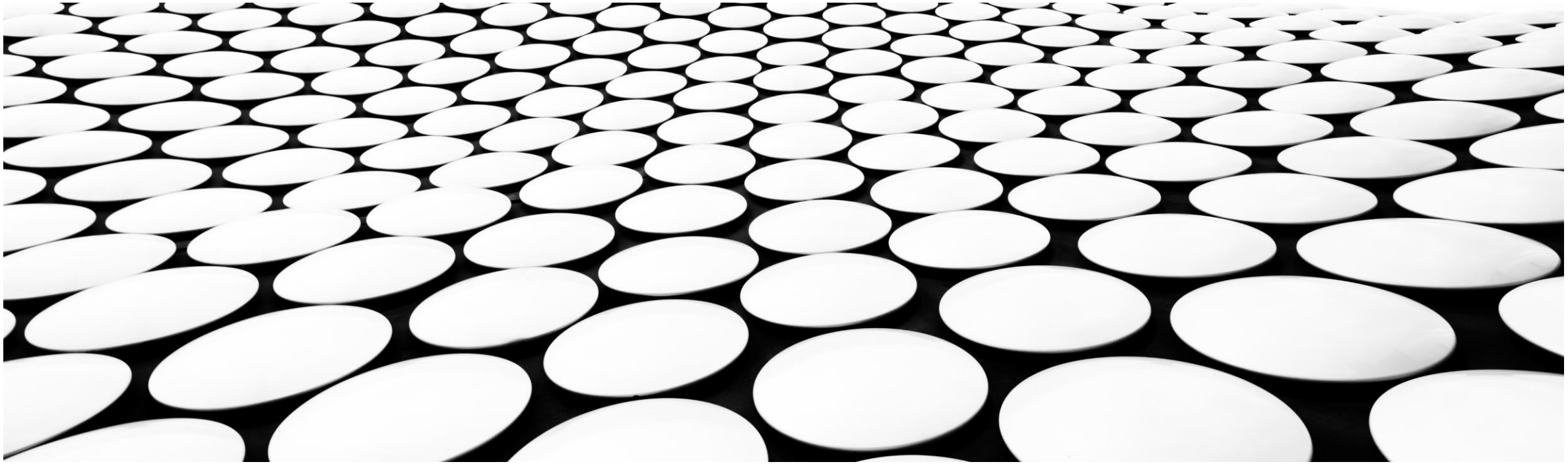
STUDENT DETAILS

NAME : A.SREE CHARAN KUMAR

COLLEGE : SRM UNIVERSITY AP

STREAM : AI & ML

INTERNSHIP COMPANY : EDUNET FOUNDATION



PROJECT TITLE/PROBLEM STATEMENT

PROJECT TITLE : EMPLOYEE BURNOUT PREDICTION USING LINEAR REGRESSION

PROBLEM STATEMENT:

- Employee burnout is a critical issue in modern workplaces, leading to reduced productivity, increased absenteeism, and higher turnover rates.
- This project aims to Determine the key factors contributing to employee burnout by analyzing various workplace variables such as work hours, job role, stress levels, and support from management.

AGENDA

- PROJECT TITLE/PROBLEM STATEMENT
- PROJECT OVERVIEW
- WHO ARE THE END USERS OF THIS PROJECT?
- YOUR SOLUTION AND ITS VALUE PROPOSITION
- MODELLING
- RESULTS

PROJECT OVERVIEW

Employee burnout is a significant issue affecting productivity, job satisfaction, and overall well-being in the workplace. This project aims to develop a predictive model using linear regression to identify employees at risk of burnout, enabling organizations to take proactive measures to prevent it.

Objectives

- **Identify Key Factors**: Determine the primary factors contributing to employee burnout by analyzing workplace variables.
- **Develop Predictive Model**: Create a linear regression model to predict the likelihood of burnout based on the identified factors.
- **Provide Insights and Recommendations**: Offer actionable insights and recommendations for organizations to mitigate burnout and improve employee well-being.

PROJECT OVERVIEW

Methodology

- **Data Preprocessing**: Cleaning and preparing the data by handling missing values, encoding categorical variables, and normalizing the data.
- **Exploratory Data Analysis (EDA)**: Analyzing the data to uncover patterns and relationships between features.
- **Feature Selection**: Identifying the most relevant features for the prediction model using statistical methods and domain knowledge.
- **Model Development**: Building and training the linear regression model using the training dataset.
- **Model Evaluation**: Assessing the model's performance using metrics such as R-squared, Root Mean Squared Error (RMSE), and Mean Absolute Error (MAE).

WHO ARE THE END USERS OF THIS PROJECT?

- **Human Resources (HR) Managers**

  Utilize the predictive model to identify employees at risk of burnout.

- **Line Managers and Team Leaders**

  Monitor the well-being of their team members using the model's predictions.

- **Company Executives and Leadership**

  Leverage insights from the model to inform strategic decisions regarding workforce management and resource allocation.

- **Employee Wellness Programs**

  Design and implement wellness initiatives based on the model’s insights.

- **Data Analysts and Researchers**

  Continuously improve the model by analyzing its performance and updating it with new data.

YOUR SOLUTION AND ITS VALUE PROPOSITION

- The solution involves developing a predictive model using linear regression to forecast employee burnout based on various workplace factors.
- The solution involves a comprehensive analysis of employee burnout within an organization, employing both quantitative and qualitative research methods. and to develop actionable recommendations to mitigate burnout and enhance employee well-being.
- This model will provide early warnings, enabling targeted interventions and support programs to be implemented proactively.
- The value of this solution lies in its ability to manage employee well-being effectively, increase job satisfaction, reduce turnover, and enhance organizational productivity.
- the solution fosters continuous improvement, ensures ethical and fair practices, and builds trust among employees.

HOW DID YOU CUSTOMIZE THE PROJECT AND MAKE IT YOUR OWN

- **Employee-Centric Approach:** Focusing on creating a solution that truly benefits employees by prioritizing their well-being and providing actionable, personalized interventions.
- **Continuous Learning:** Staying updated with the latest research and advancements in predictive analytics and employee well-being to keep improving the project.

MODELLING

- **Data Collection and Integration**

  Data Sources: Gather employee data from surveys, HR records, performance evaluations, and attendance logs.

- **Data Preprocessing**

  Cleaning: Handle missing values, outliers, and inconsistencies to prepare data for analysis.

- **Model Development**

  Utilize linear regression to predict burnout risk based on identified features.Split data into training and testing sets, train the model on the training set, and validate its performance

- **Model Evaluation**

  Performance Metrics: Assess model performance using metrics like R-squared, RMSE, and MAE to ensure accuracy and reliability.

- **Model Deployment and Maintenance**

  Monitoring: Continuously monitor model performance and update it with new data to maintain relevance and effectiveness.

STATISTICAL TECHNIQUES

Linear Regression Model :

Linear Regression Formula: $y=\beta0+\beta1x1+\beta2x2+\cdots+\beta n\ xn+\epsilon$

Where:

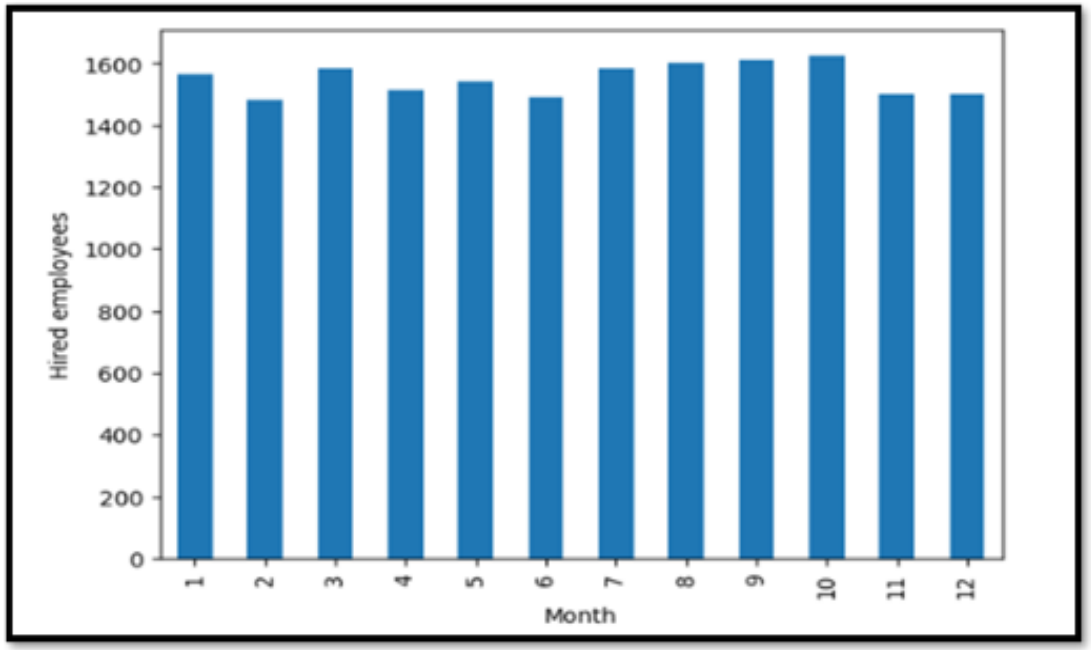
Y is the dependent variable.

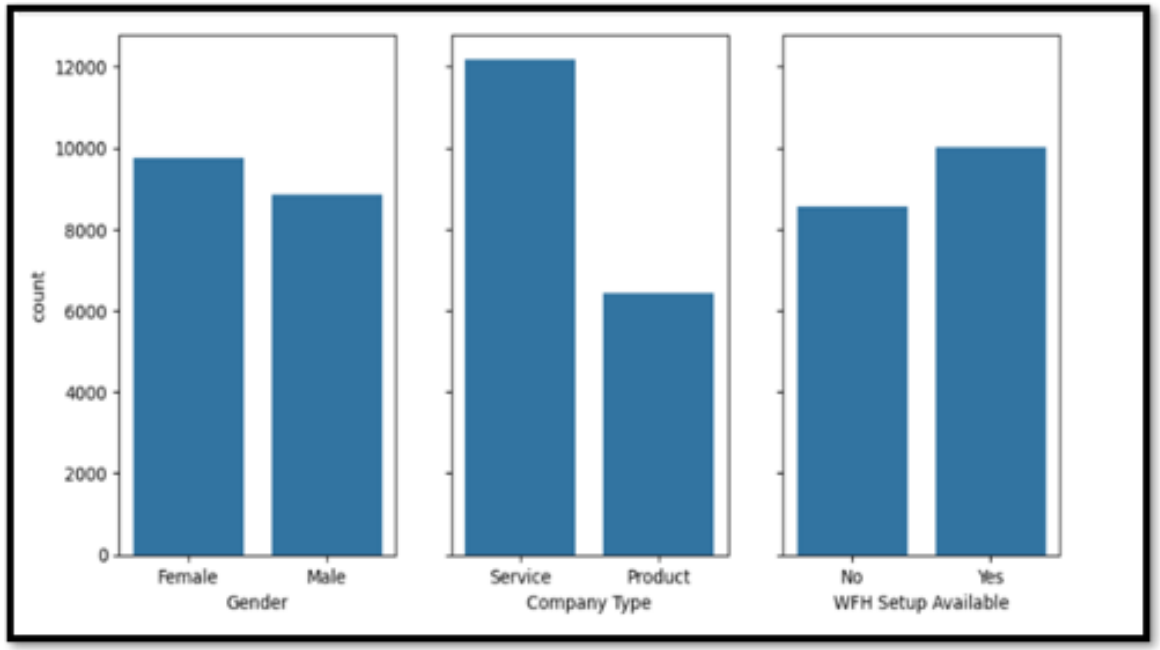
x1,x2,…,xn are the independent variables.

β1,β2,…,βn are the coefficients.

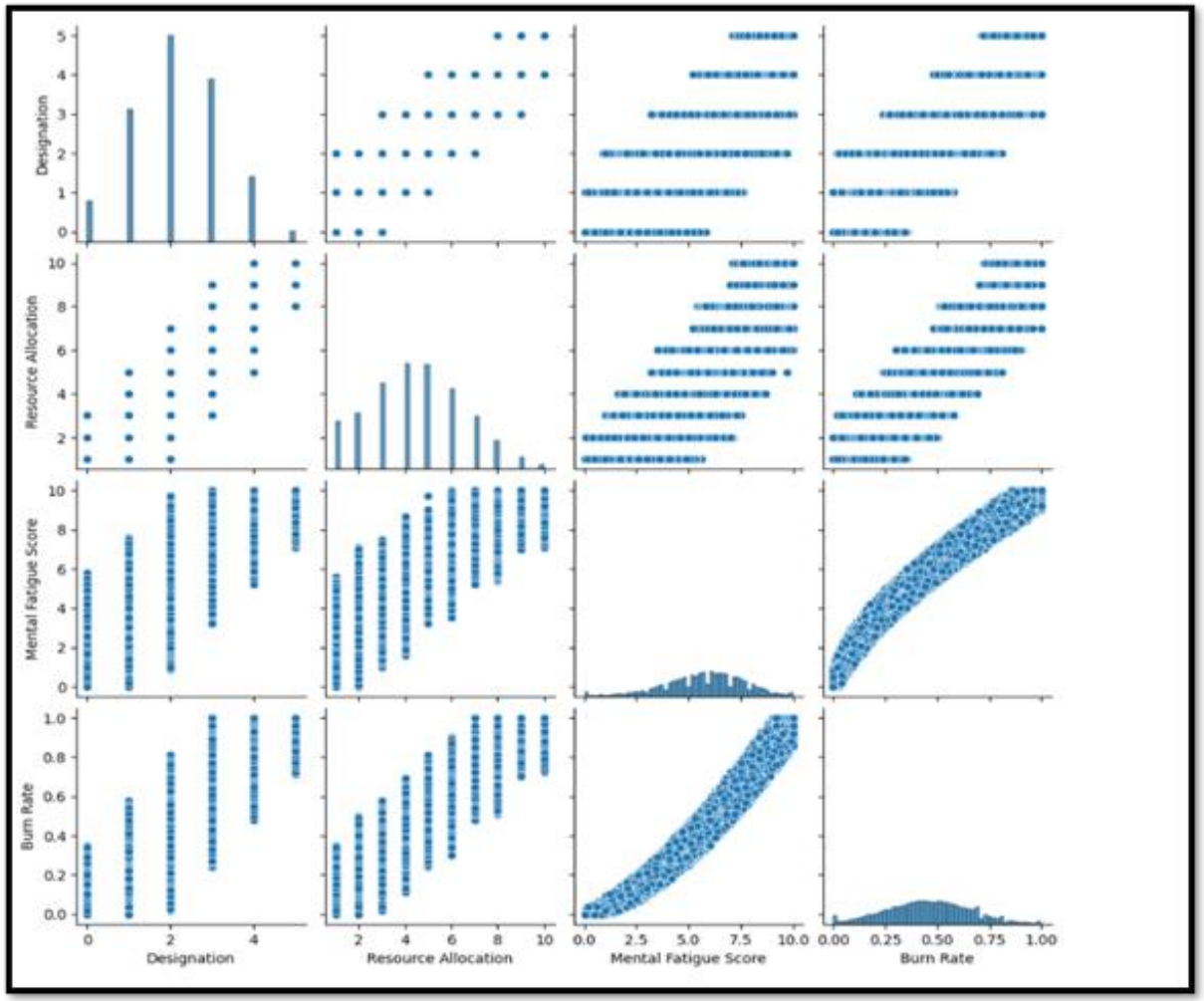
ϵ is the error term.

RESULTS


Hired employees
1600
1400
1200
1000
800
600
400
200
0
1
2
3
4
5
6
7
8
9
10
11
12
Month


count
12000
10000
8000
6000
4000
2000
0
Female
Male
Gender
Service
Product
Company Type
No
Yes
WFH Setup Available

RESULTS



Linear Regression Model Performance Metrics:

Mean Squared Error:
0.0031569779113610717

Root Mean Squared Error:
0.056186990588223 1

Mean Absolute Error:
0.04595032032644773

R-squared Score:
0.918822674247248

RESULTS IN OTHER METHODS

Support Vector Machine (Linear Kernel) Performance Metrics

Mean Squared Error: 0.003242003391288931
Root Mean Squared Error: 0.056938593162185976
Mean Absolute Error: 0.04613085977730072
R-squared Score: 0.9166363614901809

RandomForestRegressor Performance Metrics

Mean Squared Error: 0.0033848745846064457
Root Mean Squared Error: 0.058179675012898155
Mean Absolute Error: 0.045972449807427554
R-squared Score: 0.9129626261248232

Support Vector Machine (RBF Kernel) Performance Metrics

Mean Squared Error: 0.00341204574122986
Root Mean Squared Error: 0.058412719002199
Mean Absolute Error: 0.04962241383880883
R-squared Score: 0.9122639573681112

LINKS

https://colab.research.google.com/drive/1UNcIG3pULDxXAt2_xapuhMHk8eTA2MVy?usp=sharing