Project: -Resume Screening using k-Nearest Classifier.

Problem:

Hiring the right talent is a challenge for all businesses. This challenge is magnified by the high volume of applicants if the business is labor-intensive, growing, and facing high attrition rates. Typically, large companies do not have enough time to open each CV, so they use machine learning algorithms for the Resume Screening task.

Outcome expectations:

The main expectation of this project will be **classifying the resumes based on similarities and how closely they are connected and categorize into same groups of categories**. And visualize the data and analyze it and take hiring decisions.

Steps (with screenshots):

1. Data cleaning and convert the key words into categorical values: this is the challenging part where I was supposed to clean data in order to analyze it to take hiring decisions.

* Removing URLs, Hashtags, Special Letters, and punctuation:

A picture containing table

Description automatically generated

1. Model Training: Accuracy: 0.99 (Both on training and testing dataset):

* 0.99 on both training and testing dataset

Table

Description automatically generated

1. Data visualization:

Chart, pie chart

Description automatically generated

Tools:

* Python notebook