# CORPORATE EMPLOYEE ATTRITION ANALYTICS

## DATA ANALYTICS DOMAIN

**TEAM ID:** PNT2022TMID11482

## A PROJECT REPORT

***Submitted by***


**ARAVINTHARAJ T (910619106004)**

**GOWTHAM R (910619106018)**

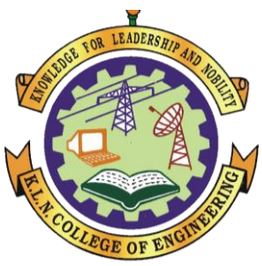
**HARIPRASATH M (910619106019)**

**NITHISH KUMAR A (910619106035)**

**ELECTRONICS AND COMMUNICATION ENGINEERING**

**K.L.N. COLLEGE OF ENGINEERING**

**POTTAPALAYAM**


**TEAM MENTOR**

**- DR. B.BUVANESWARI**

**INDUSTRY MENTOR**

**- SHANAWAZ ANWER**

**EVALUATOR**

**- L.MEENAKSHI**

**November 2022**

**ANNA UNIVERSITY : CHENNAI 600025**

**BONAFIDE CERTIFICATE**

**Certified that this project report "WEB PHISHING DETECTION USING MACHINE LEARNING" is the bonafide work of**

**"ARAVINTHARAJ T (910619106004),**

**GOETHAM R (910619106018),**

**HARIPARASTH M (910619106019),**

**NITHISH KUMAR A (910619106035)",**

Who carried out the project work under our supervision.

| **SIGNATURE** | **SIGNATURE** |
|---|---|
| **FACULTY MENTOR** | **FACULTY EVALUATOR** |
| DR. T.R. MUTHU | DR. S.SUBHA |
| ASSISTANT PROFESSOR (Sr.Gr) | ASSISTANT PROFESSOR |
| ELECTRONICS AND | ELECTRONICS AND |
| COMMUNICATION ENGINEERING | COMMUNICATION ENGINEERING |
| K.L.N COLLEGE OF ENGINEERING | K.L.N COLLEGE OF ENGINEERING |

**SIGNATURE**

**DR.V.KEJALAKSHMI**

**HEAD OF THE DEPARTMENT**

**ELECTRONICS AND COMMUNICATION ENGINEERING**

**K.L.N COLLEGE OF ENGINEERING**

**PROJECT REPORT**

**CORPORATE EMPLOYEE ATTRITION**

**ANALYTICSTEAM ID:PNT2022TMID11482**

## TABLE OF CONTENTS

7. **CODING & SOLUTIONING (Explain the features added in the project along with code)**
    1. Feature 1
    2. Feature 2
8. **TESTING**
    1. Test Cases
    2. User Acceptance Testing
9. **RESULTS**
    1. Performance Metrics
10. **ADVANTAGES & DISADVANTAGES**
11. **CONCLUSION**
12. **FUTURE SCOPE**
13. **APPENDIX**

Source Code

GitHub & Project Demo Link

# 1. INTRODUCTION

## PROJECT OVERVIEW

Employee attrition is defined as the unpredictable and uncontrollable, but normal,reduction of the workforce due to resignations, retirement, sickness, or death. The
employee attrition rate measures the number of people who move out of a company and are not replaced. Employee retention is crucial to your company's success and the attrition rate is the metric that provides insight into how well you're retaining your
employees.

Measuring and analysing your company's attrition rate will allow you to determinehow many employees left your company within a certain period and understand why
they left. Then, you can develop relevant and effective retention strategies to reduceattrition within your organisation.

## PURPOSE:

The key to success in any organization is attracting and retaining top talent. As anHR analyst one of the key task is to determine which factors keep employees at the company and which prompt others to leave. Given in the data is a set of data points on the employees who are either currently working within the company or have resigned.
The objective is to identify and improve these factors to prevent loss of good people.

## 2. LITERATURE SURVEY

### EXISTING PROBLEM:

So why is attrition such a problem? Among the most common reasons forturnover in the contact centre are:

- **Inappropriate candidate selection for the role and the culture:** Regardless of industry, survey after survey finds employers complaining about how difficult hiring is. When organizations are unable to identify individuals who will do theirjob well within their specific organization, attrition skyrockets.

- **Employee disengagement:** Employees who aren't engaged are more likely toleave. One Gallup survey found that organizations whose employees reportedhigh engagement had 25% to 65% less attrition than their peers.

- **Low job satisfaction:** Employee satisfaction is directly linked to attrition, both inthe short-term (immediately after hire) and the long-term.

- **Excessive pressure and stress:** Workplace stress in America is estimated at over $300 billion annually, and some experts estimate that up to 40% of turnover isdue to stress.

- **Inflexible working environment:** Employees who are unable to

balance theirprofessional and personal lives, including taking time off for doctor's
appointments and to care for children, are more likely to search out other, moreflexible, employers. In fact, one survey reported by the Society for Human
Resource Management found that a majority of employees (78%) said a flexiblework arrangement would allow them to live a healthier life, while 86% said theywould be less stressed.

- **No career growth or development opportunities:** Global Talent Monitor's report on workforce activity shows that the lack of future career development remains akey driver of employee attrition — 40% of departing employees say it led them tobe dissatisfied with their jobs, Gallup reports.

- **Lack of recognition:** More than 65 percent of employees report they don't feelrecognized at work, according to Human Resources Today, and this fuels attrition.

- **Abusive calls:** One researcher found that agents can average up to 10 hostile encounters a day, which takes a toll on morale and encourages employees toleave the organization.

On their own, each of these factors has a significant impact on attrition. Together, theymagnify the challenge contact centres face in keeping attrition at a manageable level.

**REFERENCES:**


**1.** Abdulquadri Ade, Ajagbe Musibau Akintunde, Sholanke Anthony Babatunde, Sani Tolani Abdulrahman, (2015), "Impact of Employee Turnover in Small and Medium Construction Firms: A Literature Review", International Journal of Engineering Research & Technology (IJERT), Volume 4, Issue 2, (p. 976)

**2.** Chandrasekar, K.(2011). Managing Attrition: The Real Problem Behind the Growth of Business Process Outsourcing (BPO) Companies. Journal of Social Welfare &Management, 3.

**3.** David Shine, (2015), "Attrition in IT", International Journal Of Core Engineering &Management (IJCEM), Volume 2, Issue 1, (p. 236)

**4.** Dorance Jeen S. Batty, (2014), "A Study on Attrition – Turnover Intentions in retail Industry", International Journal of Business and Administration Research Review",Volume 1, Issue 3, (p. 55)

**5.** Farooque Asma, Habibuddin, (2015), "Employee Attrition Analysis of Hotel Industry", International Journal of Advancement in Engineering Technology, Management & Applied Science, Volume 2, Issue 12, (p. 34)

## PROBLEM STATEMENT DEFINITION:

**I am** an employee. **I am trying to** collect and analyse the data sets with a proper understandablilty and process, implement for further analyses.**But** I am unable to collect the proper datasets and procceed to net steps. **It is because** of the improper datasets collected which may have some missing values , false values or null values.**This makes me to feel** to maintain the proper records of datasets further to process the implementation of attiriton analytics.

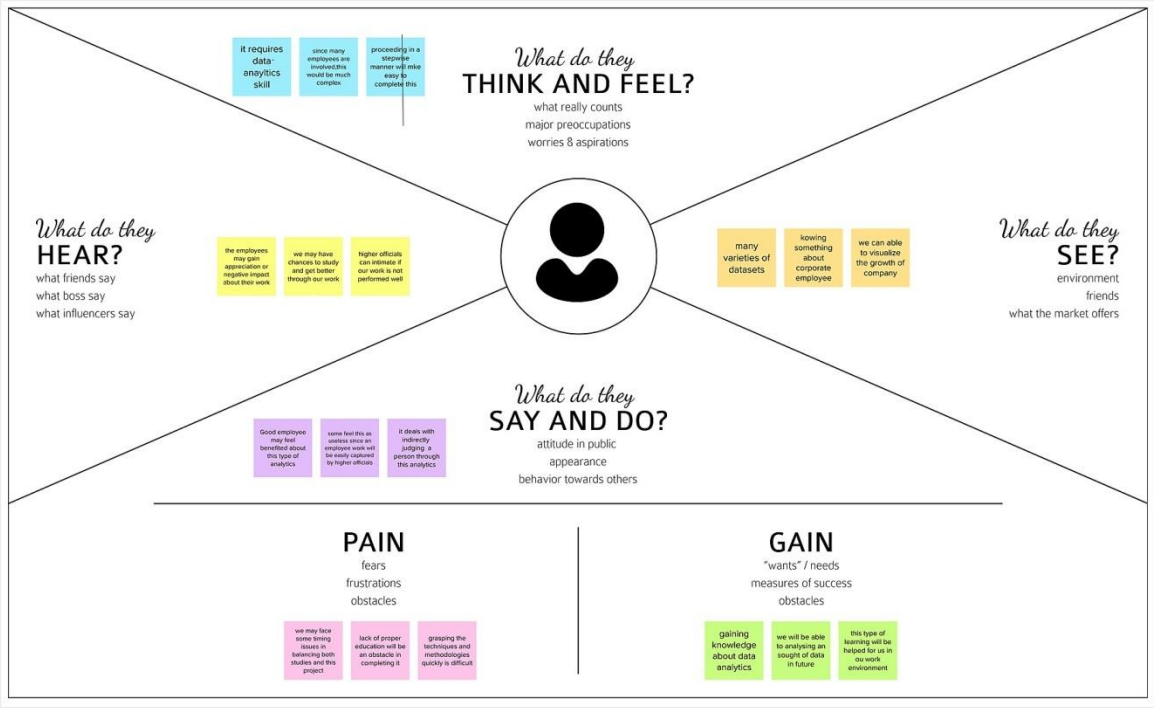
## 3. IDEATION &PROPOSED SOLUTION

### EMPATHY MAP CANVAS:

An empathy map is **a collaborative tool teams can use to gain a deeperinsight 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.

1 CORPORATE EMPLOYEE ATTRITION ANALYTICS

Build empathy and keep your focus on the user by putting yourself in their shoes.

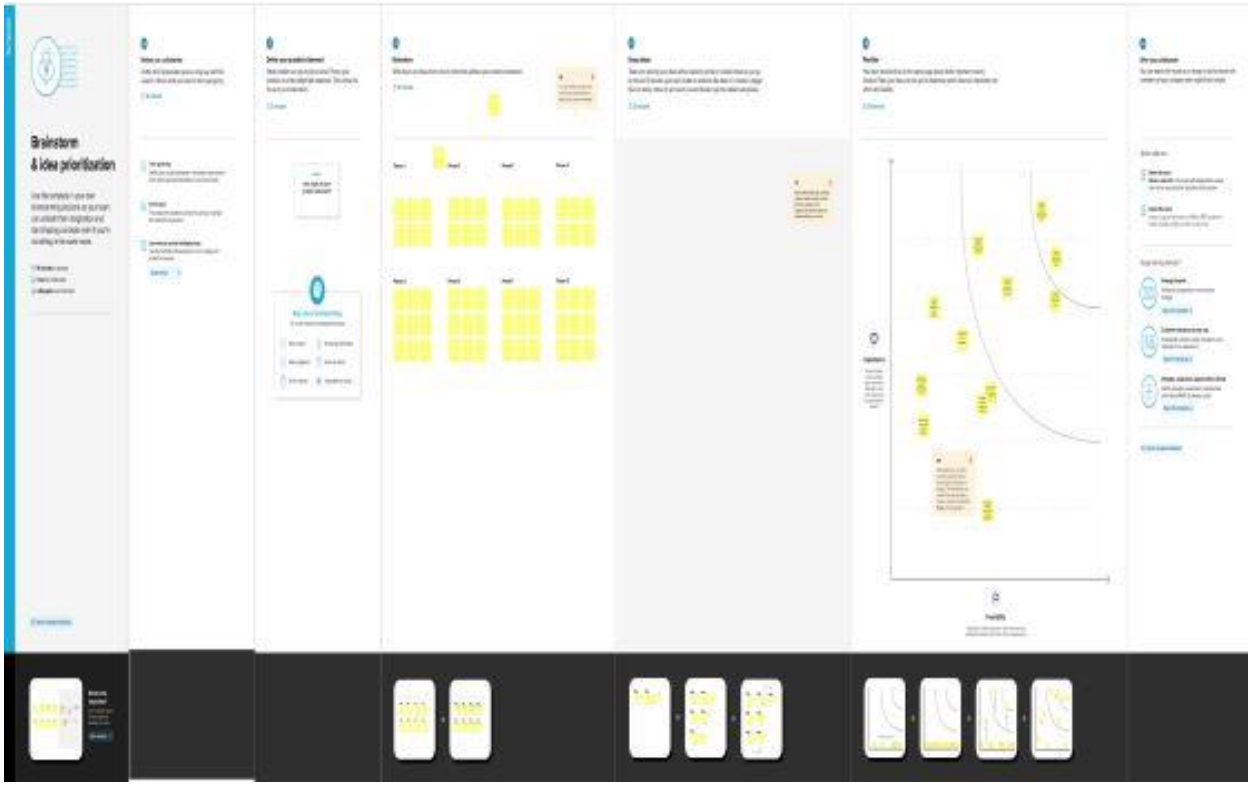



Share your feedback

## IDEATION AND BRAINSTORMING:

Ideation is **the process of forming ideas from conception to implementation, most often in a business setting**. Ideation is expressed via graphical, written, or verbal methods, and arises from past or present knowledge, influences, opinions, experiences, and personal convictions.

Ideation is often closely related to the practice of brainstorming, a specific technique that is utilized to generate new ideas. A principal difference between ideation and brainstorming is that **ideation is commonly more thought of as being an individual pursuit, while brainstorming is almost always a group activity**.


Brainstorm
& idea prioritization

## 3.4 PROBLEM SOLUTION FIT:

### A. CUSTOMER SEGMENTS

- The customer of this project will be the HR professionals,the administration orthe person with the higher power authority who are responsible for their lower level employees.
- The customer uses the employee data.

### B.JOBS-TO-BE-DONE / PROBLEMS

- Initially the data has to be collected and formatted in a proper way.
- A deep analysis of the employee data should be done In order to gain the results.
- The problem which may arise here is sometimes the data may be an invalid orincorrect data which affects the results.

### C. TRIGGERS

- With this analysis,the employee will be more aware of hisresponsibilities being done.
- It encourages good employees to step forward in their career and it serves as awarning for those employees who are not being responsible in their work.

### D. EMOTIONS: BEFORE / AFTER

- The good employees will be encouraged and the irresponsible one will benoticed.

## E. AVAILABLE SOLUTIONS

- Initially the performance of the employee is observed manually by the higherofficials.
- But this may led to imbalance in treating all employees as same.
- But the analysis will be completely digital so that there any not occur anyfavourism.

## F. CUSTOMER CONSTRAINTS

- The constraints which the customer would face may be the lack of skilled employee or the amount of surplus employee would bring the issue in decisionmaking in taking the appropriate results.

## G. BEHAVIOUR

> Directly related with the higher authorities.

> Indirectly related with the knowledge of the employees.

## H. CHANNELS OF BEHAVIOUR

> ONLINE:

The customers can perform visualization using different graphs,can drawsmany useful insights from it.

> OFFLINE:

Using the results which was collected the action may be taken offline. Preparing data sets can be done offline.

## I. PROBLEM ROOT CAUSE

The main thing to do an analytics of this is :

> to identify the potential employees

> To find the reason of employee attrition

> To improve the organization profit by retaining good talents

> To consider every employee performance in bias.

## J. SOLUTION

> The solution would be the attrition anayltics which gains the useful resultswhich may be beneficial both to the employees as well as to the organization.

# 4. REQUIREMENT ANALYSIS

## FUNCTIONAL REQUIREMENTS:

| FR No. | Functional Requirement (Epic) | Sub Requirement (Story / Sub-Task) |
|---|---|---|
| FR-1 | Data collection | 1. Collecting data from companyadministration<br>2. Collecting data from company database |
| FR-2 | Representation of data | 3. Presenting data in excel sheets |
| FR-3 | Analysing data | 4. Performing various visualizationtechniques<br>5. Creating dashboards,stories,graphs,charts,etc. |
| FR-4 | Preparing reports | 6. Analysing the results from the dataanalysis performed.<br>7. Preparing reports<br>8. Taking necessary action if required. |

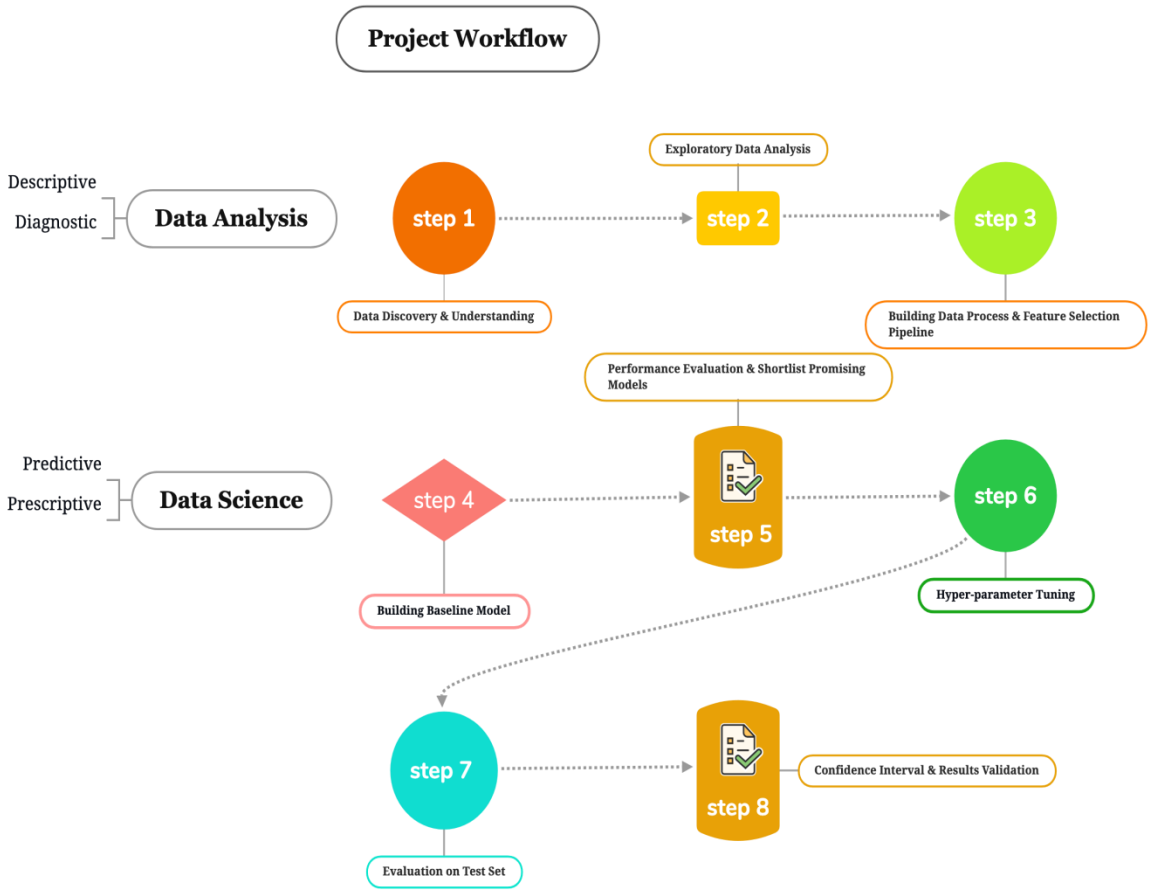
## NON FUNCTIONAL REQUIREMENTS:

| FR No. | Non-Functional Requirement | Description |
|---|---|---|
| NFR-1 | **Usability** | The analysis platform should provide thecapacity to perform the task safely. |
| NFR-2 | **Security** | The data must be maintained in a securedmanner |
| NFR-3 | **Reliability** | The data analysis performed must be consistent and be trustworthy to take necessary actions when required. |
| NFR-4 | **Performance** | To perform analysis,the data analytics platform must have a good performancecriteria. |
| NFR-5 | **Availability** | The collected data or the performedanalysis must be available whenever required. |
| NFR-6 | **Scalability** | It must be scalable since the data size may vary at any time. |

# 5. PROJECT DESIGN

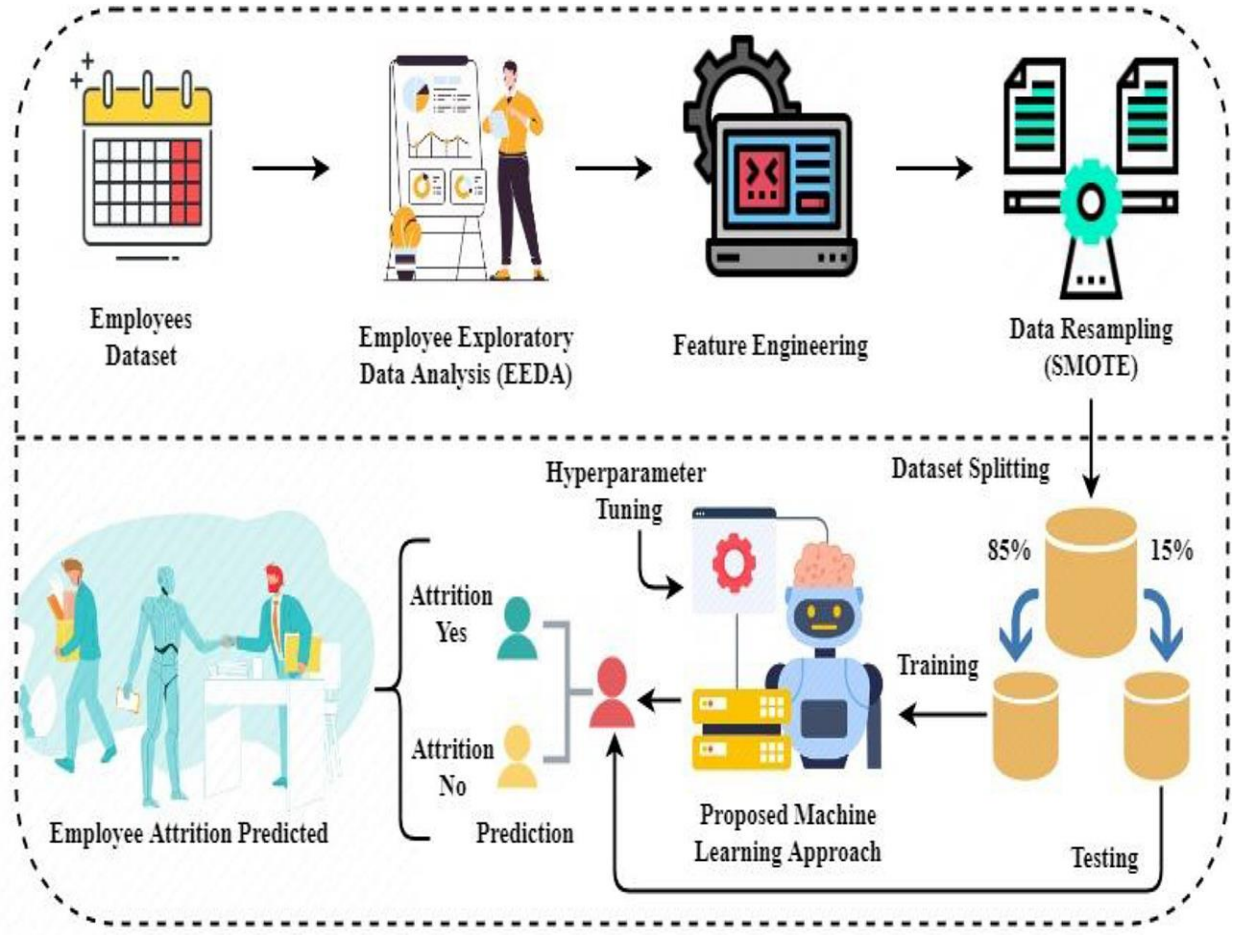
## DATAFLOW DIAGRAMS:

A data-flow diagram is a way of representing a flow of data through a process or asystem. The DFD also provides information about the outputs and inputs of each entity and the process itself. A data-flow diagram has no control flow — there are no decision rules and no loops.



## SOLUTION AND TECHNICAL ARCHITECTURE:

Based on the complexity of the deployment, a solution architecture diagram mayactually be **a set of diagrams documenting various levels of the architecture**. The diagram relates the information that you gather on the environment to both physical andlogical choices for your architecture in an easily understood manner.



**USER STORIES:**

- **CUSTOMER(SUPPORT ROLE WORKERS):**

As a user, I collect the required information about the corporate employee from the higher officials or from the office administration. As a user, I can also get the employee details through the company database. As a user, I segregate the data in a
representable form which is used for the further steps.

- **CUSTOMER(DATA ANALYST):**

As a user, I analyse the data through visualization. As a user, I analyse the data through dashboards As a user, I analyse the data in the form of stories,graph,reports,etc. As a user, I finally represent the results gained from the data analytics.

- **HR:**

As a user , I can prepare reports from the data analysis process. From the reports, I can take necessary actions which results in employee attrition.

## 6. PROJECT PLANNING AND SCHEDULING

### SPRINT PLANNING ND SCHEDULING:

| Sprint | Functional Requirement (Epic) | User Story Number | User Story / Task | Story Points | Priority |
|---|---|---|---|---|---|
| Sprint-1 | Collecting and preparing datasets | USN-1 | As a user, I collect the required information about the corporate employee from the higher officials or from the office administration. | 2 | low |
| Sprint-1 | | USN-2 | As a user, I can also get the employee details through the | | High |

| | | | | | |
|---|---|---|---|---|---|
| | | | company database. | | |
| Sprint-1 | | USN-3 | As a user, I segregate the data in a representable form which is used for the further steps. | 1 | high |
| Sprint-2 | Data visualization | USN-1 | As a user,I analyse the data through visualization | 2 | medium |
| Sprint-2 | | USN-2 | As a user, I analyse the data through dashboards | | high |
| Sprint-2 | | USN-3 | As a user, I analyse the data in the form of stories,graph,reports ,etc. | | low |
| Sprint-3 | Data analysing | USN-1 | As a user, I finally represent the results gained from the data analytics using | 2 | high |

| | | | | | |
|---|---|---|---|---|---|
| | | | python | | |
| Sprint-3 | | USN-2 | Through python,I can calculate the attrition results | | medium |
| Sprint-4 | Reporting the results | USN-1 | As a user , I can prepare reports from the data analysis process | 1 | medium |
| Sprint-4 | | USN-2 | From the reports, I can take necessary actions which results in employee attrition. | | low |

**SPRINT DELIVERY SCHEDULE:**

| **S**print | **Sprint Start Date** | **Sprint End Date (Planned)** | **Sprint Completed (as on Planned End Date)** | **Sprint Release Date (Actual)** |
|---|---|---|---|---|
| Sprint 1 | 24 Oct 2022 | 29 Oct 2022 | 29 OCtober 2022 | 05 Novembet 2022 |
| Sprint 2 | 31 Oct 2022 | 05 Nov 2022 | 05 November 2022 | 06 November 2022 |
| Sprint 3 | 07 Nov 2022 | 12 Nov 2022 | 08 November 2022 | 09 November 2022 |
| Sprint 4 | 14 Nov 2022 | 19 Nov 2022 | 11 November 2022 | 14 November 2022 |

## 7. CODING AND SOLUTIONING

### FEATURE 1:

In this project , we have done visualization by considering several criterias like

- age
- gender
- department
- business travel
- number of companies worked
- monthly income, etc.

Considering all this during the visualization process makes it more accurate toexactly identify the root caue for the attrition of the employees.

### FEATURE 2:

The dataset is also understood by various factors to consider the missing or unnecessary values in it. Python is used inorder to make the process quite easy and visualisation is also performed using python.Random forest classifier is used in trainingand testing the datasets which yields almost 97% of accuracy.

# 8. TESTING

## TESTCASES:

A test case is a set of actions performed on a system to determine if it satisfies software requirements and functions correctly.A test case is a document,which has a set of test data, preconditions, expected results and postconditions, developed for a particular test scenario.

In this project the system is tested and trained based on the empoyees attrition.

- What are all the conditions which lead to employee attrition?
- Which factor is more responsible for the attrition to occur?
- What are the root causes which lead to the attrition?
- How to retain talented employees?
- What necessary steps to be taken to avaoid these conditions?

## USER ACCEPTANCE TESTING:

User acceptance testing (UAT), also called application testing or end-user testing, is a phase of software development in which the software is tested in the realworld by its intended audience.UAT is often the last phase of the software testing
process and is performed before the tested software is released to its intended

market.The goal of UAT is to ensure software can handle real-world tasks and perform up to development specifications.

In UAT, users are given the opportunity to interact with the software before its official release to see if any features have been overlooked or if it contains any bugs. UAT can
be done in-house with volunteers, by paid test subjects using the software or by makingthe test version available for download as a free trial. The results from the early testers are forwarded to the developers, who make final changes before releasing the softwarecommercially.UAT is effective for ensuring quality in terms of time and software cost, while also increasing transparency with users.

# 9. RESULTS

## 9.1 PERFORMANCE METRICS:

**1. Attrition status by age :**

- visualization performed by column chart
- Age by status = 92%

**2. Employee count by department :**

Visualization performed by bar

chartEmployee count by

department wise

- Human resource = 17%
- R&D = 89%
- Sales = 60%

**3. Attrition based on business travel :**

Visualization performed by waterfall chart percentage by business travel

- Non-travel =25%
- Travel frequently =75%
- Travel rarely =35%
- Sum=100%

**4. Attrition based on department ,job role ,education &marital status:**

Visualization performed by line & column chart

percentage byDepartment wise

**i.** Human resource =15%

**ii.** R&D =85%

**iii.** Sales =45%

**iv.** Education =69%

**v.** Job role =100%

**vi.** Marital status

- Male =80%
- Female =20%

**5. Attrition based on salary hike percentage :**

- Visualization performed by pie chart
- Salary hike percentage (overall) =95%(based on department wise)

6. **Based on No.of companies worked:** Visualization performed by stocked column chartNo. of companies worked based on attrition

**i.** Human resource =15%

**ii.** R&D =65%

**iii.** Sales =35%

**7. Visualization based on monthly income groups :**

Visualization performed by scatterplot chart

- Monthly income percentage = 100%

**8. Prediction based on employee working groups :**

Visualization performed by network chart employee working groups

- percentile =75%

**DASHBOARDS:**

**1. Attrition based on department by age department (visualization performed by bar chart)**

- Human resource =17%
- R&D =89%
- SALES =60%
- OVERALL =91%

**2. Analysis based on job involvement in daily rate :**

Visualization performed by heat plot chart job involvement

- percentage =99%

**3. Based on attrition :**

- Visualization performed by scatterplot
- Attrition percentage =66%
- Business travel =99%

**4. Calculating the employee performance :**

Visualization performed by column chart

- No.of companies worked =59%
- Performance rating =84%

## 1O.ADVANTAGES &DISADVANTAGES

**ADVANTAGES:**

- It brings to fore the cause of employee disengagement.
- Enables HR managers develop long-term strategies to reduce attrition
- Competitive measures to enhance company brand image
- Develops and shapes drills that benefit both the management and theemployees
- Enhanced work culture
- When employers do not fill up the vacant position, departmental workflows canbe changed. The company can assign new duties to the rest of the employeesand shift resources allocated for that position within the organization.
- When employees leave the company voluntarily, labor costs on the company reduce that can be allocated in other areas. Labor costs refer to the wages ofthose employees, along with appraisals and bonuses.
- Organizational culture improves if the employees that are troublesome andnegatively impact the company culture leave through attrition.
- When some employees retire or resign, this can be considered a fresh start forthe company, and the current employees can be offered new opportunities or
  new positions can be created that would add to the productivity of the company.
- If underperforming employees leave the company, this can pose a positive

impact on the company’s productivity. Underperforming employees often hinder their work efficiency and influence other employees if they are hired at important
positions. Through attrition, the company becomes able to solve this problemwithout turnover or layoffs.

One of the known facts about attrition is it cannot be eliminated from thecompany. It is a part of every organisation as employees are free to leave the organisation. However, through attrition analysis, one can only reduce the rate ofemployee exits.

**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 departingmay lead to lot of various losses to company on different aspects. The disadvantages can be

**1. 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 totrain the new employee who is to take over the job and the whole team gets affected. It can directly be seen in an overall decrease of performance of the team.Sometimes this may even lead to drastic change in customer relationship. Customers connect with
employees in an organization and those leaving all of a sudden may lead to doubts incustomer's minds as well.

**2. Daily task management:**

Sudden attrition may lead to difficulty in managing daily tasks. Even largeorganization struggle to manage their task when employees leave jobs, getting small information and managing daily tasks become difficult as they cannot be managed bysmall current team which is left behind. Organization generally have notice period to
ensure there is a smooth transition but attrition states otherwise, employees who

leavesuddenly leads to unmanageable daily routines.

### 3. 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 payingthe previous employee, hiring a new one, training cost for the new employee**.Research shows that these costs are way more than the losses incurred in managing and missingout on work.

**4. 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 highattrition 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.Older employees with their years of experience can take over critical matters which can never be trusted with these new employees.Even with employees who haveexperience are hired they may suffer at taking care of critical business matter as they are new to company's policies, culture and current employees.

**5. Create a Negative image:**

It is not just that employees are looking for job, even organizations are on theoutlook of qualified professionals. When any company has high attrition rate it negatively impact the brand of the organization.Recruiters' state that they find it difficultto map qualified candidates to the organization, as candidates opts out fearing the attrition rate. The reasons may vary but a negative image work against the organization.

**6. Employee development:**

Many organizations have various employee development plans and higher attrition rate means losing out on it.Employee development plans takes

time and hugeinvestments. When there is disturbance within the organization due to employees leaving the organization it affects the development process for all.The money investedon the employee who leave is wasted; also it affects others who have to jump in to fill infor the lost employee affecting their career plan and growth.These plans are structured and require dedication and time to reach the goals.

## 11. CONCLUSION

The study on attrition analysis highlighted, so many factors which will help to the employees. The study was conducted among 1241 employees and collected information through structured questionnaire and the project wascarried out in a good manner and has met the expectations of the organization.

The main objectives of a program to reduce a attrition in a corporate and to find the reason for the same. So that employees are better equipped to do this job for increase asalary and employees should satisfied with career development opportunity from the corporate side too.

## 12. FUTURE SCOPE

Further enhancement can be made by preserving the datasets prior in need.Several other methods can be implemented for analysing the datasets.Performingattrition anaytics in every corporate will help to achieve a better prfit to the organizationand the better environment which exists within the organization.

## 13. APPENDIX

### SOURCE CODE:

```
#Import the
libraries import
numpy as np import
pandas as pd import
seaborn as sns# load
the data
from google.colab import
filesuploaded =
files.upload()
#Store the data int a dataframe
df = pd.read_csv('WA_Fn-UseC_-HR-Employee-Attrition.csv')
#Print the first 100 rows
df.head(100)
#Get the rows and columns
df.shape
#Get the column data types
df.dtypes
# Get a count of the empty values of each column
df.isna().sum()
```

**#Check for any missign or null vlues in the data**

df.isnull().values.any

()**# View some**

**statistics**

df.describe()

**#Get a count of the number of employee that stayed and left the company**

df['Attrition'].value_counts()

**#Visualize the number of employees that stayed and left the company**

sns.countplot(df['Attrition'])

**#Checking the accuracy**

print((1233-237)/1233)

**#Show the number of employees that left and stayed by age**

```
import matplotlib.pyplot as
plt
plt.subplots(figsize=(12,4))
sns.countplot(x='Age',hue='Attrition',data=df,palette='colorblind')
```

**#Print all of the datatypes and their unique values**

```
for column in df.columns:
  if df[column].dtype == object:
    print(str(column) + ' : ' +
    str(df[column].unique()))
    print(df[column].value_counts())
    print('_____')
```

**#Removing some unnecessary**

```
columns df=df.drop('Over18',axis=1)
df=df.drop('EmployeeNumber',axis=1
) df=df.drop('StandardHours',axis=1)
df=df.drop('EmployeeCount',axis=1)
```

**#Get the corelation**

```
df.corr()
```

**#Visualize the corelation**

```
plt.figure(figsize=(14,14))
sns.heatmap(df.corr(),annot=True,fmt= '.0%')
```

**#Transform the data**

**#Transform non-numerical into numerical columns**

```
from sklearn.preprocessing import
```

```
LabelEncoderfor column in df.columns:
  if df[column].dtype == np.number:
    continue
  df[column] = LabelEncoder().fit_transform(df[column])
#Create a new
column
df['Age_Years'] =
df['Age']#Drop the
age column df =
df.drop('Age',axis=1)
#Show the data frame
```

```
df
#Split the data
X =
df.iloc[:,1:df.shape[1]].values
Y = df.iloc[:,0].values

#Split the data into 75% training and 25% testing
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=0)
# Use the Random forest classifier
from sklearn.ensemble import RandomForestClassifier
forest = RandomForestClassifier(n_estimators=10,criterion='entropy',random_state=0)
forest.fit(X_train,Y_train)
#Get the accuracy on the training dataset
forest.score(X_train,Y_train)
#Show the confusion matrix and accuracy score for the model on the test data
from sklearn.metrics import confusion_matrix
cm=
confusion_matrix(Y_test,forest.predict(X_test))
TN=cm[0][0]
TP=cm[1][1]
FN=cm[1][0]
FP=cm[0][1]
```

```
print(cm)
print('Model testing Accuracy = {}'.format((TP+TN) / (TP+TN+FN+FP)))
```

**GITHUB LINK:**

https://github.com/IBM-EPBL/IBM-Project-26724-1660034895

**PROJECT DEMO LINK:**

https://youtu.be/DhO0jfAfrr0