University of New Haven
(Tagliatela College of Engineering & Technology)
CSCI 6660: Introduction to Artificial Intelligence

Term Project:

HR ANALYTICS

Guided By:- **Dr. Shivanjali Khare**

Group Members:
RIDDHI JOSHI
SAITEJA GUNDA

Project Topic:-

1. The main Objective of this project is to predict which employee is looking for job change
2. To predict the percentage of candidates leaving job and anticipate the pool required to fill these positions
3. Also, to analyze where HR team needs to improve candidate experience in organization to decrease the attrition rate of the company

Inspiration

- Predict the probability of a candidate will work for the company
- Interpret model(s) such a way that illustrate which features affect candidate decision

Statement of Project Objectives:-

- A large company named XYZ, employs, at any given point of time, around x employees.

- However, every year, around ~15% of its employees leave the company and need to be replaced with the talent pool available in the job market.

- The management believes that this level of attrition (employees leaving, either on their own or because they got fired) is bad for the company, because of the several reasons.

Statement of Project Objectives:-

- A sizeable department has to be maintained, for the purposes of recruiting new talent
- The former employees' projects get delayed, which makes it difficult to meet timelines, resulting in a reputation loss among consumers and partners
- More often than not, the new employees have to be trained for the job and/or given time to acclimatize themselves to the company
- Hence, the management has contracted an HR analytics firm to understand what factors they should focus on, in order to curb attrition.

Approach (Tools & Techniques):-

- To fulfill the project's core goal of space writing, we must use the following technique:

  - ✓ Performed classification using 4 models Logistic Regression
  - ✓ SVM
  - ✓ Random Forest
  - ✓ XGBoost algorithm

Deliverable (Tools & Techniques):-

- We need to predict the target variable which is categorical and implement supervised learning models to determine the probability of a candidate to look for a new job or will work for the company as well as interpreting affected factors on employee decision

- We need to deal with an imbalanced data and most features are categorical(nominal, ordinary, binary) with some high cardinality.

- In order to handle this kind of data we do feature engineering.

- We will use scikit learn metrics such as confusion matrix, classification matrix, accuracy, precision, recall to evaluate a model

Evaluation Methodology:-

- We will use the Python for feature engineering/pre-processing the data and several libraries like matplotlib, scikit-learn to analyze the data and built the model.

- In order to attain the better accuracy of our model and results, we need to deal with the missing values and outliers within our data before fitting the model.

- We have implemented different ML models like Logistic regression, Decision tree, Support Vector Machine, XGBoost and also hyper tuned the parameters for better results.

Thank You !