VIT

Vellore Institute of Technology

(Deemed to be University under section 3 of UGC Act, 1956)

School of Computer Science and Engineering

J Component report

Programme : B.Tech

Course Title : Essentials of Data Analytics

Course Code : CSE3506

Slot : G1/G2

Title: Employee Layoff Analytics

Team Members: SUMEGH GONUGADE (20BE1100)

NITHISH KANNA G (20BCE1304)

I NILAVAN (20BCE1080)

J MIZBA (20BCE1004)

Faculty: Dr. Vergin M

Sign:

Date: 12-4-2023

Employee Layoff Analytics:

I Nilavan, Sumegh Sadashiv Gonugade, G Nithish Kanna and J Mizba

Vellore Institute of Technology, Chennai, India

ABSTRACT:

This project focuses on Employee layoff data analytics which is the examination of data related to workforce reductions to gain insights into the reasons, effects, and outcomes of layoffs. The analytics includes data on employee demographics, job performance, organizational structure, financial performance, and workforce trends. The purpose of this analytics is to understand the factors that contribute to layoffs and to evaluate the impact of layoffs on the organization and its employees. The results of the analytics can inform decisions on layoffs, help to minimize the impact of workforce reductions, and provide a basis for planning for future workforce needs. The methods used in employee layoff data analytics includes statistical analysis, machine learning models for prediction and data visualization. This work will help both the freshers as well as the employees who are associated with the specific company.

KEYWORDS:

HR Analytics, Machine Learning Models, Data Visualization, Performance Evaluation

1. INTRODUCTION:

The involuntary removal of employees from work initiated by the employers for economic and financial reasons is called a layoff. Recent times has seen unprecedented growth in mass layoffs in firms belonging to different fields. Employees have lost jobs overnight and often their only source of income. This is a pressing issue and with serious repercussions. However, it is important to understand the pattern of employee layoffs to gain insights into the reasons and analyse the situations leading to the same. To do so, the key attributes of an employee like their personal and professional details can be assessed and modelled. This can be done quantitatively and qualitatively using various data analytics and machine learning techniques. Personal characteristics include age, gender while professional factors can include experience, domain of expertise, salary expectations for instance. The contribution of each element, if any, can be used to benefit both the employees and firms in preventing the situation from occurring.

Over tens of thousands of tech employees have been laid off across the globe. As the world emerges from the pandemic era, over-hiring, cost pressures and funding challenges have posed fresh problems for companies. The layoffs which started last year has only increased since the beginning of 2023.

By offering useful insights into an organization's financial and operational performance, data analytics can play a significant role in deciding when to make layoffs. Organizations may need to make unpleasant choices, like decreasing their staff through layoffs, in times of financial difficulty or downturns. Yet, carrying out layoffs without conducting adequate data analysis might have unforeseen implications that are detrimental to the firm and its workers. Organizations can use data analytics to identify the teams, individuals, or departments that are least helpful in achieving their goals and objectives and can spot places where production is lagging or where there is redundant staffing by monitoring employee performance metrics. Making decisions on which positions or roles are most suitable for layoffs can be done with the help of this data. Organizations can identify the optimum level of workforce reduction without jeopardising the company's financial stability or operational capabilities by studying the relationship between employment levels and key performance metrics like revenue and profit margins. Data analytics can also assist businesses in finding alternatives to layoffs, such as reassigning staff to different divisions or providing voluntary separation packages. Organizations can identify which employees are the most adaptable and which positions they can be reassigned to by examining employee skills, experience, and performance data. Organizations may find data analytics to be a crucial tool for making thoughtful decisions about layoffs. Organizations can lessen the detrimental effects of layoffs on both its employees and the whole business by employing data analysis to identify which parts of the organisation are the least productive, forecast the potential impact of layoffs, and identify alternative solutions.

By pointing out areas where employees may enhance their performance and add more value to the company, data analytics can also play a part in preventing employee layoffs. Employees can seek to improve their skills and expertise, making them more valuable to the firm and less likely to be targeted for layoffs, by assessing employee performance data and finding areas for improvement. Machine learning algorithms can also be used to find employees that are working hard and have a history of supporting the mission and goals of the company and help boost these workers' enthusiasm and engagement by praising and rewarding them, which may then result in increased output and help the business's bottom line.

As a whole, both the businesses that implement them and the people who are let go may suffer as a result. Layoffs can result in lower job satisfaction, more stress and anxiety, as well as lower productivity and motivation among surviving employees. They can harm the company's brand, reduce consumer loyalty, and raise the costs of hiring and training new staff. Furthermore, layoffs may have a detrimental effect on the neighbourhood economy and the areas where the impacted employees reside. Hence, considering the strong effects of such a situation, our project is relevant and required for betterment of the society.

2. LITERATURE SURVEY:

In order to get a detailed view on Layoffs, its causes, after effects on employees and company, a wide range instances from different points of time were studied. The cause of layoffs can either be internal or external of the company, while internal causes can br avoided by well-defined management external causes are unexpected and can be devastating. The onset of COVID 19 was one such cause. It made all job sectors to shut down and move to virtual means, while many sectors were able to cope up, sectors like tourism and retail got affected the most. They companies in those sectors had to carry out layoffs, which affected the employees adversely [5] [20]. The main outcome every company expects from layoff is the prevention of the downfall and the proper functioning of the organization. But this is not seen in most situations. The reaction of external market can turn over the outcome of layoffs [17]. For a company the reaction of their investors matters the most. And it's observed that the initial reaction of investors is against the decision of layoffs, but with time investors are convinced about the decision [16]. The markets also respond positively to layoff in certain instances [14], because layoff removes detached workers and reduces financial stress on company [6]. This outcome of the external factors is actually influenced by how the internal factors, the employees react to layoff. The layoffs affect both the victims and the survivors. The survivors are prone to experience a feeling of job insecurity and a communicational hole develops between the employees [18]. A study on white and black workers found out that even when the employer ensured to avoid layoff the feeling of job insecurity prevailed among both the section of workers. The effects of layoff on survivors are studied via various means, and all converge to give a conclusion that layoff tips off the balance [12]. The survivors experience depression [4] and detachment from work [8]. This leads to voluntary turnover from the survivors. But it's also found that not all survivors leave the company and the level of emotional intelligence influences the voluntary turnover of survivors [3]. The layoffs also disturb the work life balance of survivors [7]. These negative impacts within the company can make the company to crumble from within. In order to avoid such happenings managers and HR should work out ways on handling the layoffs effectively. The managers and HR have to ensure that the existing employees don't get detached from company and they don't feel threatened within the company [10]. Our study revealed a instance where the frequent listening to employees by manager helps in creating a secure environment for survivors [9]. And a study was conducted to formulate an efficient model to properly announce layoffs decision to employees [11]. The impact of layoff can be very severe for the victims. The impact follows them even to their new job. The victims even after getting new job feel insecure and are prone to take voluntary turnover [19]. And for a victim who have suffered multiple layoffs, the effects can be worse. They get used to layoffs and show very low levels of job security [13]. The negative impacts of layoff can be reduced in victims with the help of social support [2]. The adverse effects of layoff induce researchers to devise mechanism to predict layoffs. The social networks and data mining techniques have been used to predict layoffs [15]. The expanded literature survey helped us in getting useful insights about layoffs. This motivated us to develop a efficient model to predict layoff, thereby helping companies and employees.

3. PROPOSED WORK:

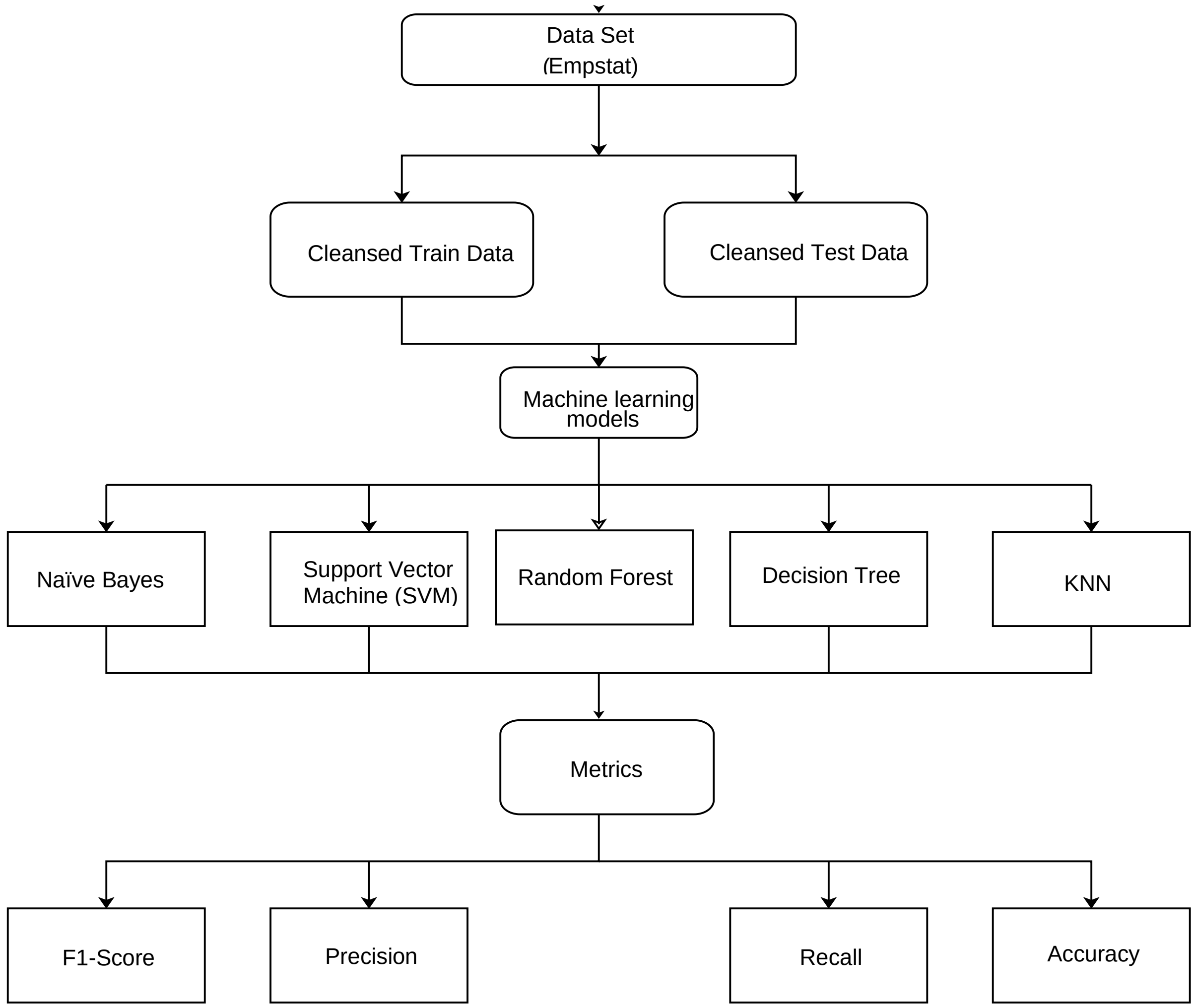

Figure 1 Proposed Framework for HR Analytics

The proposed framework as mentioned in **Figure 1** is implemented in this paper considers different statistical models to give a comparative analysis on the model accuracy. The Machine Learning algorithms implemented in this paper are- Support Vector Machine (SVM), Random Forest, K-Nearest Neighbor (KNN), Decision Tree and Naïve Bayes. Further this work includes computation of metrics (F1-score, Accuracy, Precision, Recall) for each of the algorithms to analyze the layoff patterns and compare with other models. This paper also includes visualization with different graphs in order to understand trends in Layoff and the vital features considered by the company while laying off people.

3.1. DATA SET DESCRIPTION

The dataset **'empstat'** used in this analysis was taken from a Kaggle public post / competition, 'HR Analytics Case Study, by Vijay Choudhary and Aman Kumar'. The dataset comprises of 4410 rows and 24 columns with no NULL values. The rows represent individual employees, and the columns the parameters measured. The parameters that huge impact during layoff like Job Role, Education level, Job performance, Manager review, are present in the dataset. In addition to this the dataset also has many personal, educational, technical and job related parameters of the employees.

3.2. MACHINE LEARNING METHODS

3.2.1. Decision tree

Decision trees are a supervised learning technique that can be used for both classification and regression problems, but they are mostly suitable for solving classification problems. It is a tree-structured classifier, with internal nodes representing characteristics of the dataset, branches representing decision rules, and each leaf node representing a result.

In this scenario, the target variable is Layoff Status and the predictor variables include, Age, Gender, Department, Monthly Income, Number of Companies Worked. The model gave an accuracy of 74%.

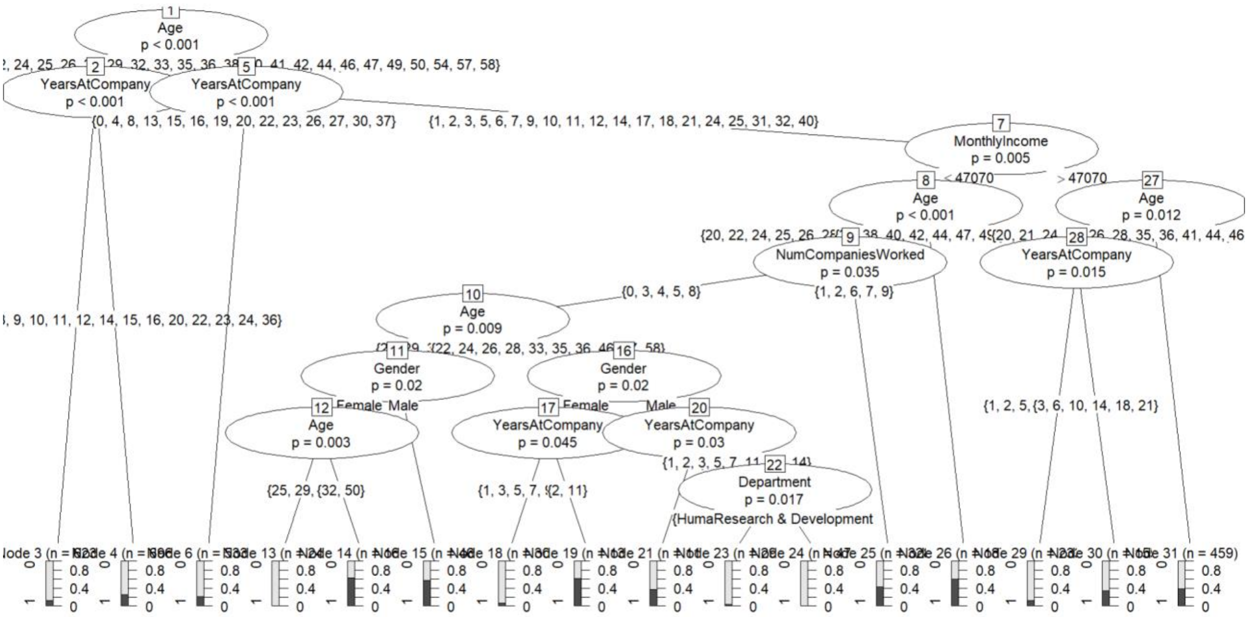

Figure 2 Decision tree for (Age, Gender, Department, Monthly Income, Number of Companies Worked) v/s Layoff Status

predict_model

	0	1
0	668	24
1	225	35

Figure 3 Confusion Matrix for the Decision Tree

3.2.2. Naïve Bayes

The Naive Bayes is an algorithm that learns the probability of every combination of each data value and its features by grouping them appropriately. It is also known as a probabilistic classifier. It comes under supervised learning and is mainly used to solve classification problems. It works on the principle of conditional probability, as given by the Bayes theorem. Given below is a representation of the Bayes Theorem:

$$P(A \mid B) = \frac{P(B \mid A) \cdot P(A)}{P(B)}$$

A, B = events

$P(A|B)$ = probability of A given B is true

$P(B|A)$ = probability of B given A is true

$P(A), P(B)$ = the independent probabilities of A and B

Figure 4 Conditional Probability Function for Naïve Bayes

In this scenario, the target variable is Layoff Status and the predictor variables include, Age, Gender, Department, Monthly Income, Number of Companies Worked. The model gave an accuracy of 72%.

y_pred	0	1
0	631	3
1	238	5

Figure 5 Confusion Matrix for the Naïve Bayes

3.2.3. KNN

The k-nearest neighbors algorithm (k-NN) is a non-parametric supervised learning method that is used for classification and regression. In k-NN classification, the output is a class membership. An object is classified by a plurality vote of its neighbors, with the object being assigned to the class most common among its k nearest neighbors (k is a positive integer, typically small). If k = 1, then the object is simply assigned to the class of that single nearest neighbor.

In this scenario, the target variable is Layoff Status and the predictor variables include, Age, Years of experience, Job Level, Monthly Income and Number of Companies Worked. The model gave an accuracy of 71% for K=3 and 75% for K=7.

3.2.4. Random Forest

Random Forest is a popular machine learning algorithm that belongs to the supervised learning technique. It can be used for both Classification and Regression problems. As the name suggests, "Random Forest is a classifier that contains a number of decision trees on various subsets of the given dataset and takes the average to improve the predictive accuracy of that dataset." Instead of relying on one decision tree, the random forest takes the prediction from each tree and based on the majority votes of predictions, and it predicts the final output.

In this scenario, the target variable is Layoff Status and the predictor variables include, Age, Gender, Department, Monthly Income, Number of Companies Worked. The model gave an accuracy of 99%.

y_pred	0	1
0	619	7
1	7	243

Figure 6 Confusion Matrix for the Random Forest

3.2.5. SVM

Support vector machines, or SVMs, are one of the most popular supervised learning algorithms used for both classification and regression problems. However, it is mainly used for machine learning classification problems. Layoff is the y label and the classificationof the specified x features. The goal of the SVM algorithm is to create optimal lines or decision boundaries that can divide n-dimensional space into classes so that new data points can be easily placed in the correct category in the future. This optimal decision boundary is called a hyperplane. SVM selects extrema/vectors to help create hyperplanes. These extreme cases are called support vectors, and the algorithm is called a support vector machine. But the model gives poop accuracy of 27%.

3.3. DATA VISUALIZATION

A qualitative overview of our problem statement has been drafted using Tableau, a visual analytics software. Key attributes were selected to highlight underlying patterns through visualizations.

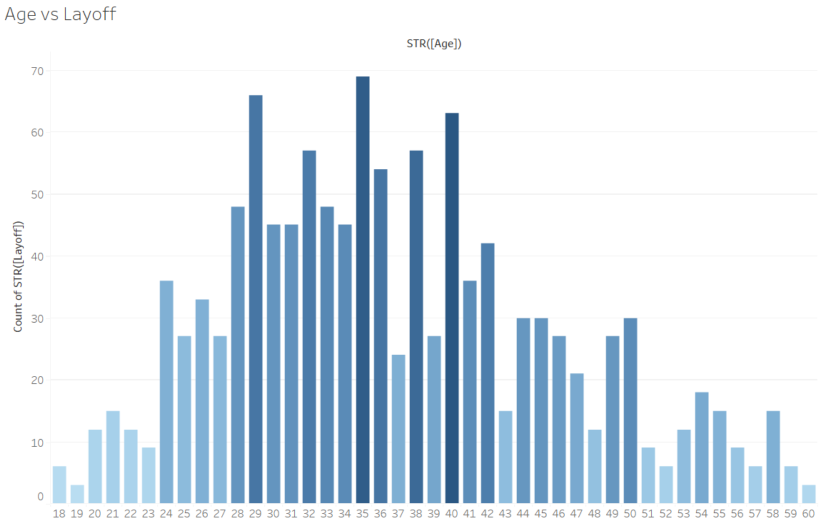

Figure 7: Age of employee vs number of layoffs

Figure 7 conveys that, employees belonging to their middle age - from the age of 28 years to 40 years - are mostly laid off. This can be explained by the following: freshers may get an edge due to three reasons – bringing new perspectives, lower pay expectation and generalized job profile. The older employees may come with several years of experience or association with the company excluding them from layoffs.

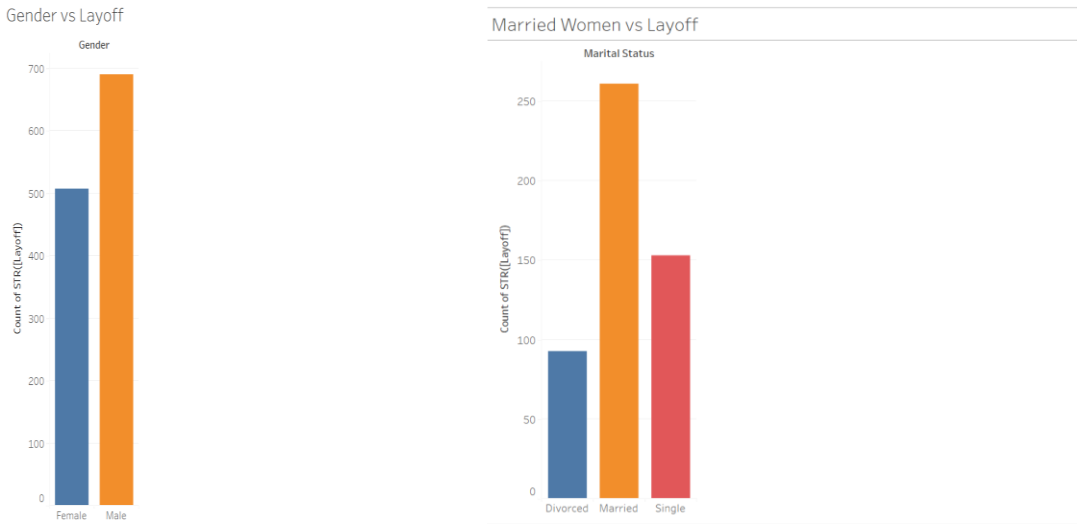

Figure 8: Gender of employee vs number of layoffs and Marital status of female employees vs number of layoffs

Figure 8 concludes that male employees are more likely to be laid off. This motivated a deeper delving into the status of female workers to understand their numbers leading to Visualization 3 that highlights that married women are more likely to be laid off a possible reason being their personal commitments and other duties.

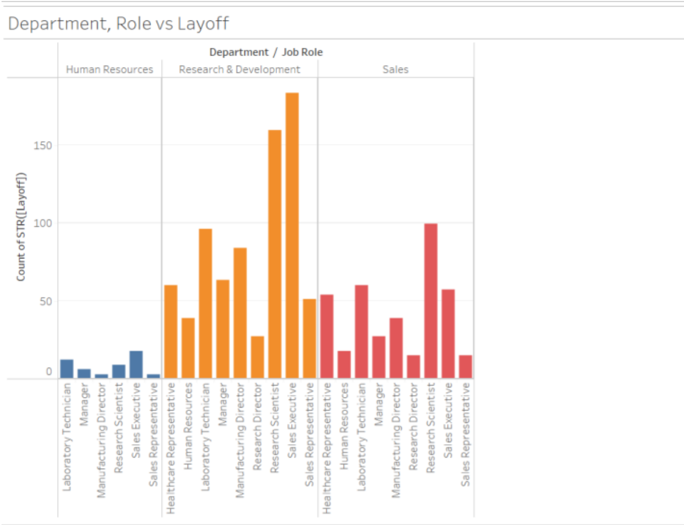

Figure 9: Department and job role of employee vs number of layoffs

Figure 9 displays a higher number of layoffs for employees belonging to the Research and Development department specifically in the Sales sub-department.

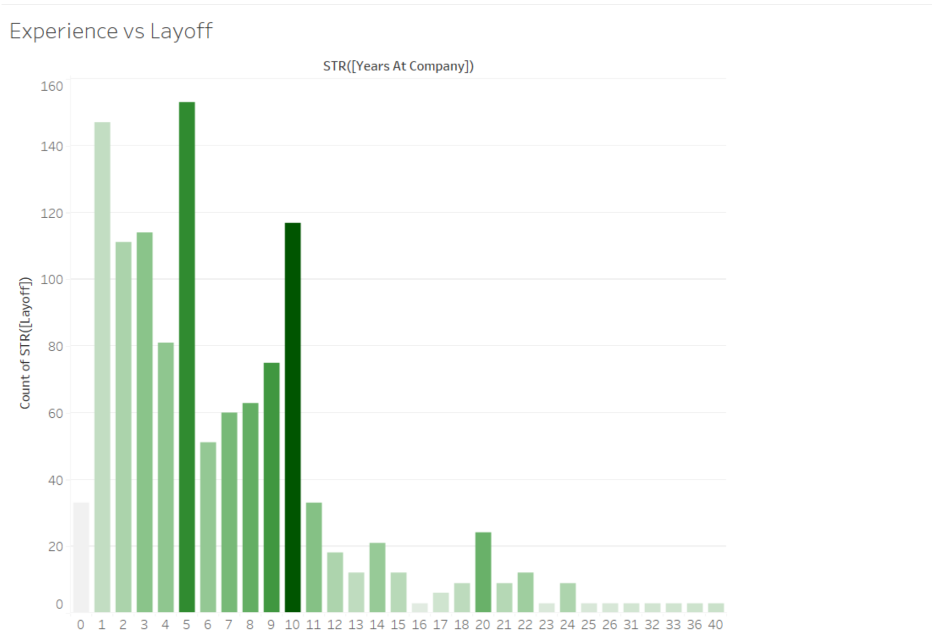

Figure 10: Experience in the field (in years) of employee vs number of layoffs

Figure 10 shed light on the relationship between employee experience and their work security. It can be noted that maximum people are laid off at 5 and 10 years of experience which is often regarded a pivot point in their career. Further, those with experience lesser than 10 years have a higher probability of being laid off than the rest of populace. This called for an extended analysis for this attribute with a subset of data.

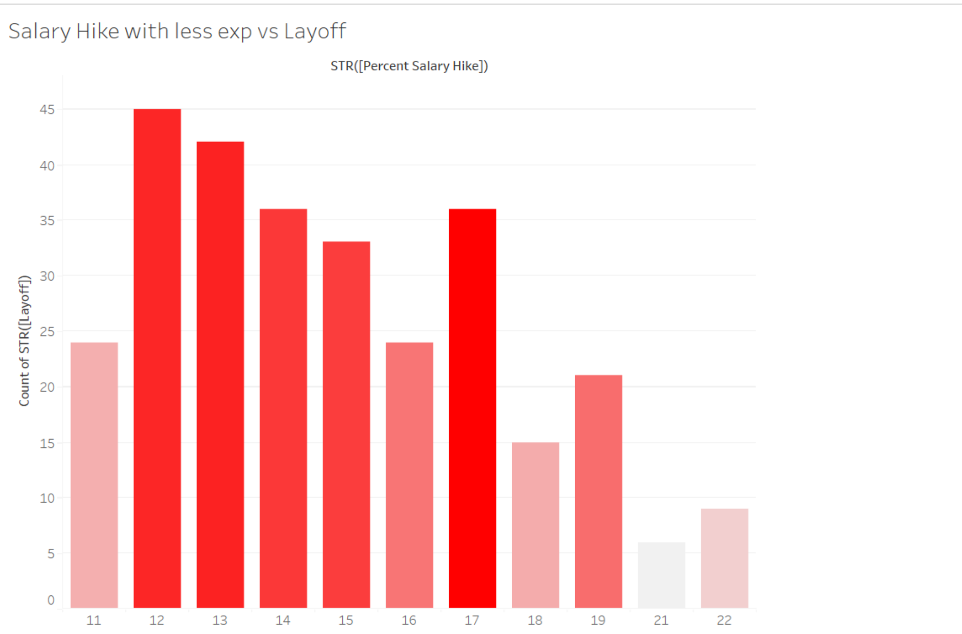

Figure 11: Demand for salary hike of employees with less than 2 years of experience vs number of layoffs

Figure 11 explores the role of salary hike percentage as a determining factor. Those with a 12-15% hike face higher chances of being laid off.

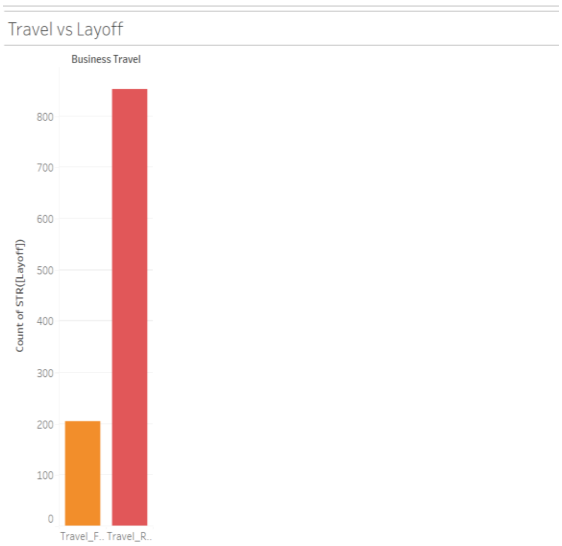

Figure 12: Work travel preferences of employee vs number of layoffs

Lastly, **Figure 12** hints that employees who travels rarely are being laid off more than 3 times of those who travels frequently.

4. RESULTS AND DISCUSSIONS:

4.1. Precision and Recall

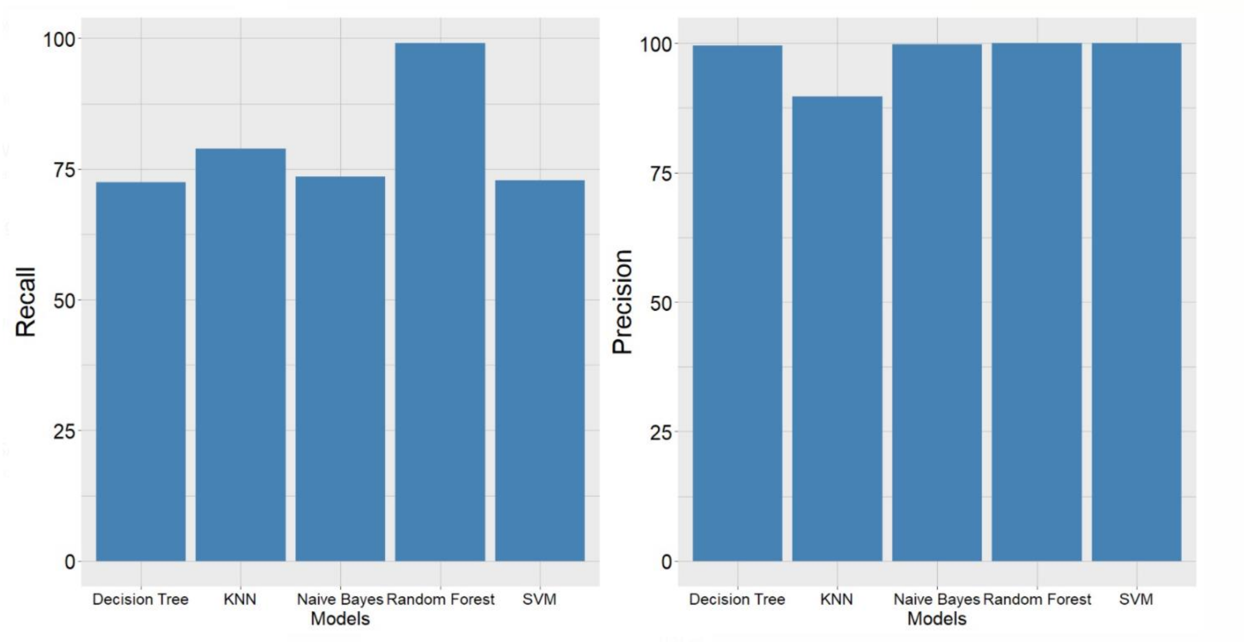

Figure 13 Precision and Recall score of various ML models used

Precision tells about how relevant the ML model classifies the data points. From the **Figure 13,** Random Forest classifier model gives the best precision score followed by other algorithms which are slightly lesser. Recall tells about how many correct positive predictions made by the model to the total number of positive predictions. In this work, Random Forest gives the best recall value.

4.2. F1 Score and Accuracy

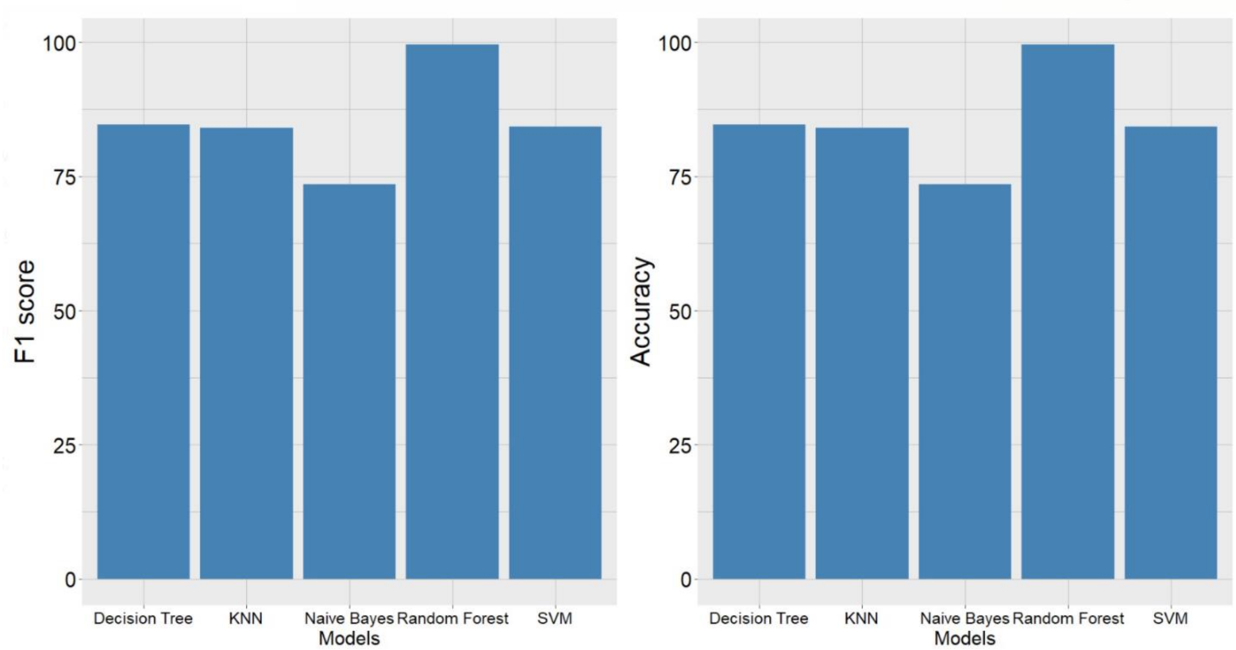

Figure 14 F1 Score and Accuracy value of various ML models used

F1 Score is the harmonic mean of Precision and Recall values. It specifies the number of correct predictions across the dataset made by the model. Accuracy denotes the number of correctly classified data points over the total number of data points. Accuracy is the most important parameter to be considered while comparing the performance of different ML models. **Figure 14** concludes that Random Forest is the best suitable data model for this dataset, which gives an accuracy of 99% and the highest F1 Score.

Table 1 Machine Learning models with its metrics

Algorithms	**Precision**	**Recall**	**F1 Score**	**Accuracy**
Naïve Bayes	99.85673	73.52321	73.52321	72
Decision Tree	99.50642	72.46585	84.69016	74
KNN	89.76035	78.9272	83.99592	75
Random Forest	100	99.05956	99.52756	99
SVM	100	72.86203	84.30079	27

Table 1 includes the combined scores of all the computed metrics for each of the ML algorithms. The results interpret that Random Forest algorithm gives the best values for all the metrics and is the most suitable model for this work.

5. CONCLUSION:

This work analyzes the impact of factors such as age, gender, years of experience, salary, and number of companies worked on an individual's layoff probability. The results suggest that these factors play a significant role in determining the likelihood of an individual being laid off, which will help the freshers also get information regarding the layoff metrics of a particular company. From the five used ML models, Random Forest has yielded good results for our dataset. Decision Tree and SVM models haven't been up to the mark in this work. To better understand the dataset, various parameters have been visualized in the form of a Bar chart, Stacked Bar chart, Line chart, and Pie chart, highlighting the impact of each factor on the layoff probability and also advanced visualization techniques such as Tableau can be used to gain better insights into the dataset.

This methodology can be used to assess their chances of being laid off. It can also be extended to other industries and sectors to analyze the impact of these factors on layoff probability. Deep Learning models can be implemented for improved performance over traditional ML models, thereby increasing interpretability.

6. REFERENCES:

[1] Yang, S. and Zheng, L., 2015. Perceived job insecurity of white and black workers: An expanded gap in organizations with layoff prevention commitment. Sociological Spectrum, 35(6), pp.483-503.

[2] Asante, E.A., Oduro, F., Danquah, B., Dei Mensah, R., Dartey-Baah, K. and Affum-Osei, E., 2022. From being sacked to being unwell: A conservation of resources view on the effects of psychological contract violation on layoff victims' wellbeing. Human Resource Management Journal.

[3] Bartock, A., 2019. A study on the relationship between emotional intelligence and employee turnover. Performance Improvement, 58(5), pp.6-21.

[4] Andreeva, E., Magnusson Hanson, L.L., Westerlund, H., Theorell, T. and Brenner, M.H., 2015. Depressive symptoms as a cause and effect of job loss in men and women: evidence in the context of organisational downsizing from the Swedish Longitudinal Occupational Survey of Health. BMC public health, 15, pp.1-11.

[5] Wedaningsih, L.S., Vipriyanti, N.U., Maba, W. and Partama, I.G.Y., 2021. Mapping the Employee Layoff of Star Hotels in Denpasar City: An Effort to Reduce the Impacts of the Covid-19 Pandemic. Soshum: Jurnal Sosial dan Humaniora, 11(1), pp.100-111.

[6] Mujtaba, B.G. and Senathip, T., 2020. Layoffs and downsizing implications for the leadership role of human resources. Journal of Service Science and Management, 13(02), p.209.

[7] Virick, M., Lilly, J.D. and Casper, W.J., 2007. Doing more with less: An analysis of work life balance among layoff survivors. Career Development International, 12(5), pp.463-480.

[8] Brockner, J., Grover, S.L. and Blonder, M.D., 1988. Predictors of survivors' job involvement following layoffs: A field study. Journal of applied psychology, 73(3), p.436.

[9] Kriz, T.D., Jolly, P.M. and Shoss, M.K., 2021. Coping with organizational layoffs: Managers' increased active listening reduces job insecurity via perceived situational control. Journal of occupational health psychology, 26(5), p.448.

[10] Korman, K. and Mujtaba, B.G., 2020. Corporate responses to COVID-19 layoffs in North America and the role of human resources departments. Reports on Global Health Research, 3(2), pp.1-17

[11] Smeltzer, L.R. and Zener, M.F., 1992. Development of a model for announcing major layoffs. Group & Organization Management, 17(4), pp.446-472.

[12] Carrington, L., 2016. A qualitative phenomenological study of employee perceptions of the impact of layoffs (Doctoral dissertation, Walden University).

[13] Moore, S., Grunberg, L. and Greenberg, E., 2004. Repeated downsizing contact: the effects of similar and dissimilar layoff experiences on work and well-being outcomes. Journal of occupational health psychology, 9(3), p.247.

[14] Chalos, P. and Chen, C.J., 2002. Employee downsizing strategies: Market reaction and post announcement financial performance. Journal of Business Finance & Accounting, 29(5-6), pp.847-870.

[15] Chang, H.T., Wu, H.J. and Ting, I.H., 2009, July. Mining organizational networks for layoff prediction model construction. In 2009 International Conference on Advances in Social Network Analysis and Mining (pp. 411-416). IEEE.

[16] Velásquez, S., Kanniainen, J., Mäkinen, S. and Valli, J., 2018. Layoff announcements and intra-day market reactions. Review of Managerial Science, 12, pp.203-228.

[17] Hillier, D., Marshall, A., McColgan, P. and Werema, S., 2007. Employee layoffs, shareholder wealth and firm performance: Evidence from the UK. Journal of Business Finance & Accounting, 34(3-4), pp.467-494.

[18] Susskind, A.M., Miller, V.D. and Johnson, J.D., 1998. Downsizing and structural holes: Their impact on layoff survivors' perceptions of organizational chaos and openness to change. Communication Research, 25(1), pp.30-65.

[19] Davis, P.R., Trevor, C.O. and Feng, J., 2015. Creating a more quit-friendly national workforce? Individual layoff history and voluntary turnover. Journal of Applied Psychology, 100(5), p.1434.

[20] Hossin, M.S., Sentosa, I. and Miah, M.S., 2020. The impact of COVID-19 outbreak on human resource operation: Empirical evidence from the perspective of Malaysian retail employees in Klang Valley. International Journal of Organizational Leadership, 9(4), pp.304-320.