

Employee Layoff Prediction Model using LightGBM Regressor

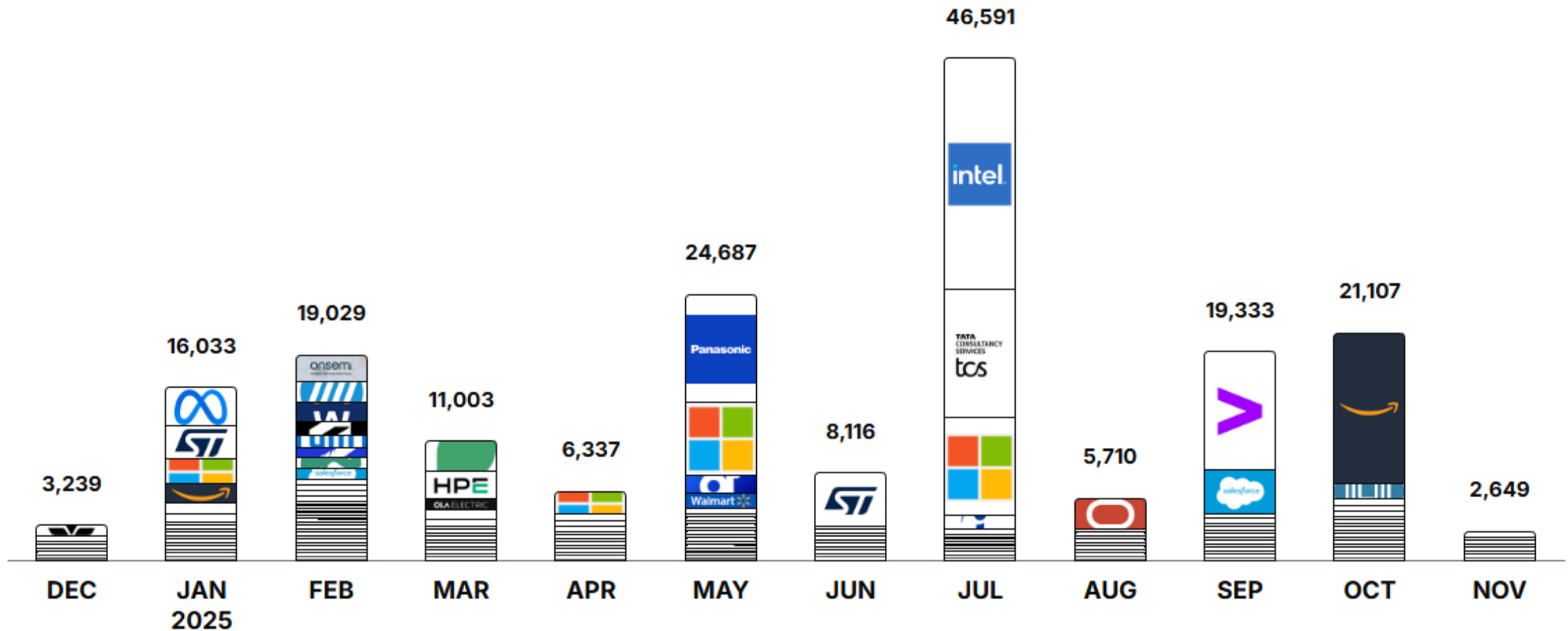
- **I have created a Machine Learning Approach for the Workforce Analytics**

Sudharshan S

GRADTWIN

Problem Statement:

- In the rapidly evolving tech industry, unexpected workforce reductions have become increasingly common, affecting thousands of employees and creating uncertainty in the job market. Between 2020 and 2024, over 500,000 tech employees worldwide faced layoffs, with limited predictive tools available to anticipate these events.
- The challenge is to develop a machine learning model that can accurately predict the number of employee layoffs in tech companies based on organizational characteristics, funding status, geographic location, and temporal patterns, enabling proactive workforce planning and risk mitigation.

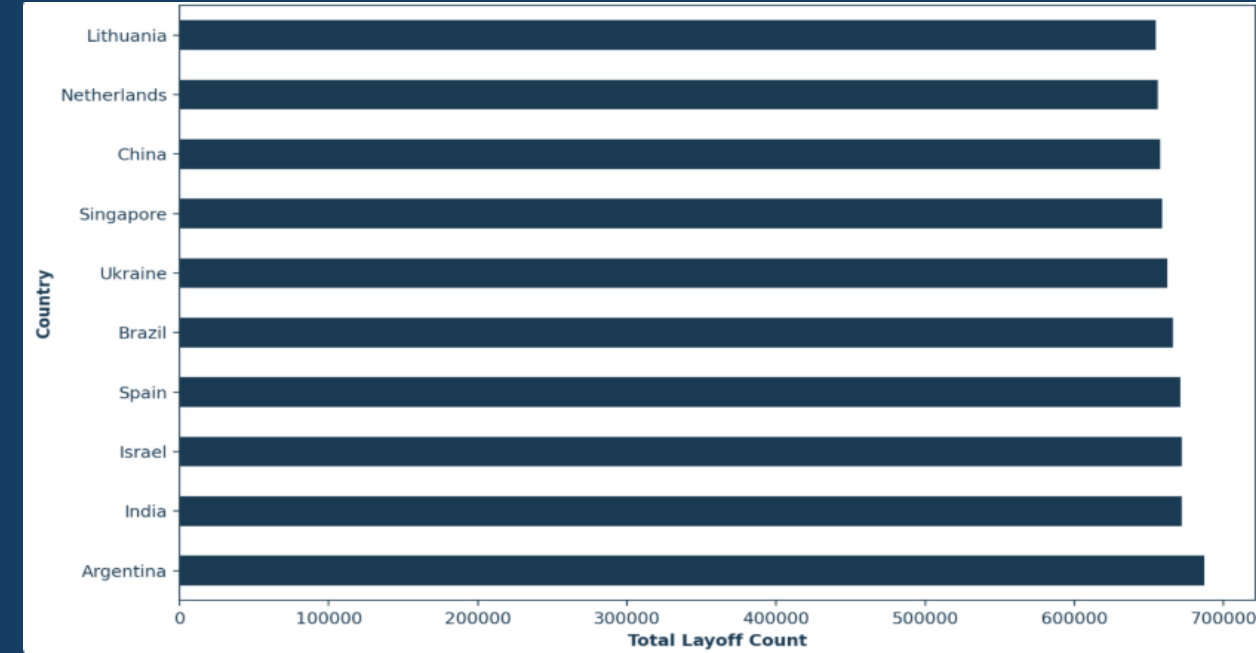
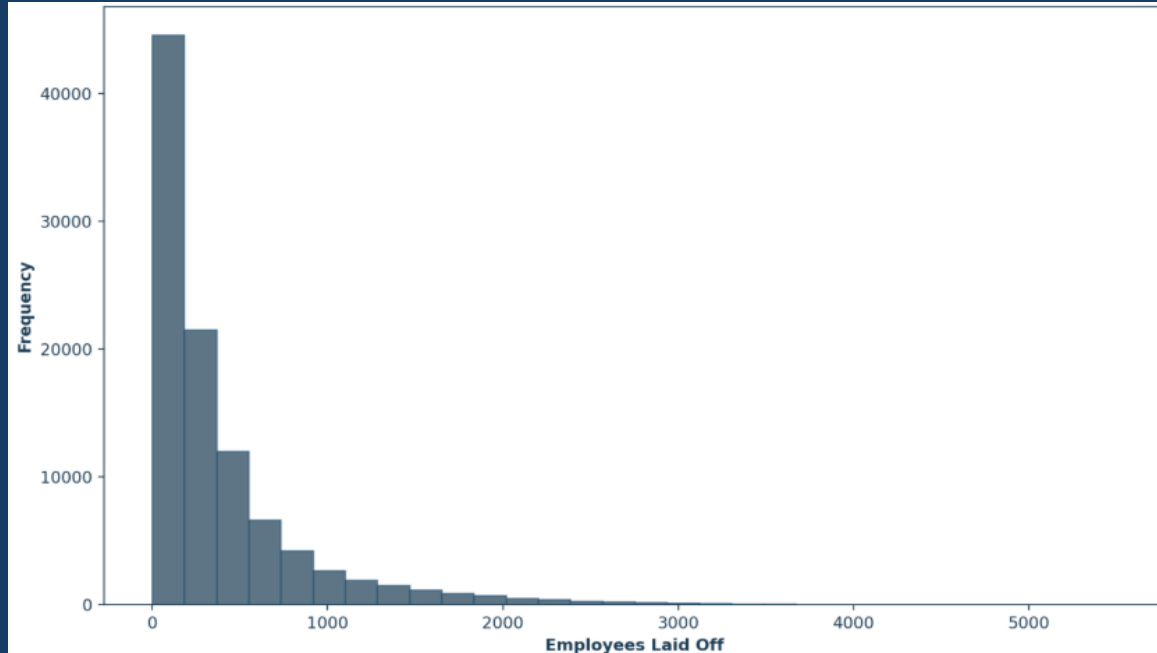
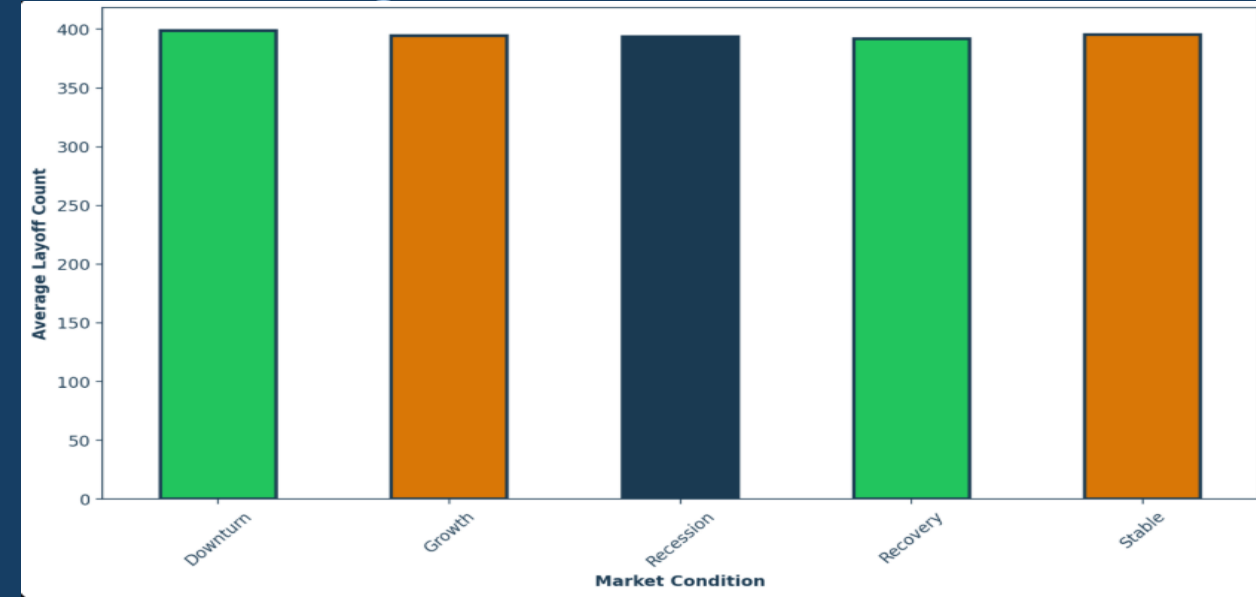
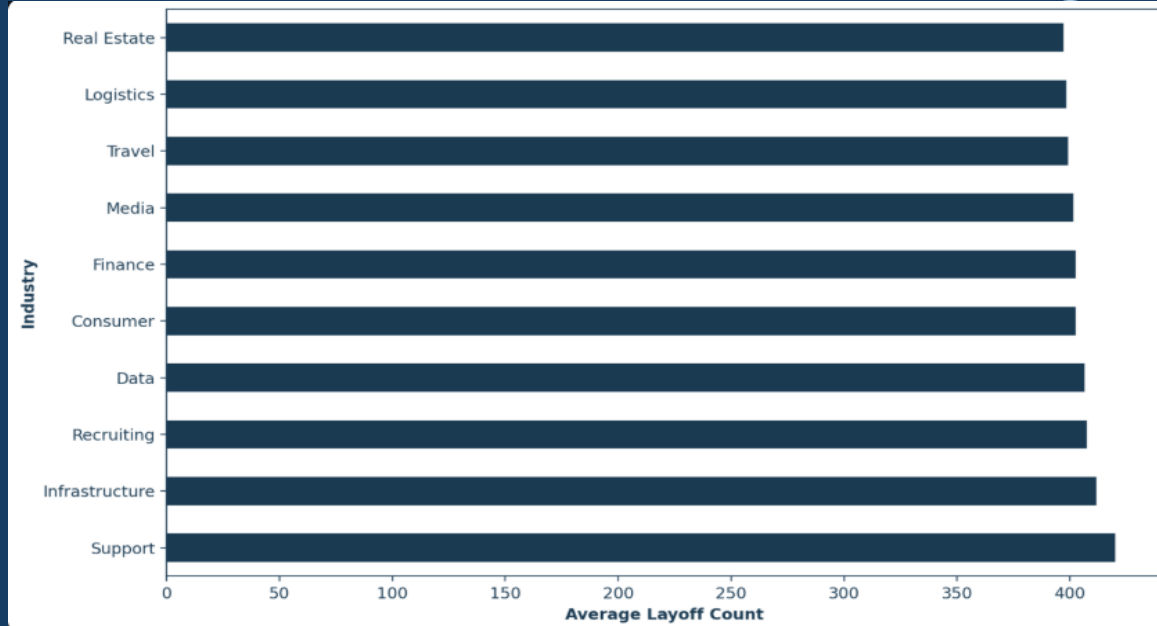


- The above is a Bar graphical representation of **LAYOFFS** done by some top tech companies in 2025.
- That also includes the largest layoff done by **TCS** i.e, layoff of **2%** in their total workforce this year(approx. **12,000 employees**).

Dataset:

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	
1	Company	Location	Industry	Laid_Off_Count	Date	Funds_Raised	Stage	Country	Percentage	Company_Value	Remaining_Value	Year_Founded	Company_Status	Reason_Closed	Quarter	Year	Month	Layoff_Severity	Notes
2	Arch Onco	Karachi	Consumer	10	#####	10.5	Series A	Pakistan	0.18	56	46	2019	2	Acquisition	Q2	2021	4	Medium	Product
3	DigitalVent	Moscow	Hardware	75	#####	300	Series E	Russia	0.09	838	763	2015	8	Cost Cutting	Q2	2023	6	Low	Product
4	SmartSolut	Nairobi	Travel	849	#####	1327.4	Subsidiary	Kenya	0.2	4228	3379	1994	29	Market Cor	Q4	2023	11	Medium	Service
5	DataGroup	Bangkok	Energy	653	#####	722.1	Subsidiary	Thailand	0.37	1786	1133	1999	24	Restructur	Q1	2023	3	High	Product
6	PerkSpot	Accra	Education	1635	#####	2167.8	Post-IPO	Ghana	0.25	6645	5010	1977	47	Cost Cutting	Q2	2024	5	Medium	Product
7	Kenko Hea	Budapest	Data	5	#####	3.1	Seed	Hungary	0.14	41	36	2021	1	Economic	Q4	2022	10	Medium	Product
8	AIVentures	Cape Town	Fitness	1173	#####	1031.8	Private Equ	South Africa	0.15	7949	6776	1992	31	Market Cor	Q4	2023	10	Medium	Service
9	TechInc	Hanoi	AI	84	#####	112.7	Series D	Vietnam	0.19	445	361	2018	5	Bankruptcy	Q3	2023	7	Medium	Product
10	CloudLabs	Dubai	Healthcare	749	#####	1421.9	Series I	United Arab	0.36	2072	1323	2007	15	Economic	Q1	2022	2	High	Service
11	ApexSoluti	Manila	Consumer	9	#####	8.2	Series A	Philippines	0.16	61	52	2020	2	Cost Cutting	Q2	2022	5	Medium	Product
12	ExtraHop	Barcelona	Media	389	#####	517.1	Series F	Spain	0.3	1295	906	2006	14	Failed Fun	Q3	2020	7	High	Product
13	DataTechn	Nairobi	Finance	173	#####	656.8	Series G	Kenya	0.19	914	741	2009	11	Acquisition	Q2	2020	6	Medium	Product
14	TechLabs	Jakarta	Travel	82	#####	356.9	Series E	Indonesia	0.09	927	845	2012	9	Product Pi	Q2	2021	6	Low	Product
15	Meta	Gothenburg	Aerospace	803	#####	1187.8	Series I	Sweden	0.32	2497	1694	2007	17	Acquisition	Q4	2024	12	High	Product
16	AIVentures	Kuala Lumpur	Security	49	#####	34	Series B	Malaysia	0.24	207	158	2014	6	Economic	Q1	2020	2	Medium	Service
17	QuantumL	Copenhagen	Sales	218	#####	686.9	Series G	Denmark	0.14	1593	1375	2008	12	Market Cor	Q3	2020	9	Medium	Product
18	Candy Digi	Istanbul	Support	693	#####	1255.9	Series J	Turkey	0.24	2927	2234	2009	15	Failed Fun	Q1	2024	2	Medium	Service
19	Flux Syster	Raleigh	Constructi	131	#####	325.9	Series F	United States	0.16	820	689	2011	11	Restructur	Q3	2022	9	Medium	Product
20	The Modist	Wellington	Retail	669	#####	598.3	Acquired	New Zealand	0.22	3091	2422	2004	18	Economic	Q1	2022	2	Medium	Service
21	ChargePoi	Prague	Food	52	#####	75.1	Series C	Czech Rep	0.23	226	174	2019	4	Restructur	Q2	2023	6	Medium	Product
22	Le Tote	Wellington	Constructi	134	#####	131.8	Series D	New Zealand	0.25	546	412	2012	8	Market Cor	Q3	2020	7	Medium	Product
23	Zeitgold	Warsaw	Recruiting	105	#####	181	Series D	Poland	0.16	647	542	2013	8	Product Pi	Q1	2021	1	Medium	Product
24	Allnnovatio	Kyiv	HR	624	#####	881.1	Acquired	Ukraine	0.16	3862	3238	2003	20	Acquisition	Q2	2023	4	Medium	Product
25	TechDynam	Toronto	Healthcare	96	#####	25.5	Series B	Canada	0.37	264	168	2018	6	Acquisition	Q3	2024	8	High	Product
26	Compass	Kitchener	Fitness	374	#####	1566.9	Series J	Canada	0.17	2189	1815	2000	22	Failed Fun	Q4	2022	10	Medium	Product
27	VTEX	Vilnius	Crypto	23	#####	29.2	Series B	Lithuania	0.15	152	129	2020	3	Cost Cutting	Q2	2023	6	Medium	Product
28	Q	London	Consulting	50	#####	88.4	Series C	United Kingdom	0.18	878	688	2018	4	Product Pi	Q1	2022	4	Medium	Service

Exploratory Data Analysis:



Feature Engineering:

New Features Added:

- Added new features like layoff_severity rate, CEO_change and more.
- Used label encoder to encode categorical values and MinMax scalar for Scaling numerical values.

...	Company	Location_HQ	Industry	Laid_Off_Count	Date	Funds_Raised	Stage	Country	Percentage	Company_Size	...	Market_Condition	CEO_Change	Previous_Layoffs	Remote_Policy	Revenue_Milli	
	0	153	106	3	10	2021-04-11	0.001903	4	39	18	0.003082	...	3	True	False	2	0.001
	1	712	149	12	75	2023-06-21	0.059896	8	45	9	0.055474	...	2	True	False	0	0.030
	2	2117	152	29	849	2023-11-12	0.265705	14	29	20	0.282594	...	4	False	False	0	0.251
	3	654	12	7	653	2023-03-26	0.144451	14	55	37	0.118987	...	2	True	True	0	0.091
	4	1750	1	6	1635	2024-05-16	0.434054	1	19	25	0.444526	...	3	False	False	2	0.320
...
99995	987	101	6	26	2024-02-12	0.001442	4	24	22	0.007303	...	4	True	False	0	0.004	
99996	136	121	18	409	2024-07-05	0.056691	0	40	15	0.178748	...	3	False	False	1	0.134	
99997	663	46	13	470	2024-02-26	0.287300	14	59	10	0.326075	...	4	False	False	0	0.174	
99998	1872	96	21	364	2024-05-04	0.380248	13	37	15	0.157309	...	3	False	False	0	0.138	
99999	688	151	12	120	2024-06-11	0.010477	15	18	10	0.079727	...	3	False	True	1	0.015	
100000 rows x 28 columns																	

```
from sklearn.preprocessing import LabelEncoder

le=LabelEncoder()
layoff["Company"] = le.fit_transform(layoff["Company"])
layoff["Location_HQ"] = le.fit_transform(layoff["Location_HQ"])
layoff["Industry"] = le.fit_transform(layoff["Industry"])
layoff["Stage"] = le.fit_transform(layoff["Stage"])
layoff["Country"] = le.fit_transform(layoff["Country"])
layoff["Reason_Category"] = le.fit_transform(layoff["Reason_Category"])
layoff["Quarter"] = le.fit_transform(layoff["Quarter"])
layoff["Layoff_Severity"] = le.fit_transform(layoff["Layoff_Severity"])
layoff["Market_Condition"] = le.fit_transform(layoff["Market_Condition"])
layoff["Remote_Policy"] = le.fit_transform(layoff["Remote_Policy"])
```

```
from sklearn.preprocessing import MinMaxScaler

scaler = MinMaxScaler()
num_df = layoff[["Funds_Raised", "Company_Size", "Year_Founded", "Company_Age", "Year", "Month", "Revenue_Millions", "Burn_Rate_Months", "Valuation_Change", "Industry_Growth_Rate"]]
layoff[["Funds_Raised", "Company_Size", "Year_Founded", "Company_Age", "Year", "Month", "Revenue_Millions", "Burn_Rate_Months", "Valuation_Change", "Industry_Growth_Rate"]] = scaler.fit_transform(num_df)
```

Feature Importance Analysis:

Top 10 Most Important Features:

- Revenue_Millions (23.8%)
- Burn_Rate_Months (21.3%)
- Industry Growth Rate (16.2%)
- Company_Size (8.5%)
- Market_Condition (6.9%)
- Stage (5.7%)
- Company_Age (4.4%)
- Quarter (3.6%)
- Industry (3.1%)
- Country (1.8%)

Based on the above analysis, it gives us that
REVENUE MILLIONS plays the vital role in the prediction.

Model Building, Training & Evaluation:

Gained Evaluations:

- **R2** => 0.8005385977946707
- **MSE** => 53666.94086255893
- **RMSE** => 231.66126318950893

```
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
import seaborn as sns
from sklearn.model_selection import train_test_split
from sklearn.ensemble import RandomForestRegressor, GradientBoostingRegressor
from sklearn.metrics import r2_score, mean_squared_error, root_mean_squared_error
from sklearn.model_selection import GridSearchCV
import pickle
from lightgbm import LGBMRegressor

layoff = pd.read_csv("Layoff Dataset III.csv")

layoff["Date"] = pd.to_datetime(layoff["Date"])
layoff["Percentage"] = layoff["Percentage"]*100
layoff["Percentage"] = layoff["Percentage"].astype(int)

from sklearn.preprocessing import LabelEncoder

le=LabelEncoder()
layoff["Company"] = le.fit_transform(layoff["Company"])
layoff["Location_HQ"] = le.fit_transform(layoff["Location_HQ"])
layoff["Industry"] = le.fit_transform(layoff["Industry"])
layoff["Stage"] = le.fit_transform(layoff["Stage"])
layoff["Country"] = le.fit_transform(layoff["Country"])
layoff["Reason_Category"] = le.fit_transform(layoff["Reason_Category"])
layoff["Quarter"] = le.fit_transform(layoff["Quarter"])
layoff["Layoff_Severity"] = le.fit_transform(layoff["Layoff_Severity"])
layoff["Market_Condition"] = le.fit_transform(layoff["Market_Condition"])
layoff["Remote_Policy"] = le.fit_transform(layoff["Remote_Policy"])
```

```
from sklearn.preprocessing import MinMaxScaler

scaler = MinMaxScaler()

num_df = layoff[["Funds_Raised", "Company_Size", "Year_Founded", "Company_Age", "Year", "Month", "Revenue", "Revenue_Millions"]]
layoff = scaler.fit_transform(num_df)

x = layoff[["Company", "Location_HQ", "Industry", "Stage", "Country", "Company_Size", "Company_Age", "Quarter"]]
y = layoff["Laid_Off_Count"]

x_train, x_test, y_train, y_test = train_test_split(x, y, test_size=0.2, random_state=42)

LGBM = LGBMRegressor(
    n_estimators=500,
    learning_rate=0.05,
    random_state=42
)

LGBM.fit(x_train, y_train)

y_pred = LGBM.predict(x_test)

r2 = r2_score(y_test, y_pred)
mse = mean_squared_error(y_test, y_pred)
rmse = np.sqrt(mse)

print("R2:", r2)
print("Mean Squared Error:", mse)
print("Root Mean Squared Error:", rmse)
```


Model Evaluation Metrics:

➤ The evaluation metrics obtained from the Random Forest Regressor model is listed below:

Metric	Value	Interpretation
R2 Score	0.8005385977946707	80% Variance explained.
MSE	53666.94086255893	Mean Squared Error.
RMSE	231.66126318950893	Root Mean Squared Error.

Employee Layoff Prediction Model:

RUNNING... Stop Deploy

 **Employee Layoff Predictor**

AI-Powered Forecasting • Market Intelligence • Risk Assessment

Prediction

Analytics

Features

Visualizations

Help

Enter Company Details

Quick Summary

Company

E Inc.

Industry

AI

Stage

Acquired

Country

Argentina

Market Condition

Downturn

Remote Policy

Remote

Company: E Inc.

Industry: AI

Stage: Acquired

Size: 1,237 employees

Revenue: \$329M

Market: Downturn

Deploy

Company Size

1237

Company Age

11

Location HQ

1

Funds Raised (\$M)

427.90

Revenue (\$M)

329.00

Burn Rate

13

Industry Growth (%)

-0.01

Year

2024

Month

6

Predict Layoffs

Employee Layoff Prediction System

Powered by LightGBM • AI-Driven Risk Analysis • Advanced ML Forecasting

Disclaimer: For informational purposes. Predictions are estimates based on historical patterns.



Thank You!