HUMAN RESOURCES ANALYTICS

(PROJECT PROPOSAL)

1. PROBLEM DEFINITION AND PLANNING:

Description:

Human resource analytics (HR analytics) is an area in the field of analytics that refers to applying analytic processes to the human resource department of an organization. The goal of the HR Analytics is to provide an organization with insights for effectively managing employees so that business goals can be reached quickly and efficiently. Here we are trying to get some insights about why valuable employees are leaving prematurely and who will leave next so that company can either take appropriate measures to retain them or they can plan so that company's performance will not get affected.

i. Problem Statement:

- Why Valuable employees are leaving prematurely?
- Which employees will leave next?

ii. Project deliverables:

- What makes an employee valuable?
- Who all are valuable employees?
- Define premature exit
- Categorize employees based on their major reason of exit
- Based on the above categories suggest measures to retain valuable employees
- Select an effective model to predict premature exit of employees

iii. Generate Success factors:

- Accuracy measures (Laplace and M estimate)
- ROC
- Confusion matrix (precision and recall)

iv. Resources:

(1) Software tools

- Asana- Project Management Tool
- GitHub- Version Control Tool
- R Studio and R
- Tableau and Plotly for visualization

(2) Project Members

- Divya Vasireddy
- Naveen Sampath
- Nivedita Merla
- Snehal Gawas

2. DATA PREPARATION:

v. *Data Source:*

Kaggle Dataset: https://www.kaggle.com/ludobenistant/hr-analytics

vi. *Summarize data:*

Number of observations = 15,000
Number of variables = 10

Variable Name	Description	Type
satisfaction_level	Level of satisfaction	Continuous/Interval
last_evaluation	Evaluation of employee performance	Continuous/Interval
number_project	Number of projects completed while at work	Ratio
average_montly_hours	Average monthly hours at workplace	Ratio
time_spend_company	Number of years spent in the company	Ratio
Work_accident	Whether the employee had a workplace accident	Binary
left	Whether the employee left the workplace or not	Binary
promotion_last_5years	Whether the employee was promoted in the last five years	Binary
sales	Department in which they work for	Nominal
salary	Relative level of salary	Ordinal

vii. *Transform data:*

- We need to convert categorical variables(Sales) to numerical
- Rename variables to meaningful names
- Add a unique identifier for each employee

viii. *Segment:*

- Filter the employees based on the definition of valuable employee

3. ANALYSIS:

- Summarize data - We will perform a descriptive analysis on the data and check for outliers.
- Relationships - We will study correlations between variables
- Regularization - We will perform feature selection using techniques like Lasso
- Grouping Data - We will use Clustering to group employees based on their reason for premature exit
- Find Pattern - Finding hidden patterns from the data for better analysis
- Predicting Model - Implementation of Classification models (Logistic Regression, Decision Trees, Random Forest, KNN, Gradient boosting). Compare them and choose the best among these

4. DEPLOYMENT:

- Report
- Predicting model