Project Proposal: Impact Analysis of Mass Layoff in Various Domains Using Machine Learning

By Mira Radhakrishnan [MXR220049]

Introduction:

The COVID-19 pandemic has caused significant disruptions in the global economy, leading to many companies implementing cost-cutting measures, including mass layoffs. The tech industry has been pushed to meet these challenges, with many notable companies announcing significant layoffs in recent years. Understanding the factors that contribute to mass layoffs in the tech industry can provide valuable insights for both companies and policymakers.

Objective:

The objective of this project is to analyze the layoffs dataset and identify potential factors that contribute to mass layoffs in the tech industry. The project will aim to study the pattern and relationship between company financial performance and layoffs., Reasons cited to layoff for better understanding of the company's situation and pattern among the reasons cited along with trends formed with specific to geographic location.

Dataset Description:

The dataset used in this project is provided by Kaggle and contains information on Tech companies lay off details. The dataset comprises 2430 observations with 9 explanatory variables, including numerical, categorical, and binary variables. The target variable is a numerical variable that indicates the future prediction of like how much percentage can be layed off to increase the fund and can further study the domain of maximum and minimum impact using clustering.

COMPAY	Name of the Company
LOCATION_HQ	Location of the headquarters
INDUSTRY	Domain or industry
LAID_OFF_COUNT	No of employees laid off
DATE	Date of lay off
SOURCE	Source of Data
FUNDS_ RAISED	Fund increased
STAGE	Stage of company
DATE_ADDED	Date information added
COUNTRY	Country name

PERCENTAGE	Percentage of impact
LIST_OF_EMP_LAYEDOFF	List source of employees laid off

The project wLiill use exploratory data analysis, statistical analysis, and machine learning techniques to analyze the layoffs dataset. The dataset will be pre-processed, cleaned, and formatted to ensure that it is suitable for analysis. The project will also use data visualization techniques to communicate the results of the analysis effectively. Statistical models, such as regression analysis, clustering, and classification, will be used to identify patterns and relationships in the data.

Methodology:

The project will use exploratory data analysis techniques to analyze the layoffs dataset. The dataset will be cleaned and pre-processed to ensure that it is suitable for analysis. Statistical and machine learning techniques will be used to identify patterns and relationships in the data. The project will also incorporate data visualization techniques to communicate the results of the analysis effectively.

Below are the stages that are planned to build this project.

- Pre processing the data
- Exploratory data analysis
- Feature selection
- Model selection
- Model evaluation

Conclusion

The project will provide valuable insights into the factors that contribute to mass layoffs in the tech industry. Along with trends followed and future prediction of layoff and its impact incase of economic shortage.

Purpose of choosing the dataset:

When I was wondering about the layoff news everywhere in LinkedIn and other sources popped up the messages with counts of lay off where the whole world is being impacted now. I got my interest to study and analyze the trend behind this. So I have decided to use my machine learning knowledge to bring out some insights to the real time scenario.

My main aim is to learn deeper on machine learning using real time scenarios and I found this to be one among that so that I can just get information on it.

Source : https://www.kaggle.com/datasets/theakhilb/layoffs-data-2022