizing volume over value, out-of-the-box ideas are welcome and built upon, and all participants are encouraged to collaborate, helping each other develop a rich amount of creative solutions.



3.3 Proposed Solution:

S.No.	Parameter	Description
1.	Problem Statement (Problem to be solved)	Employee attrition is a significant expense to an organization and in many cases, the Human Resources department's top priority is to fore see such attritions. The goal in solving this problem is to forecast an organization's employee attrition rate.
2.	Idea / Solution description	Employee attrition must be decreased for a firm as it increases the high training cost and the crucial business time of an organization. Focus on employee engagement through meaningful work, goal-setting, and communication of value. Additionally, they should notice the causes for employees to leave the organization. It is more responsible to find the better solution to get into the work and find the various useful points by analyzing the employees who are leaving.
3.	Novelty / Uniqueness	The uniqueness is that it increases the way to predict the employees leaving from the organization due to the dissatisfaction in the work and the various employees leaving from the organization due to their personnel problems can be determined

4.	Social Impact / Customer Satisfaction	This leads to the various dissatisfaction between the employees and the organization. So this can be solved in the organization by the various mechanisms to solve the employees needs from the company to ensure their needs from the organization.
5.	Business Model (Revenue Model)	This model may increase the rate of efficiency to find the employees moving out from the organization for various reasons and they can be resolved. The top skilled persons may move from the companies that can create huge loss to organization by this model we can retain the employees in the organization.
6.	Scalability of the Solution	This ensures the Various Needs that are put forward by the employees in an organization to stay. This makes the solution more powerful to solve the problem with the most suitable format that is needed to be faced. It is most suitable to make the employees stay as far as possible and meet their needs and expectations.

3.4 Problem Solution Fit:

Define CS, fit into CC	1. CUSTOMER SEGMENT(S) CS <small>Who is your customer? i.e. working parents of 0-5 y.o. kids</small>	6. CUSTOMER CONSTRAINTS CC <small>What constraints prevent your customers from taking action or limit their choices of solutions? i.e. spending power, budget, no cash, network connection, available devices.</small>	5. AVAILABLE SOLUTIONS AS <small>Which solutions are available to the customers when they face the problem or need to get the job done? What have they tried in the past? What pros & cons do these solutions have? i.e. pen and paper is an alternative to digital notetaking</small>	Explore AS, differentiate
	<p>Our customer is HR of an organization, who is responsible for an employee to be in their organization. HR need to make the employee meet their satisfaction in their job.</p>	<p>Feeling burnt out and underappreciated</p> <p>Pay and benefit</p> <p>Lack of communication</p>	<p>HR Staff provide information regarding employee benefits, leave of absence, employee assistance programs and worker's compensation benefits.</p>	
Focus on J&P, tap into BE, understand RC	2. JOBS-TO-BE-DONE / PROBLEMS J&P <small>Which jobs-to-be-done (or problems) do you address for your customers? There could be more than one; explore different sides.</small>	9. PROBLEM ROOT CAUSE RC <small>What is the real reason that this problem exists? What is the back story behind the need to do this job? i.e. customers have to do it because of the change in regulations.</small>	7. BEHAVIOUR BE <small>What does your customer do to address the problem and get the job done? i.e. directly related: find the right solar panel installer, calculate usage and benefits; indirectly associated: customers spend free time on volunteering work (i.e. Greenpeace)</small>	Focus on J&P, tap into BE, understand RC
	<p>JOB TO BE DONE:</p> <p>Increasing employee satisfaction</p> <p>Attracting talented professionals</p> <p>Boosting employee retention</p> <p>PROBLEM:</p> <p>Poor work-life balance.</p> <p>Not fitting in and feeling a sense of belonging.</p> <p>Poor job satisfaction and pay.</p> <p>Not enough career opportunities.</p>	<p>Dissatisfaction of Salary.</p> <p>Better job opportunities in the market.</p> <p>Hiring wrong Candidates.</p> <p>Lack of Appreciation.</p> <p>Unrealistic expectations from Employees.</p>	<p>Suggesting an online portals where they will get part time jobs online tutorials and websites for improving their skills</p>	
Define CS, fit into CL	3. TRIGGERS TR <small>What triggers customers to act? i.e. seeing their neighbour installing solar panels, reading about a more efficient solution in the news.</small>	10. YOUR SOLUTION SL <small>What kind of solution suits Customer scenario the best? Adjust your solution to fit Customer behaviour, use Triggers, Channels & Emotions for marketing and communication.</small>	8.1 ONLINE CHANNELS CH <small>What kind of actions do customers take online? Extract online channels from box #7 Behaviour</small>	Explore AS, differentiate
	<p>Other Employees getting better salary than another employee.</p> <p>Some employees getting a better job position than that employee.</p> <p>Top skilled employees getting an apt job than a employee in same skill set level.</p>	<p>The employee attrition can be minimized by various methods like increasing the efficiency of the workers by satisfying their requirements. This could be resolved with the employee attrition by the following constraints improvement such as:</p> <p><i>Pay Competitive Benefits and Perks.</i></p> <p><i>Find The Reason.</i></p> <p><i>Recruit The Right Candidate.</i></p> <p><i>Offer Flexibility.</i></p> <p><i>Provide A Positive Workplace Environment.</i></p> <p><i>Improve Employee Engagement.</i></p> <p><i>Appreciate.</i></p>	<p>Employees do work from home jobs until they get a job for their requirements.</p> <p>They searching for a new job for their need and skill satisfying jobs.</p>	
	4. EMOTIONS: BEFORE / AFTER EM <small>How do customers feel when they face a problem or a job and afterwards? i.e. lost, insecure > confident, in control - use it in your communication strategy & design.</small>		8.2 OFFLINE CHANNELS CH <small>What kind of actions do customers take offline? Extract offline channels from box #7 Behaviour and use them for customer development.</small>	
	<p>BEFORE</p> <p>Lost</p> <p>Insecure</p> <p>Frustrated</p> <p>AFTER</p> <p>Satisfied</p> <p>Comfortable</p> <p>Enthusiastic</p>		<p>Employees do a part time jobs for the money needs.</p> <p>Employees will work on their skill sets for the satisfied job position.</p>	

4. REQUIREMENT ANALYSIS

4.1 Functional requirement:

Functional requirements may involve calculations, technical details, data manipulation and processing, and other specific functionality that define what a system is supposed to accomplish. Behavioral requirements describe all the cases where the system uses the functional requirements, these are captured in use case.

Following are the functional requirements of the proposed solution.

FR No.	Functional Requirement (Epic)	Sub Requirement (Story / Sub-Task)
FR-1	Employee Details	Employee Details through Feedback Forms Employee Details through Registrations Employee Details through Query Forms
FR-2	Employee Confirmation	Confirmation via Email Confirmation via OTP
FR-3	HR Analysis	Analyzing through Predictions Confirmation with IBM Cognos Analytics
FR-4	Confirmation of Analysis	Through Phone Number Through Email

4. 2 Non - Functional requirement:

Non-functional requirements are global constraints on a software system e.g., development costs, operational costs, performance, reliability, maintainability, portability, etc.

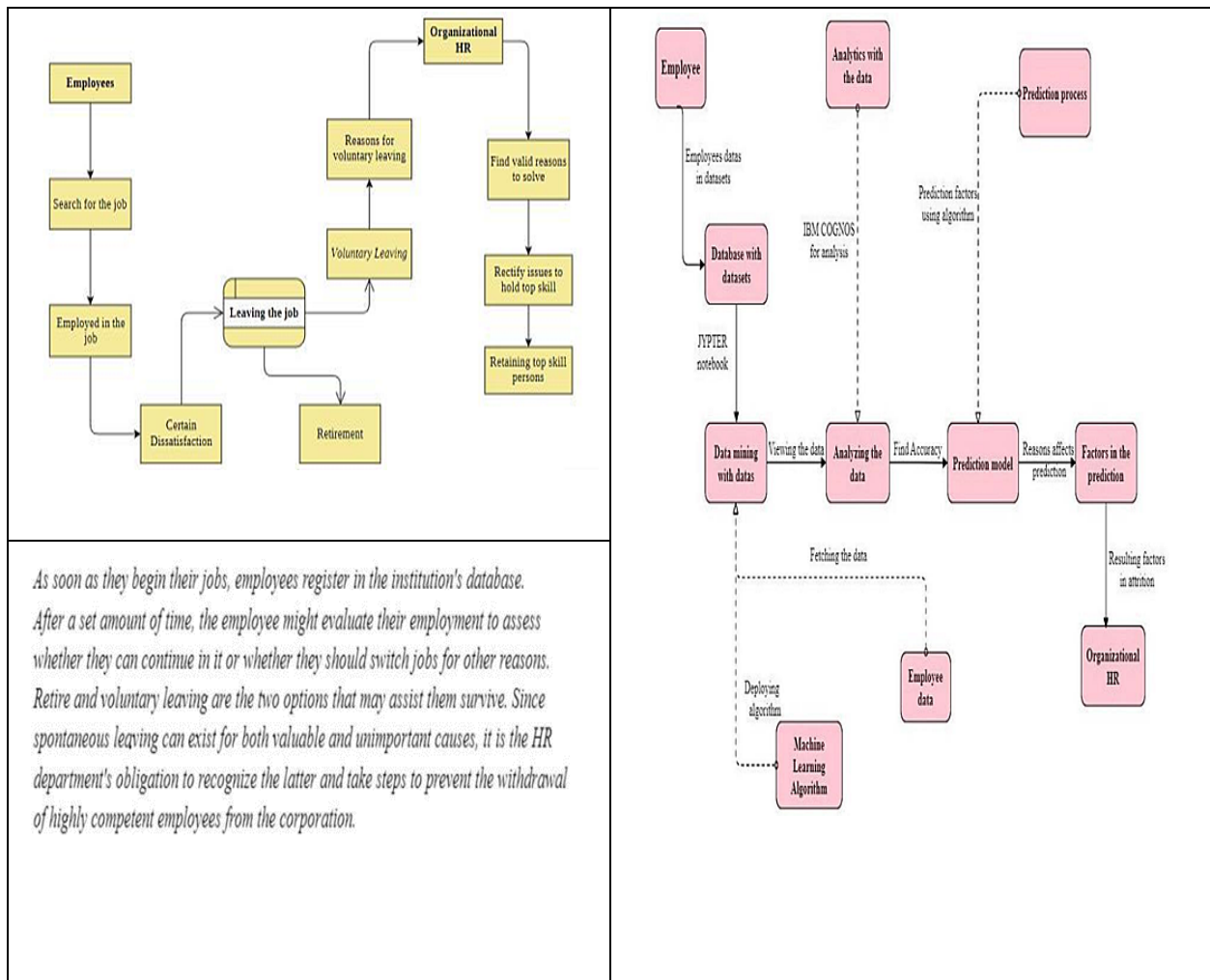
Following are the non-functional requirements of the proposed solution.

FR No.	Non-Functional Requirement	Description
NFR-1	Usability	Employee attrition is a major cost to an organization and predicting such attrition is the most important requirement of the Human Resources department in many organizations
NFR-2	Security	Attrition is a key measurement in the security industry. Because revenue is measured in terms of recurring monthly revenue (RMR), your company's attrition rate – the percentage of accounts you lose over a set period of time – directly affects your revenue.
NFR-3	Reliability	The ERI tool measures how well a candidate is likely to perform in seven critical areas: freedom from disrupted job performance, courtesy, emotional maturity, conscientiousness, trustworthiness, job commitment and safety.
NFR-4	Performance	A high one may mean something in your organization needs to change immediately. Or, if it's growing over time, it may be reflecting a deep, concerning problem with your organization and company culture.
NFR-5	Availability	Availability is time an employee is exclusively available to work for an employer, and cannot be scheduled elsewhere during these times.

5. PROJECT DESIGN

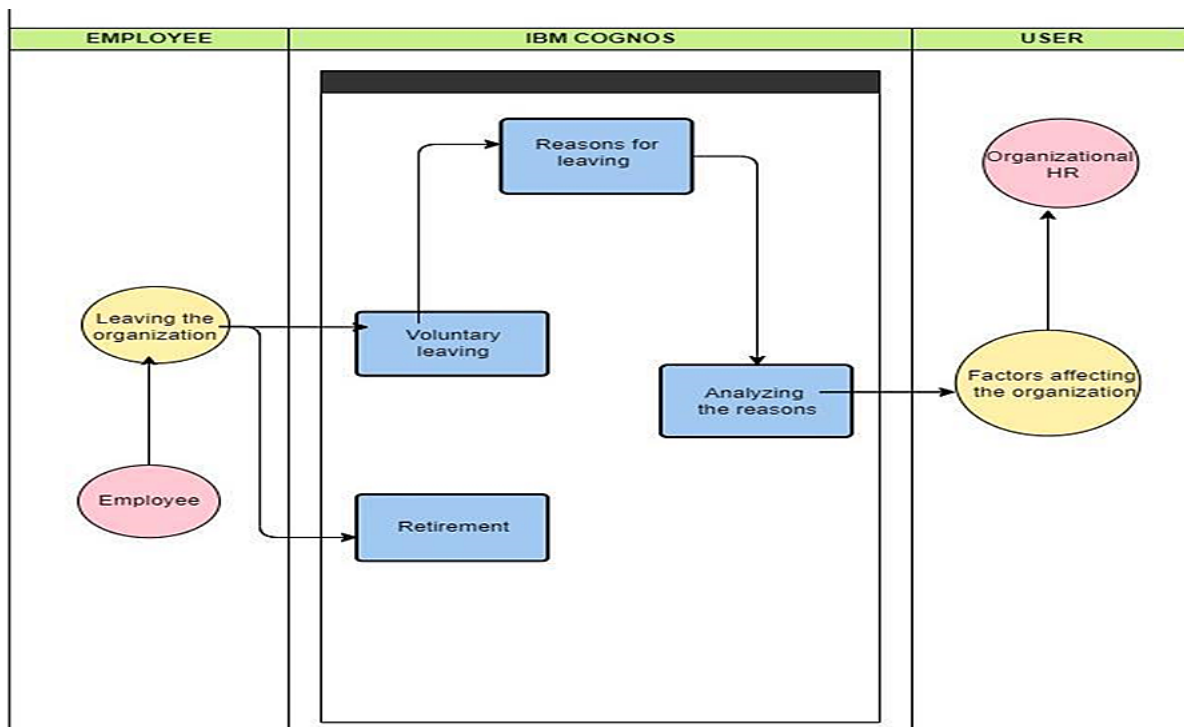
5.1 Data Flow Diagrams:

A Data Flow Diagram (DFD) is a traditional visual representation of the information flows within a system. A neat and clear DFD can depict the right amount of the system requirement graphically. It shows how data enters and leaves the system, what changes the information, and where data is stored.



5.2 Solution & Technical Architecture:

Architecture is defined as: The general layout of a central processing unit and the logical and physical connections between its parts, as applied to computers, software, or networks. The system's hardware, software, access procedures, and protocol use are all documented in the design.



- As shortly after beginning their jobs, staffers register in the institution's database.
- After a specified period of time, the employees leaving from the organization in which uses machine learning model like Decision Tree, Random Forest model, K-Nearest Neighbour training accuracy, etc. to the training the dataset to get the accuracy by predicting the value.
- User interacts with the application using website UI HTML, CSS, JavaScript, React Js etc.
- This logic depends on the extracting the needed contents into the dataset using Python.

Table-1 : Components & Technologies:

S.No	Component	Description	Technology
1.	User Interface	User interacts with the application using website UI, which is used to get the various user needed various user information details from the website UI	HTML, CSS, JavaScript, React Js etc.
2.	Application Logic-1	This logic depends on the extracting the needed contents into the dataset.	Python
3.	Application Logic-2	This logic depends on the training the dataset to get the accuracy by predicting the value	Python Jupiter
4.	Database	Data Type, Configurations etc.	Python Jupiter
5.	Cloud Database	Database Service on Cloud	IBM DB2, IBM Cloudant etc.
6.	File Storage	File storage requirements	IBM Block Storage or Other Storage Service or Local Filesystem
7.	Machine Learning Model	It allows the user to feed a computer algorithm an immense amount of data and have the computer analyse and make data-driven recommendations and decisions based on only the input data.	Decision Tree, Random Forest model, K-Nearest Neighbour training accuracy, etc
8.	Infrastructure (Server / Cloud)	Application Deployment on Local System / Cloud Local Server Configuration: Google server (Collab)	Local, Cloud Foundry, etc.

Table-2: Application Characteristics:

S.No	Characteristics	Description	Technology
1.	Open-Source Frameworks	A software for which the original source code is made freely available and may be redistributed and modified according to the requirement of the user.	Python, Google collab, Python Jupiter
2.	Open-Source Frameworks	IBM Cognos Application Firewall provides security features that are in addition to many of the components identified in the recommended security framework. Firewall architecture is based on a shared library that can be easily updated when new security threats are identified.	Encryptions, IAM Controls, OWASP, SSL Transport Security etc.
3.	Scalable Architecture	IBM Cognos Application Firewall provides security features that are in addition to many of the components identified in the recommended security framework. Firewall architecture is based on a shared library that can be easily updated when new security threats are identified.	Technology used in the architecture is that with the Python and the IBM Cognos.

4.	Availability	Availability is the ability of a system to withstand or recover from exceptional situations, such as a computer failure. The Jupiter Notebook is a web based interactive computing platform. The notebook combines live code, equations, narrative text, visualizations, etc.	Technology used in the architecture is that with the Python and the IBM Cognos.
5.	Performance	This is a fundamental step if we need to achieve the greatest benefit with the least amount of work. Designing for capacity means determining the hardware needed for your system to perform well under its anticipated workload.	Technology used in the architecture is that with the Python and the IBM Cognos.

5.3 User Stories:

The smallest piece of work in an agile system is a user story. It is a final objective, not a feature, as seen through the eyes of a software user. A user story is a casual, all-inclusive description of a software feature written from the viewpoint of the client or end user.

A user story's objective is to describe how a piece of work will provide the customer with a specific value. Keep in mind that "customers" don't always have to be end users on the outside in the conventional sense; they might also be colleagues or internal customers within your company who depend on your team.

User stories are short, straightforward statements that describe the desired result. They don't get specific. Requirements are added after the team has approved

User Type	Functional Requirement (Epic)	User Story Number	User Story / Task	Acceptance criteria	Priority	Release
Employee	Retirement	USN-1	As an Employee, I will leave the organization when the retirement period is occurred.	I can leave the organization for retirement	High	Sprint-1
	Voluntary Attrition	USN-1	As an Employee, I can retire early and become independent consultants due to factors other than age.	I can leave the organization for other factors	High	Sprint-1
		USN-2	As an Employee, I will leave the organization because of unsatisfied salary	I can leave the organization for the better salary	High	Sprint-2
		USN-3	As an Employee, I will leave the organization for the good environment.	I can leave the organization for the better working environment.	Medium	Sprint-1

	Internal Attrition	USN-5	As an Employee, I can leave the current positions to work in other departments within the same company	I can leave the finance department and join the HR department	High	Sprint-1
HR (Representing Organization)	Analysing the Employee Attrition Reasons	USN-1	As an HR, I can analyse the attrition reasons to rectify the Employee Attrition	I will understand and analyse the reasons through analysis of Employee Attrition.	High	Sprint-1
		USN-2	As an HR, I want to analyse what are the requirements for the employees.	I want know the reasons to satisfy the requirements of the employees to rectify attrition.	Medium	Sprint-1
Administrator	Login	USN-1	As an Administrator, I can login into the analysis page.	I can login to the analysis page using username and password.	High	Sprint-2
	Dashboard	USN-2	As an Administrator, I can access the Dashboard.	I can update the details of the Employee Attrition Reasons	High	Sprint-2

6. PROJECT PLANNING & SCHEDULING

6.1 Sprint Planning & Estimation:

Sprint	Functional Requirement (Epic)	User Story Number	User Story / Task	Story Points	Priority	Team Members
Sprint-1	Datasets	USN-1	As a user, I can enter the details of the employees working in our organization for the attrition detail.	2	High	Harini. S
Sprint-1		USN-2	As an Analyst, I will check the dataset and clean the dataset to create an efficient model.	3	High	Kokila. M
Sprint-2	Exploring data and creating model	USN-3	As an Analyst, I can make Exploratory Data Analysis to analyze the important factors for the attrition.	2	Low	Keerthi.K
Sprint-2		USN-4	As an Analyst, I will create a prediction model for predicting the attrition.	3	Medium	Eswari. V
Sprint-3	Prediction	USN-5	As an Analyst, I will create different type of model to identify which give the correct prediction.	5	High	Harini. S

Sprint-4	Creation of webpage	USN-6	As an Analyst. I will create a HTML report.	2	Medium	Kokila. M
Sprint-4		USN-7	As an Analyst, I will create the story in IBM cognos to analyze the model.	3	High	Eswari. V

6.2 Sprint Delivery Schedule:

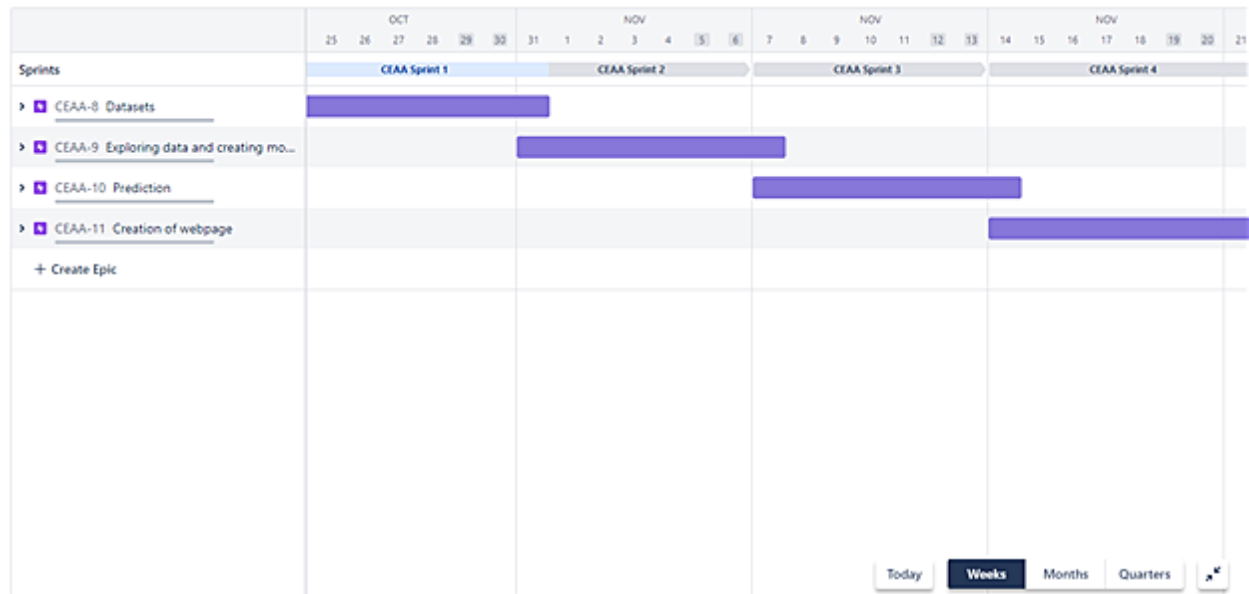
`Sprint	Total Story Points	Duration	Sprint Start Date	Sprint End Date (Planned)	Story Points Completed (as on Planned End Date)	Sprint Release Date (Actual)
Sprint-1	5	7 Days	22 Oct 2022	29 Oct 2022	5	29 Oct 2022
Sprint-2	5	7 Days	29 Oct 2022	05 Nov 2022		
Sprint-3	5	7 Days	05 Nov 2022	12 Nov 2022		
Sprint-4	5	7 Days	12 Nov 2022	19 Nov 2022		

Velocity:

We have a 8-day sprint duration, and the velocity of the team is 5 (points per sprint). To calculate the team's average velocity (AV) per iteration unit (story points per day).

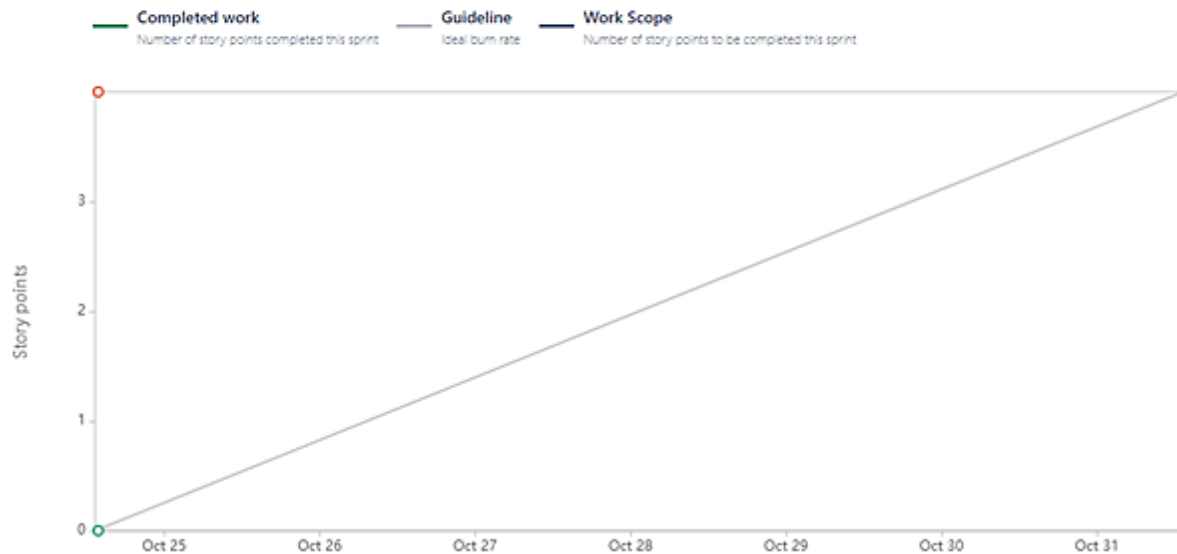
$$AV = \frac{\text{Sprint Duration}}{\text{Velocity}} = \frac{7}{5} = 1.4$$

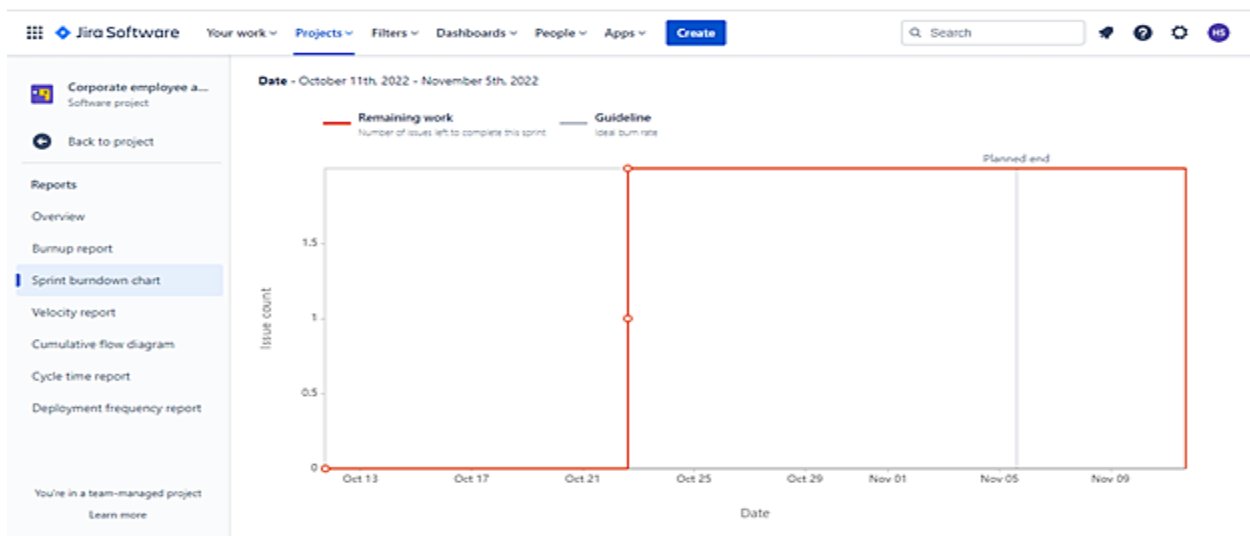
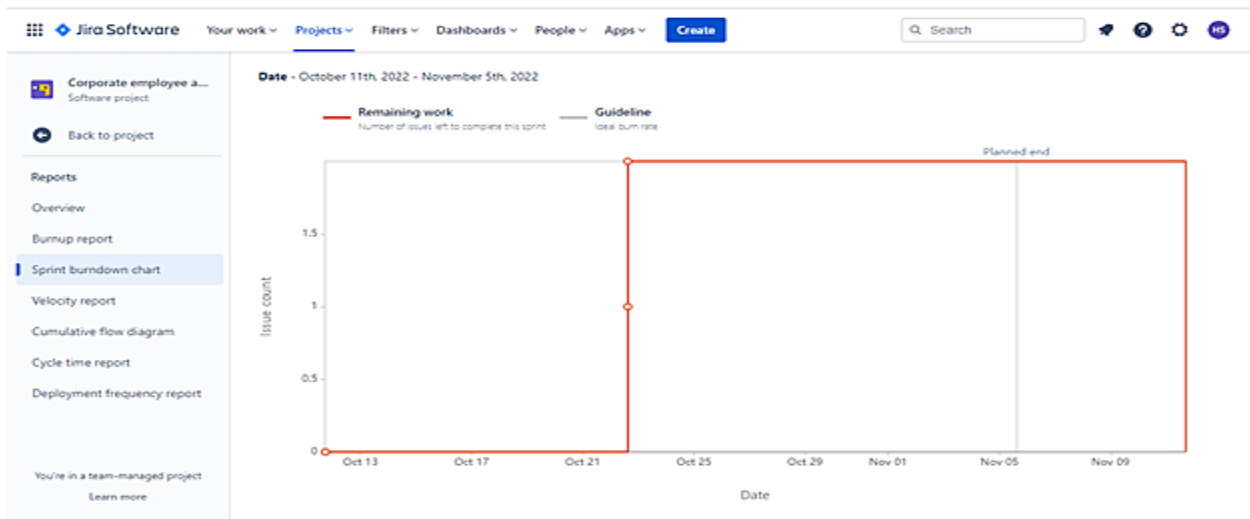
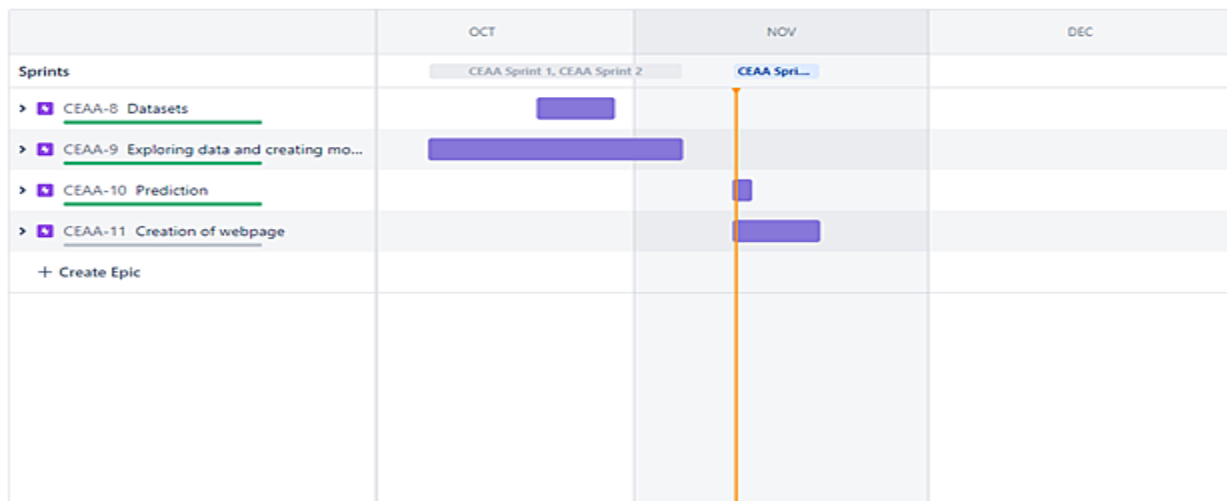
6.3 Reports from JIRA:

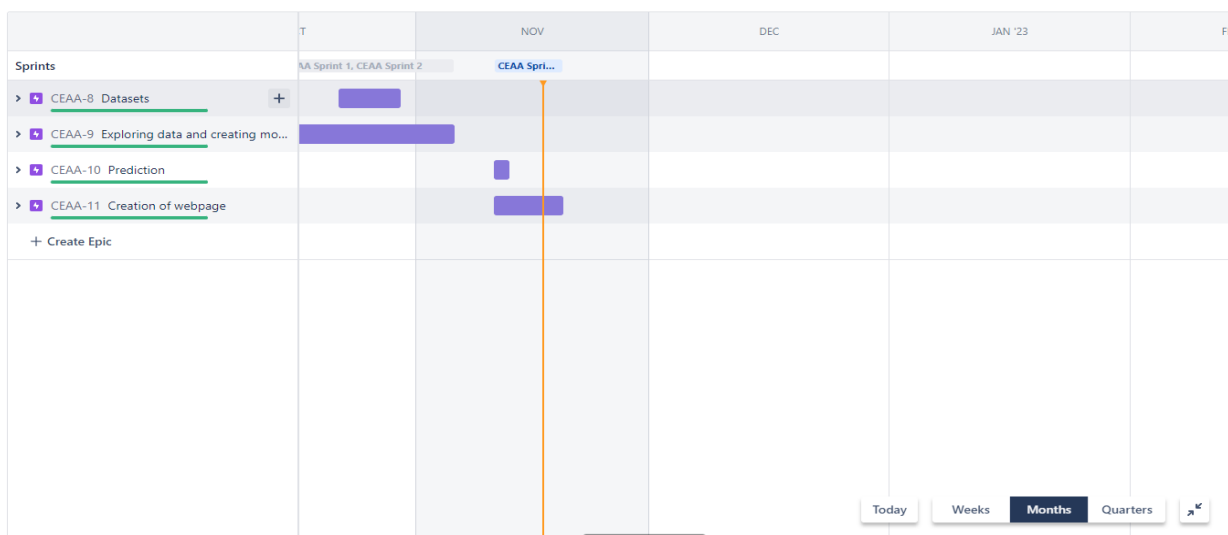
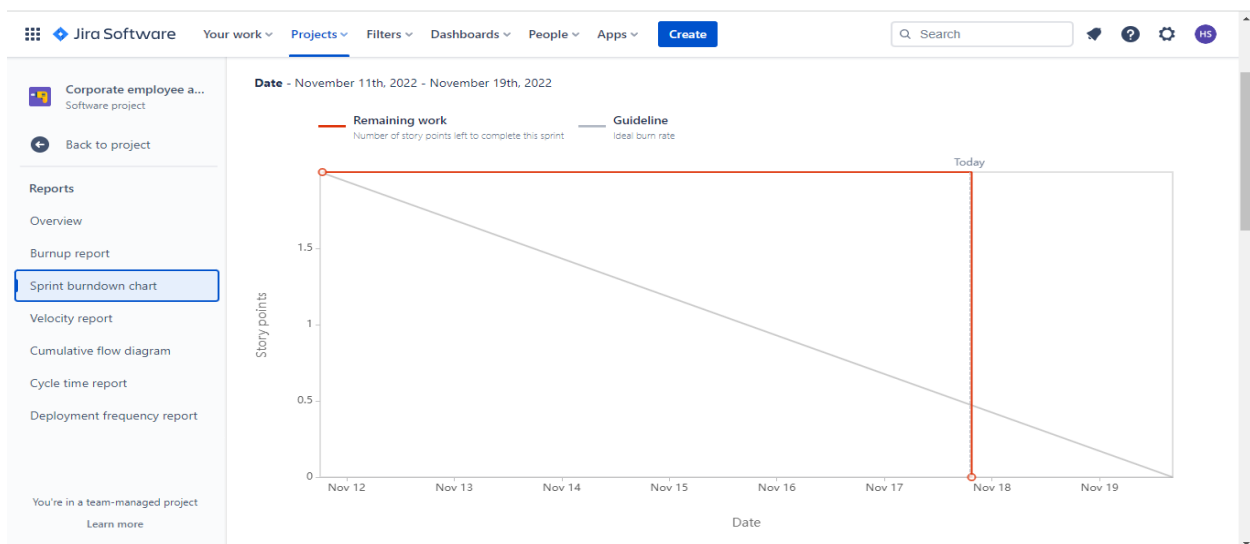
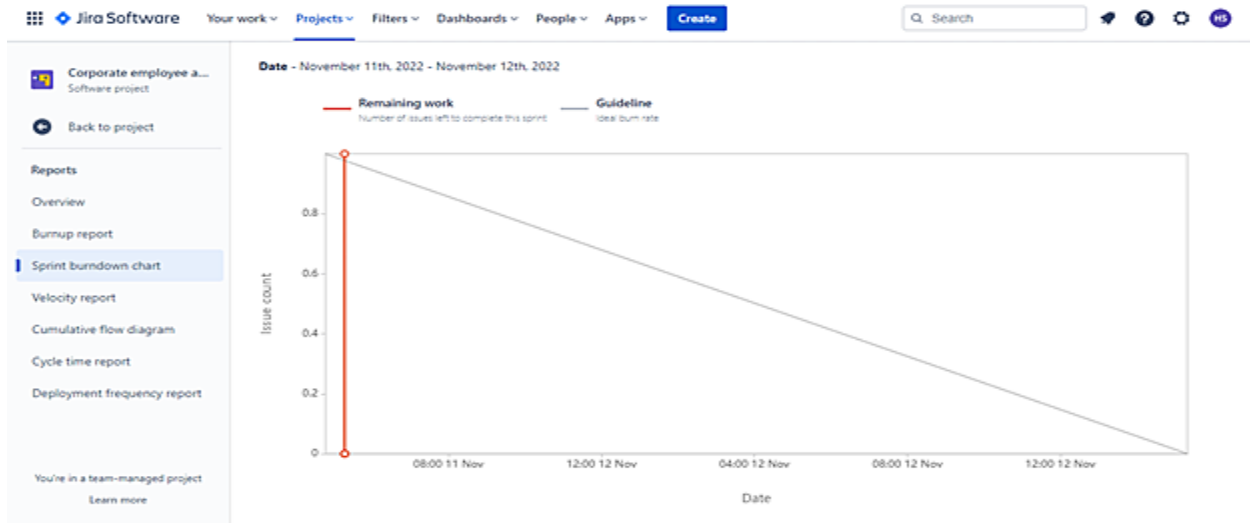


Sprint: CEEA Sprint 1
Estimation field: Story points

Date - October 24th, 2022 - October 31st, 2022







7. CODING & SOLUTIONING

7.1 Feature 1:

general_data prediction.ipynb:

1. To import the libraries that are required for the prediction and various analysis:

```
import numpy as np  
  
import pandas as pd  
  
import matplotlib.pyplot as plt  
  
%matplotlib inline  
  
import seaborn as sns
```

2. To import the directory to import the csv file

```
import os  
  
os.chdir("E:/IBM/IBM/Datasets")  
  
df = pd.read_csv('general_data.csv')
```

3. To describe the dataset that is defined as “df”

- To display the df
df
- To display the columns that are present in the “df”
df.columns
- To display the shape of the dataset
df.shape
- To display the information of the dataset
df.info()

- To display the information of the dataset

```
df.describe()
```

- To display the null values present in the dataset

```
df.isnull().sum()
```

- To remove the null values present in the columns

```
df['NumCompaniesWorked']=df['NumCompaniesWorked'].fillna(df['NumCompaniesWorked'].mean())
```

```
df['TotalWorkingYears']=df['TotalWorkingYears'].fillna(df['TotalWorkingYears'].mean())
```

- To display the value count present in the column Attrition

```
df['Attrition'].value_counts()
```

4. To display the visual representation of the column.

- To display the representation of the column in countplot.

```
sns.countplot(df['Attrition'])
```

```
fig_dims = (12, 4)
```

```
fig, ax = plt.subplots(figsize=fig_dims)
```

```
sns.countplot(x='Age', hue='Attrition', data = df, palette="colorblind", ax = ax,
```

```
edgecolor=sns.color_palette("dark", n_colors = 1));
```

- To display the heatmap representation of the column.

```
plt.figure(figsize=(14,14)) #14in by 14in
```

```
sns.heatmap(df.corr(), annot=True, fmt='.0%')
```

6. To describe the testing and training model:

- To drop the unwanted columns in the dataset.

```
df = df.drop('EmployeeCount', axis = 1)
```

```
df = df.drop('Over18', axis = 1)
```

- To split the dataset

```
X = df.iloc[:, 0:1].values
```

```
y = df.iloc[:, 1]
```

- To describe the shape of the X and y of the splitted dataset

```
X_train.shape
```

```
y_train.shape
```

- To train and the split the dataset

```
from sklearn.model_selection import train_test_split
```

```
X_train, X_test, Y_train, Y_test = train_test_split(X, y, test_size = 0.4, random_state = 0)
```

6. The models to predict the dataset using various models:

- Using Random Forest:

```
from sklearn.ensemble import RandomForestClassifier
```

```
forest = RandomForestClassifier(n_estimators = 100, criterion = 'entropy',  
random_state = 0)
```

```
forest.fit(X_train, Y_train)
```

- Predicting the forest model:

```
forest.score(X_train, Y_train)
```

- Using Decision Tree:

```
from sklearn.tree import DecisionTreeClassifier
tree_classifier = DecisionTreeClassifier(max_depth = 10, random_state = 42)
tree_classifier.fit(X_train,y_train)
```

- Using the prediction in Decision Tree:

```
y_tree_pred = tree_classifier.predict(X_test)
```

- Prediction using Decision Tree

```
from sklearn.metrics import confusion_matrix, accuracy_score, f1_score,
precision_score, recall_score
acc = accuracy_score(y_test, y_tree_pred)
```

- Calculating the accuracy:

```
acc
```

7.2 Feature 2 (Web page):

- Index.html: (<https://employee-attrition-analysis.herokuapp.com/>)

This code is the base which displays as the home page. The home page is the first Web page that is displayed after starting a Web browser like Google Chrome or Microsoft's Internet Explorer. The browser is usually preset so that the home page is the first page of the browser manufacturer.

```
<!DOCTYPE html>
```

```
<html lang="en">
```

```
<head>
```

```
<meta charset="UTF-8">
```

```
<meta http-equiv="X-UA-Compatible" content="IE=edge">
```

```
<meta name="viewport" content="width=device-width, initial-scale=1.0">
```

```
<link href="https://cdn.jsdelivr.net/npm/bootstrap@5.2.1/dist/css/bootstrap.min.css"
```

```

rel="stylesheet">

<title>Corporate Employee Attrition Analysis</title>

<link rel="stylesheet" href="index.css">

</head><body>

<div class="head-wrapper">

    <h2><b><center>Corporate Employee Attrition Analysis</center></b></h2></div>

<div>

    <p id="head-text">Employee attrition is referred to as the regular, but unexpected and uncontrollable, decline in the workforce caused by retirement, illness, or death. Attrition contrasts from attrition in that the employer does not fill a vacancy that had emerged in a position after it has already been filled. The underlying cause of employee attrition are explored in this guide, along with suggestions on how employers might use attrition to their favor.</p>

</div>

<div class="content"><table width="100%" height="90%">

    <tr id="container"> <td>

        <center></center>

        <a href="dashboard.html"><center class="text">Dashboard</center></a>

    </td>

    <td>

        <center></center>

        <a href="report.html"><center class="text">Report</center></a>

    </td>

    <td>

        <center></center>

        <a href="story.html"><center class="text">Story</center></a>

    </td></tr>

</table> </div>

</body></html>

```


- **Dashboard.html:**

Here Dashboard is displayed through IBM Cognos Analytics. It integrates reporting, modeling, analysis, exploration, dashboards, stories, and event management so you can understand your organization's data, and make effective business decisions.

```
<!DOCTYPE html>
```

```
<html lang="en">
```

```
<head>
```

```
  <meta charset="UTF-8">
```

```
  <meta http-equiv="X-UA-Compatible" content="IE=edge">
```

```
  <meta name="viewport" content="width=device-width, initial-scale=1.0">
```

```
  <link href="https://cdn.jsdelivr.net/npm/bootstrap@5.2.1/dist/css/bootstrap.min.css"
rel="stylesheet">
```

```
  <title>Dashboard</title>
```

```
  <link rel="stylesheet" href="index.css">
```

```
</head>
```

```
<body>
```

```
  <div class="head">
```

```
    <center><b><h2>Dashboard</h2></b></center>
```

```
  </div>
```

```
  <div>
```

```
    <table width="100%" height="100%">
```

```
      <tr>
```

<h3 class="content-head">Employee Attrition Dashboard for general data with cognos tool </h3>

<center><p><iframe
src="https://us3.ca.analytics.ibm.com/bi/?perspective=dashboard&pathRef=.public_folders%2F
Final%2FGeneral%2Bdata%2Bdashboard&action=view&mode=dashboard&subView=model00
00018471546a1b_000000000" width="1200" height="800" frameborder="0" gesture="media"
allow="encrypted-media" allowfullscreen=""></iframe></p></center>

</tr>

<tr>

<h3 class="content-head">Employee Attrition Dashboard for employee and manager data with cognos tool </h3>

<center><p><iframe
src="https://us3.ca.analytics.ibm.com/bi/?perspective=dashboard&pathRef=.public_folders%2F
Final%2FEmployee%2B-
%2BManager%2Bdashboard&action=view&mode=dashboard&subView=model0000018470dc3
186_000000000" width="1200" height="800" frameborder="0" gesture="media"
allow="encrypted-media" allowfullscreen=""></iframe></p></center>

</tr>

</table>

</div>

</body>

</html>

- **Report.html:**

Here report is displayed through IBM Cognos Analytics. IBM Cognos Analytics Reporting tool is a web-based reporting tool that integrates Processing, Analytics, Stories, Events, Modules, Content.

<!DOCTYPE html>

<html lang="en">

<head>

<meta charset="UTF-8">

<meta http-equiv="X-UA-Compatible" content="IE=edge">

<meta name="viewport" content="width=device-width, initial-scale=1.0">

<link href="https://cdn.jsdelivr.net/npm/bootstrap@5.2.1/dist/css/bootstrap.min.css" rel="stylesheet">

<title>Report</title>

<link rel="stylesheet" href="index.css">

</head>

<body>

<div class="head">

<center><h2>Report</h2></center>

</div>

<div>

<table width="100%" height="100%">

<tr>

<h3 class="content-head">Employee Attrition Report for general data with cognos tool </h3>

<center><p><iframe
src="https://us3.ca.analytics.ibm.com/bi/?pathRef=.my_folders%2FFinal%2FReport%2Bview%2Bof%2BGeneral%2Bdata%2Breport&action=run&format=HTML&prompt=false"

width="1200" height="800" frameborder="0" gesture="media" allow="encrypted-media" allowfullscreen=""></iframe></p></center>

</tr>

<tr>

<h3 class="content-head">Employee Attrition Report for employee and manager data with cognos tool </h3>

<center><p><iframe
src="https://us3.ca.analytics.ibm.com/bi/?pathRef=.my_folders%2FFinal%2FReport%2Bview%2Bof%2BEmployee%2B-%2BManager%2Breport&action=run&format=HTML&prompt=false"
width="1200" height="800" frameborder="0" gesture="media" allow="encrypted-media" allowfullscreen=""></iframe></p></center>

</tr>

</table>

</div>

</body>

</html>

- **Story.html:**

Here story is displayed through IBM Cognos Analytics. It provides dashboards and stories to communicate your insights and analysis.

<!DOCTYPE html>

<html lang="en">

<head>

<meta charset="UTF-8">

<meta http-equiv="X-UA-Compatible" content="IE=edge">

<meta name="viewport" content="width=device-width, initial-scale=1.0">

<link href="https://cdn.jsdelivr.net/npm/bootstrap@5.2.1/dist/css/bootstrap.min.css"
rel="stylesheet">

<title>Story</title>

<link rel="stylesheet" href="index.css">

</head>

<body>

<div class="head">

<center><h2>Story</h2></center>

</div>

<div>

<table width="100%" height="100%">

<tr>

<h3 class="content-head">Employee Attrition Story for general data with cognos tool
</h3>

<center><p><iframe
src="https://us3.ca.analytics.ibm.com/bi/?perspective=story&pathRef=.public_folders%2FFinal
%2Fgeneral%2Bdata%2Bstory&action=view&sceneId=model000001847184603a_000000000&s
ceneTime=0" width="1200" height="800" frameborder="0" gesture="media" allow="encrypted-
media" allowfullscreen=""></iframe></p></center>

</tr>

<td>

<h3 class="content-head">Employee Attrition Story for employee and manager data
with cognos tool </h3>

```
<center><iframe
src="https://us3.ca.analytics.ibm.com/bi/?perspective=story&pathRef=.public_folders%2FFinal
%2FEmployee%2B-
%2BManager%2Bstory&action=view&sceneId=model00000184712564c0_000000000&sceneTi
me=13950" width="1200" height="800" frameborder="0" gesture="media" allow="encrypted-
media" allowfullscreen=""></iframe></p></center>
```

```
</tr>
```

```
</table>
```

```
</div>
```

```
</body>
```

```
</html>
```

- **Index.css:**

CSS (Cascading Style Sheets) is used to style and layout web pages — for example, to alter the font, color, size, and spacing of your content, split it into multiple columns, or add animations and other decorative features.

```
body {
    margin: 0;
    padding: 0;
    background-color: lightblue;}
```

```
#D-img
```

```
{
    margin: 0;
```

```
width: 150px;
```

```
height: 150px;
```

```
border-radius: 1rem;}
```

```
.text
```

```
{
```

```
color: black;
```

```
font-size: 20px;
```

```
font-weight: 700;
```

```
font-family: Arial, Helvetica, sans-serif;}
```

```
.head-wrapper
```

```
{
```

```
margin: 0;
```

```
padding: 3rem 4rem;
```

```
font-family: Georgia, 'Times New Roman', Times, serif;
```

```
font-weight: 900;}
```

```
.head
```

```
{
```

```
margin: 0;
```

```
padding: 1rem 4rem;}
```

```
#head-text
```

```
{
```

```
margin: 0;
```

```
padding: 20px;

font-size: 18px;

font-weight: 400;

text-align: justify;

font-family: Arial, Helvetica, sans-serif;

line-height: 45px;}
```

```
.content
```

```
{

margin: 0;

padding: 5rem;}
```

```
.content-head {
```

```
margin: 0;

padding: 3rem;

text-align: center;

font-size: 20px;}
```

- **Index.php:**

PHP(short for Hypertext PreProcessor) is the most widely used open source and general purpose server side scripting language used mainly in web development to create dynamic websites and applications.

```
<?php include_once("index.html"); ?>
```


7.3 Database Schema

Kaggle External API connectivity:

- Importing library, creating library & coping json file:

```
!pip install -q kaggle
```

```
!mkdir ~/.kaggle #creating the kaggle library
```

```
!cp kaggle.json ~/.kaggle/ #copying the json file to folder
```

```
!chmod 600 ~/.kaggle/kaggle.json
```

- For downloading the dataset from the Kaggle account

```
!kaggle datasets download -d harinis2813/employeeattritiondataset
```

- To unzip the files from the zip file uploaded in the Kaggle

```
!unzip /content/employeeattritiondataset.zip
```

```
!pip install -q kaggle
```

```
[ ] !mkdir ~/.kaggle #creating the kaggle library
```

```
mkdir: cannot create directory '/root/.kaggle': File exists
```

```
[ ] !cp kaggle.json ~/.kaggle/ #copying the json file to folder
```

```
[ ] !chmod 600 ~/.kaggle/kaggle.json
```

```
[ ] !kaggle datasets download -d harinis2813/employeeattritiondataset
```

```
employeeattritiondataset.zip: Skipping, found more recently modified local copy (use --force to force download)
```

```
[ ] !unzip /content/employeeattritiondataset.zip
```

```
Archive: /content/employeeattritiondataset.zip
inflating: data_dictionary.xlsx
inflating: employee_survey_data.csv
inflating: general_data.csv
inflating: in_time.csv
inflating: manager_survey_data.csv
inflating: out_time.csv
```

8. TESTING

8.1 Test Cases:

1. To view dashboard, story and report icons
2. Click on dashboard icon
3. Click on report icons
4. Click on story icon

8.2 User Acceptance Testing:

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	19
External	11	2	4	20	37
Fixed	2	3	0	1	6
Skipped	0	0	1	1	2
Won't Fix	0	5	2	1	8
Totals	23	14	9	26	72

Test case Analysis:

Section	Total Cases	Not Tested	Fail	Pass
Print Engine	12	0	0	12
Web app view	3	0	0	3
Security	12	0	0	12
Post Redirect	3	0	1	2
Post	3	0	0	3

9. RESULTS

9.1 Performance Metrics:

Dashboard designs:

By presenting critical insights and analyses about your data on one or more pages or screens, a dashboard enables you to keep track of events or actions at a glance.

Number of Visualizations / Graphs – There are two (2) dashboard tabs with 7 - 15 visualizations in each dashboard which explains the various metrics of design of dashboard.

Data Responsiveness:

It hides certain aspects of the visualization if the size is limited, to maximize the space that is available to display data.

- There was two different datasets with the common column and full outer join was done by that common column.
- There was another dataset with various continuous values , those values was grouped as common.

Effective User Story:

An example of a view is a story. A story is made up of a series of scenes that are told in chronological order. You may use stories to give your data a visual narrative.

Number of Scene Added in the story – There are two effective stories with 8 – 10 visualizations in each story explaining the metrics of impact in Employee Attrition in various measures.

Descriptive Reports:

The firm will be transparent thanks to the reporting tool. It supports industry-leading databases and offers a variety of data sources, enable to create efficient content management and reporting methods.

Number of Visualizations / Graphs – There are two reports with 3 - 5 visualization in each report to display the similar or kinds of data.

10. ADVANTAGES & DISADVANTAGES

Advantages:

Due to the high expense of hiring and training new employees, employers strive to minimise turnover. However, an employer could gain when an employee willingly leaves a company:

- **Reduction in labour costs:**

To avoid layoffs, employers may decide to institute a hiring freeze. Layoffs resulting from a shift or decline in the business lower employee morale and make it challenging to fill new roles. When workers voluntarily leave, the business can cut costs while making other hiring decisions..

- **Resource shifting:**

By opting not to fill a job, companies might adjust departmental work flows, give new responsibilities to existing workers, or reallocate resources within an organisation.

- **New dynamics:**

Attrition presents chances for new dynamics and ideas. It may revitalise a company and present current personnel with fresh chances.

- **Culture change:**

Changing an organization's culture is not always simple, especially if you don't replace the entire workforce. Without dismissing or laying off workers, businesses may renew their corporate culture through natural staff attrition.

Disadvantages:

When employees leave the organization it is a loss to the company, the team and the individuals. Employees are the backbone of any organization and their departing may lead to lot of various losses to company on different aspects. The disadvantages can be:

- **Decreased overall performance:**

The whole business process is affected when an employee leave the organization. It is even more risky when this happen all of a sudden. There is no time to train the new employee who is to take over the job and the whole team gets affected.

- **Daily task management:**

Sudden attrition may lead to difficulty in managing daily tasks. Even large organization struggle to manage their task when employees leave jobs, getting small information and managing daily tasks become difficult as they cannot be managed by small current team which is left behind.

- **Increased cost:**

This has to be the highest disadvantage to a company when employees leave their jobs. There is increased cost associated with every level of the process – **losing and paying the previous employee, hiring a new one, training cost for the new employee.**

- **Lack of knowledgeable employees:**

This goes without saying when employees leave an organization they take with them the experience they have gained overtime. With organizations which has high attrition rate the average years of experience of employees is really low. This result in low performance, lack of loyalty and cluelessness on what company has been through.

- **Create a Negative image:**

It is not just that employees are looking for job, even organizations are on thezqualified professionals. When any company has high attrition rate it negatively impact the brand of the organization.

11. CONCLUSION

In this research, a machine learning model for forecasting staff attrition is presented. The feature space's dimension was initially reduced using a feature selection approach.

In employee attrition problem, an estimation can be framed for either the employee will leave the company or not. With this analysis, the organization can choose the employees with the utmost chances of leaving the organization and then assign them confined incentives. There could also be some cases of false positives where human resource thinks that employee will leave the company in a short span of time, but actually, the employee does not. These mistakes could be identified to the HR by the analysis of various IBM Cognos Analytics tools like dashboard, story and reports are visualized.

12. FUTURE WORK

As we analysed and visualised the information given by the dataset we created as an web app. We created an visualization using IBM Cognos Analytics tool. We created an visualization to represent the analysis as Dashboard, Report and Story. We represented our analysis as an externally integrated web app.

In future it can be analysed at higher extent by using advanced machine learning techniques and it can be integrated with flask python framework as an web app that provides useful tools and features that make creating web applications in Python .

Using flask the interactive web can be created as it is a web framework and it is a Python module that lets you develop web applications easily. It has a small and easy-to-extend core which is also a microframework which can be used for the analysis.

13. APPENDIX

Source Code:

Index.html :

```
<!DOCTYPE html>

<html lang="en">

<head>

  <meta charset="UTF-8">

  <meta http-equiv="X-UA-Compatible" content="IE=edge">

  <meta name="viewport" content="width=device-width, initial-scale=1.0">

  <link href="https://cdn.jsdelivr.net/npm/bootstrap@5.2.1/dist/css/bootstrap.min.css"
rel="stylesheet">

  <title>Corporate Employee Attrition Analysis</title>

  <link rel="stylesheet" href="index.css">

</head>

<body>

  <div class="head-wrapper">

    <h2><b><center>Corporate Employee Attrition Analysis</center></b></h2>

  </div>

  <div>

    <p id="head-text">Employee attrition is referred to as the regular, but unexpected and
uncontrollable, decline in the workforce caused by retirement, illness, or death. Attrition contrasts from
attrition in that the employer does not fill a vacancy that had emerged in a position after it has already
been filled. The underlying cause of employee attrition are explored in this guide, along with suggestions
on how employers might use attrition to their favor.</p>

  </div>

  <div class="content">
```

```

<table width="100%" height="90%">

  <tr id="container">

    <td>

      <center></center>

      <a href="dashboard.html"><center class="text">Dashboard</center></a>

    </td>

    <td>

      <center></center>

      <a href="report.html"><center class="text">Report</center></a>

    </td>

    <td>

      <center></center>

      <a href="story.html"><center class="text">Story</center></a>

    </td>

  </tr>

</table>

</div>

</body>

</html>

```

Dashboard.html:

```

<!DOCTYPE html>

<html lang="en">

<head>

  <meta charset="UTF-8">

```



```

<meta http-equiv="X-UA-Compatible" content="IE=edge">

<meta name="viewport" content="width=device-width, initial-scale=1.0">

<link href="https://cdn.jsdelivr.net/npm/bootstrap@5.2.1/dist/css/bootstrap.min.css"
rel="stylesheet">

<title>Dashboard</title>

<link rel="stylesheet" href="index.css">

</head>

<body>

<div class="head">

<center><b><h2>Dashboard</h2></b></center>

</div>

<div>

<table width="100%" height="100%">

<tr>

<h3 class="content-head">Employee Attrition Dashboard for general data with cognos tool
</h3>

<center><p><iframe
src="https://us3.ca.analytics.ibm.com/bi/?perspective=dashboard&pathRef=.public_folders%2FFinal%2F
General%2Bdata%2Bdashboard&action=view&mode=dashboard&subView=model0000018471546a1b_
00000000" width="1200" height="800" frameborder="0" gesture="media" allow="encrypted-media"
allowfullscreen=""></iframe></p></center>

</tr>

<tr>

<h3 class="content-head">Employee Attrition Dashboard for employee and manager data
with cognos tool </h3>

<center><p><iframe
src="https://us3.ca.analytics.ibm.com/bi/?perspective=dashboard&pathRef=.public_folders%2FFinal%2F
Employee%2B-
%2BManager%2Bdashboard&action=view&mode=dashboard&subView=model0000018470dc3186_000
00000" width="1200" height="800" frameborder="0" gesture="media" allow="encrypted-media"
allowfullscreen=""></iframe></p></center>

```

```
        </tr>

    </table>

</div>

</body>

</html>
```

Report .html:

```
<!DOCTYPE html>

<html lang="en">

<head>

    <meta charset="UTF-8">

    <meta http-equiv="X-UA-Compatible" content="IE=edge">

    <meta name="viewport" content="width=device-width, initial-scale=1.0">

    <link href="https://cdn.jsdelivr.net/npm/bootstrap@5.2.1/dist/css/bootstrap.min.css" rel="stylesheet">

    <title>Report</title>

    <link rel="stylesheet" href="index.css">

</head>

<body>

    <div class="head">

        <center><b><h2>Report</h2></b></center>

    </div>

    <div>

        <table width="100%" height="100%">

            <tr>

                <h3 class="content-head">Employee Attrition Report for general data with cognos tool </h3>
```

```
<center><p><iframe
src="https://us3.ca.analytics.ibm.com/bi/?pathRef=.my_folders%2FFinal%2FReport%2Bview%2Bof%2
BGeneral%2Bdata%2Breport&action=run&format=HTML&prompt=false" width="1200" height="800"
frameborder="0" gesture="media" allow="encrypted-media" allowfullscreen=""></iframe></p></center>
```

```
</tr>
```

```
<tr>
```

```
<h3 class="content-head">Employee Attrition Report for employee and manager data with
cognos tool </h3>
```

```
<center><p><iframe
src="https://us3.ca.analytics.ibm.com/bi/?pathRef=.my_folders%2FFinal%2FReport%2Bview%2Bof%2
BEmployee%2B-%2BManager%2Breport&action=run&format=HTML&prompt=false" width="1200"
height="800" frameborder="0" gesture="media" allow="encrypted-media"
allowfullscreen=""></iframe></p></center>
```

```
</tr>
```

```
</table>
```

```
</div>
```

```
</body>
```

```
</html>
```

Story .html:

```
<!DOCTYPE html>
```

```
<html lang="en">
```

```
<head>
```

```
<meta charset="UTF-8">
```

```
<meta http-equiv="X-UA-Compatible" content="IE=edge">
```

```
<meta name="viewport" content="width=device-width, initial-scale=1.0">
```

```
<link href="https://cdn.jsdelivr.net/npm/bootstrap@5.2.1/dist/css/bootstrap.min.css" rel="stylesheet">
```

```
<title>Story</title>
```

```
<link rel="stylesheet" href="index.css">
```

</head>

<body>

<div class="head">

<center><h2>Story</h2></center>

</div>

<div>

<table width="100%" height="100%">

<tr>

<h3 class="content-head">Employee Attrition Story for general data with cognos tool </h3>

<center><p><iframe
src="https://us3.ca.analytics.ibm.com/bi/?perspective=story&pathRef=.public_folders%2FFinal%2Fgeneral%2Bdata%2Bstory&action=view&sceneId=model000001847184603a_000000000&sceneTime=0"
width="1200" height="800" frameborder="0" gesture="media" allow="encrypted-media"
allowfullscreen=""></iframe></p></center>

</tr>

<td>

<h3 class="content-head">Employee Attrition Story for employee and manager data with cognos tool </h3>

<center><iframe
src="https://us3.ca.analytics.ibm.com/bi/?perspective=story&pathRef=.public_folders%2FFinal%2FEmployee%2B-%2BManager%2Bstory&action=view&sceneId=model00000184712564c0_000000000&sceneTime=139500"
width="1200" height="800" frameborder="0" gesture="media" allow="encrypted-media"
allowfullscreen=""></iframe></p></center>

</tr>

</table>

</div>

</body>

</html>

Index .css:

```
body
{
    margin: 0;
    padding: 0;
    background-color: lightblue;
}
#D-img
{
    margin: 0;
    width: 150px;
    height: 150px;
    border-radius: 1rem;
}
.text
{
    color: black;
    font-size: 20px;
    font-weight: 700;
    font-family: Arial, Helvetica, sans-serif;
}
.head-wrapper
{
    margin: 0;
    padding: 3rem 4rem;
    font-family: Georgia, 'Times New Roman', Times, serif;
    font-weight: 900;
}
.head
{
    margin: 0;
    padding: 1rem 4rem;
}
#head-text
{
    margin: 0;
    padding: 20px;
    font-size: 18px;
    font-weight: 400;
    text-align: justify;
    font-family: Arial, Helvetica, sans-serif;
    line-height: 45px;
}
```

```
.content
{
  margin: 0;
  padding: 5rem;
}
.content-head
{
  margin: 0;
  padding: 3rem;
  text-align: center;
  font-size: 20px;
}
```

Index.php:

```
<?php include_once("index.html"); ?>
```

Machine Learning Code:

```
import numpy as np
import pandas as pd
import matplotlib.pyplot as plt
%matplotlib inline
import seaborn as sns

import os
os.chdir("E:/IBM/IBM/Datasets")

df = pd.read_csv('general_data.csv')

df

df.columns

df.shape

df.info()

df.describe()

df.isnull().sum()
```

```
df['NumCompaniesWorked']=df['NumCompaniesWorked'].fillna(df['NumCompaniesWorked'].mean())
```

```
df['TotalWorkingYears']=df['TotalWorkingYears'].fillna(df['TotalWorkingYears'].mean())
```

```
df.isnull().sum()
```

```
df.isnull().values.any()
```

```
df.head()
```

```
df['Attrition'].value_counts()
```

```
sns.countplot(df['Attrition'])
```

```
fig_dims = (12, 4)
```

```
fig, ax = plt.subplots(figsize=fig_dims)
```

```
sns.countplot(x='Age', hue='Attrition', data = df, palette="colorblind", ax = ax,  
edgecolor=sns.color_palette("dark", n_colors = 1));
```

```
for column in df.columns:
```

```
    if df[column].dtype == object:
```

```
        print(str(column) + ' : ' + str(df[column].unique()))
```

```
        print(df[column].value_counts())
```

```
df = df.drop('EmployeeCount', axis = 1)
```

```
df = df.drop('Over18', axis = 1)
```

```
df.corr()
```

```
df.cov()
```

```
plt.figure(figsize=(14,14)) #14in by 14in
```

```
sns.heatmap(df.corr(), annot=True, fmt='.0%')
```

```
X = df.iloc[:, 0:1].values
```

```
y = df.iloc[:, 1]
```

```
X
```

```
y
```

```
from sklearn.model_selection import train_test_split
X_train, X_test, y_train, y_test = train_test_split(X, y, test_size=0.4, random_state=1)
```

```
X_train
y_train
```

```
X_train.shape
y_train.shape
```

```
from sklearn.model_selection import train_test_split
X_train, X_test, Y_train, Y_test = train_test_split(X, y, test_size = 0.4, random_state = 0)
```

```
from sklearn.ensemble import RandomForestClassifier
forest = RandomForestClassifier(n_estimators = 100, criterion = 'entropy', random_state = 0)
forest.fit(X_train, Y_train)
```

```
forest.score(X_train, Y_train)
```

```
from sklearn.ensemble import RandomForestClassifier
forest = RandomForestClassifier(n_estimators = 200, criterion = 'entropy', random_state = 0)
forest.fit(X_train, Y_train)
```

```
from sklearn.model_selection import train_test_split
from sklearn.preprocessing import LabelEncoder
from sklearn import metrics
from sklearn.metrics import mean_absolute_error, mean_squared_error, r2_score
from sklearn.svm import SVR
from sklearn.tree import DecisionTreeRegressor
from sklearn.ensemble import RandomForestRegressor
from sklearn import neighbors
from math import sqrt
%matplotlib inline
```

```
X = df.iloc[:, 0:1].values
y = df.iloc[:, 1]
```

```
from sklearn.model_selection import train_test_split
X_train, X_test, y_train, y_test = train_test_split(X, y, test_size=0.25, random_state=1)
```

```
from sklearn.tree import DecisionTreeClassifier
```



```
tree_classifier = DecisionTreeClassifier(max_depth = 10, random_state = 42)
tree_classifier.fit(X_train,y_train)

y_tree_pred = tree_classifier.predict(X_test)

from sklearn.metrics import confusion_matrix, accuracy_score, f1_score, precision_score,
recall_score
acc = accuracy_score(y_test, y_tree_pred)

acc
```

Github & Demo link

Github link:

The Github link is provided so as to explain the various information about the final deliverables that are presented so far related to the project.

<https://github.com/IBM-EPBL/IBM-Project-45609-1660731223>

Demo Link:

The demo link is that the demonstration video of the project that is submitted.

https://drive.google.com/file/d/1asHdrS3_naQVwmbaa5q8E7ykXfkDBaT6/view

Wep App Link:

The link direct to the wep app where the analysis using IBM Cognos Analytics is done.

<https://employee-attrition-analysis.herokuapp.com/>