

A Study on the Layoff Problem After COVID-19 Based on Decision Tree Ensemble Methods

Yaxuan Wang^{1,a,*}

¹*School of Business Administration, South China University of Technology, Guangdong, 510641, China*

a. 202130380858@mail.scut.edu.cn

**corresponding author*

Abstract: The COVID-19 pandemic, a cataclysmic event in modern history, has upended the global economy, precipitating extensive layoffs across diverse sectors. This study delves into the determinants of these workforce reductions through the lens of decision tree ensemble methods. By scrutinizing a comprehensive dataset, we harnessed statistical and machine learning algorithms to assess the significance of key variables—including industry type, geographic location, company development stage, and capital acquisition—on the scale of layoffs. This analysis further entailed predicting the total number of employees laid off and examining their distribution. The use of decision tree ensemble methods, such as random forests and gradient boosting, provided robust insights into the complex interplay of factors influencing layoff decisions. We discovered that industry type and company development stage were particularly critical in predicting layoff patterns, while geographic location and capital acquisition also played notable roles. This research offers a data-driven perspective on the layoff phenomena, shedding light on the multifaceted influences at play and offering a foundation for further inquiry and strategic policy development in response to economic downturns. By understanding these dynamics, policymakers and business leaders can better navigate future economic crises, potentially mitigating the adverse impacts on the workforce and promoting more resilient economic structures.

Keywords: Layoff Problem, Decision Tree Ensemble Methods, Workforce Reductions, COVID-19 Pandemic.

1. Introduction

The COVID-19 pandemic has prompted an unparalleled economic crisis, with widespread layoffs emerging as a stark consequence of companies' struggle for survival. This paper seeks to provide an in-depth exploration of the patterns and predictors of post-pandemic layoffs, offering critical insights for policymakers and businesses. By meticulously analyzing a carefully selected dataset and employing a robust decision tree ensemble methodology, this study aims to identify the trends and factors influencing layoffs. The integration of quantitative data analysis with qualitative insights is intended to offer a comprehensive understanding of workforce dynamics during this critical period, guiding strategic decision-making and the development of resilience strategies in the face of economic challenges.

The existing literature on the impact of economic downturns on employment has been extensive [1, 2]. The COVID-19 pandemic presents a unique scenario that demands a fresh perspective [3, 4]. Previous studies have scrutinized the relationship between market contractions and job displacement, but the rapid and extensive job losses caused by the pandemic necessitate a reevaluation of these relationships [5, 6]. This study fills a significant gap in the literature by applying advanced data-driven methods, specifically decision tree ensemble techniques, to predict and analyze layoffs. These methods are particularly adept at handling the high dimensionality and nonlinearity often found in economic data, providing a nuanced understanding of the complex interplay of factors contributing to layoffs. The innovative approach of this study not only enhances the validity of the findings but also enriches the narrative by offering a more holistic view of the workforce dynamics during the pandemic.

The significance of this research extends beyond academic interest; it holds practical implications for strategic decision-making in both policy and business realms. By identifying the key drivers of layoffs, the study offers actionable insights that can aid organizations in building resilience against economic shocks. The findings are anticipated to contribute to the formulation of evidence-based policies and practices designed to alleviate the impact of economic downturns on employment. Moreover, the innovative application of ensemble methods to the analysis of layoffs, combined with a qualitative perspective, promises to provide a more accurate and nuanced understanding of the predictors of workforce reduction. This dual approach not only strengthens the validity of the conclusions drawn but also offers a more comprehensive view of the workforce dynamics during the pandemic. The study's findings can inform companies and governments on how to better prepare for and respond to economic crises, potentially reducing the severity and frequency of layoffs. It suggests that decision-makers should consider factors such as industry resilience, workforce diversity, and the adaptability of business models when formulating strategies. Additionally, policymakers could benefit from insights into how to structure social safety nets and economic stimulus packages to support industries and workers most vulnerable to economic downturns. The research underscores the importance of proactive measures, such as skill development programs and job transition support, to enhance the workforce's adaptability and reduce the long-term repercussions of layoffs.

2. Methodology

The dataset is from the tech layoffs reported on the platforms such as Bloomberg, San Francisco Business Times, TechCrunch and The New York Times, and the data availability is from when COVID-19 was declared as a pandemic, i.e. 11 March 2020 to present (1 June 2024). This research employed decision tree ensemble methods, including Decision Tree, Bagging, and Random Forest algorithms, to analyze the importance of different factors and predict the total number of layoffs. The dataset was preprocessed to handle missing values and categorical encoding was applied to categorical variables.

2.1. Visualization

Do some visualization and take a rough look at the relationship between `total_laid_off`, `funds_raised` and percentage from different perspectives, such as industry (e.g. Figure 1), stage (e.g. Figure 2), and country (e.g. Figure 3).

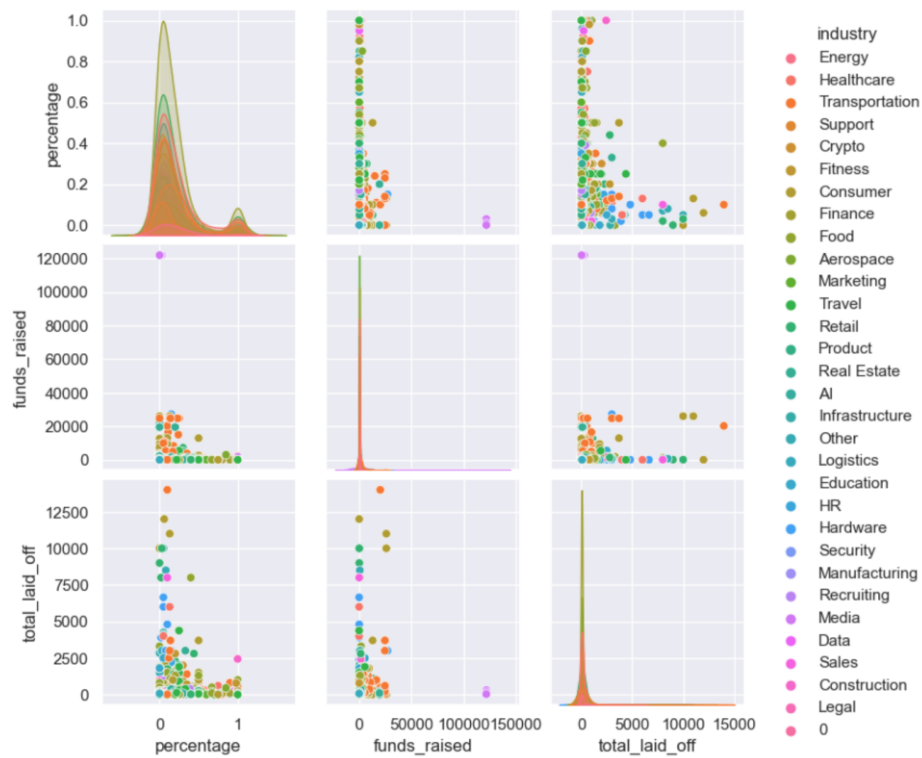


Figure 1: The relationship between total_laid_off, funds_raised and percentage across industries.
(Photo/Picture credit: Original)

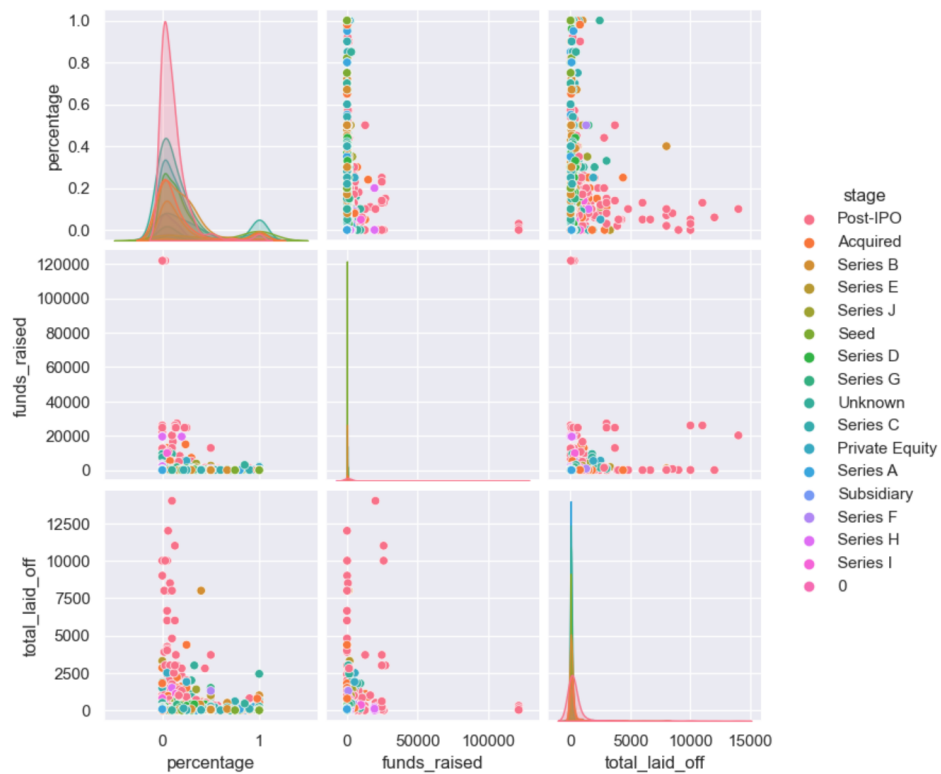


Figure 2: The relationship between total_laid_off, funds_raised and percentage across stages.
(Photo/Picture credit: Original)

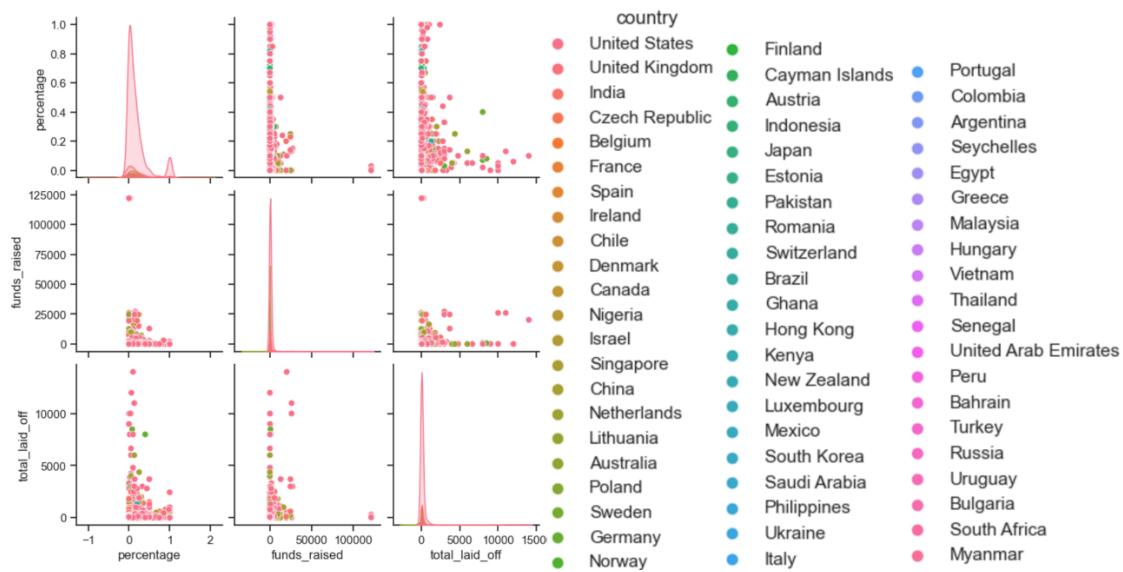


Figure 3: The relationship between total_laid_off, funds_raised and percentage across countries.
(Photo/Picture credit: Original)

2.2. Data Preprocessing

The dataset used in this study was collected from Kaggle, containing information on various companies, including their location, industry, percentage of funds raised, total laid off, stage of funding, and country. The data preprocessing steps taken include handling missing values, converting categorical data into numerical form, and feature selection.

2.3. Machine Learning

In this methodology, we harnessed Python's scikit-learn library to implement decision tree ensemble methods, specifically utilizing Bagging and Random Forest algorithms. These approaches aggregate multiple decision trees to enhance predictive accuracy and mitigate overfitting, providing a comprehensive strategy for analyzing complex datasets within our study.

2.3.1. Decision Tree

Decision tree is a flowchart-like structure where each internal node represents a feature, each branch represents a decision rule, and each leaf node represents an outcome [7]. It is used for both classification and regression by recursively partitioning the data into subsets based on feature values. The significance of decision trees lies in their ability to handle both numerical and categorical data and their ease of interpretation due to their tree-like model representation.

2.3.2. Bagging

Bagging is an ensemble technique that involves training multiple base models (usually decision trees) on different subsets of the original dataset, created by random sampling with replacement (bootstrap sampling). The final prediction is typically the average of the individual predictions (for regression) or the mode of the predictions (for classification). Bagging helps to reduce the variance of the model, leading to improved generalization and robustness against overfitting.

2.3.3. Random Forests

Random forest, an evolution of the bagging technique, introduces a sophisticated layer of randomness that sets it apart from traditional decision trees and bagging methods [8]. By training on varied bootstrap samples and incorporating a random selection of features at each split, random forests effectively mitigate issues of bias and variance. This dual mechanism not only lessens the risk of overfitting inherent in individual decision trees but also enhances the model's generalizability beyond that of bagging. Especially in high-dimensional spaces, random forests excel at uncovering intricate data relationships, offering a distinct advantage over simpler models that might struggle with complex datasets. The ability to handle non-linearities and interactions without excessive reliance on any single variable contributes to the random forest's reputation as a versatile and potent tool in the machine learning arsenal.

2.3.4. Random Forest Regression

Random forest regression, leveraging the ensemble power of multiple decision trees, offers a robust approach to forecasting continuous outcomes. This technique synthesizes the predictions of individual trees, smoothing out anomalies and enhancing overall accuracy. It excels in modeling non-linear relationships and capturing the interplay of features, making it particularly adept at handling datasets with complex underlying patterns. A significant advantage of random forest regression is its capacity to quantify the importance of each feature, providing valuable insights for feature selection and model refinement. This attribute not only aids in optimizing model performance but also deepens our understanding of the data's underlying structure, making random forest regression a preferred choice for both predictive analytics and exploratory data analysis.

3. Results and Discussion

3.1. Description of Variables

There are 8 variables, as follows.

- Total_laid_off: The number of employees laid off, reflecting the impact of downsizing.
- Funds_raised: Capital obtained from investors, indicating financial strength and growth potential.
- Stage: The development stage of a company. Common stages include early stages (Seed, Series A, Series B, etc.), growth stages (Series C, Series D, etc.), maturity stages (Post-IPO indicating the company has gone public), and being acquired (Acquired).
- Industry: The sector in which a company operates, such as technology, healthcare, finance, consumer goods, etc.
- Country: The nation of a company's headquarters..
- Location: The specific geographic location of a company, often a city.
- Company: The name of the company.
- Percentage: The proportion of the workforce laid off, showing the extent of job cuts relative to total employment.

3.2. Results

3.2.1. Accuracy of Models

The ensemble methods' high accuracy rates offer significant benefits in post-pandemic layoff analysis. The superior precision of the Random Forest model, at 91.93%, underscores its ability to capture subtle data patterns, providing nuanced insights into workforce adjustments. The Decision Tree, with

a 90.96% accuracy rate, delivers clarity in predictions, ensuring a robust framework for decision-making. Similarly, the Bagging model's 91.38% accuracy exemplifies the power of ensemble predictions, which are less prone to variance and overfitting. Collectively, these models fortify our analytical prowess, enabling a deeper comprehension of the intricate dynamics surrounding layoffs, and arming stakeholders with data-driven strategies to navigate economic uncertainties.

3.2.2.Importance of Different Features

Feature importance analysis revealed that the 'location', 'industry', 'stage', and 'country' were significant predictors of layoffs. As Table 1 and Figure 4 shown, the feature importance scores indicate that 'location' has the highest influence on layoffs, followed by 'industry', then 'stage', and finally 'country'. This suggests that the geographic area and the sector of the business are more significant predictors of layoffs than the company's development stage or the country in which it operates.

Table 1: Feature Importance.

Features	Importance
Industry	34.84%
Location	40.72%
Stage	14.32%
Country	10.12%

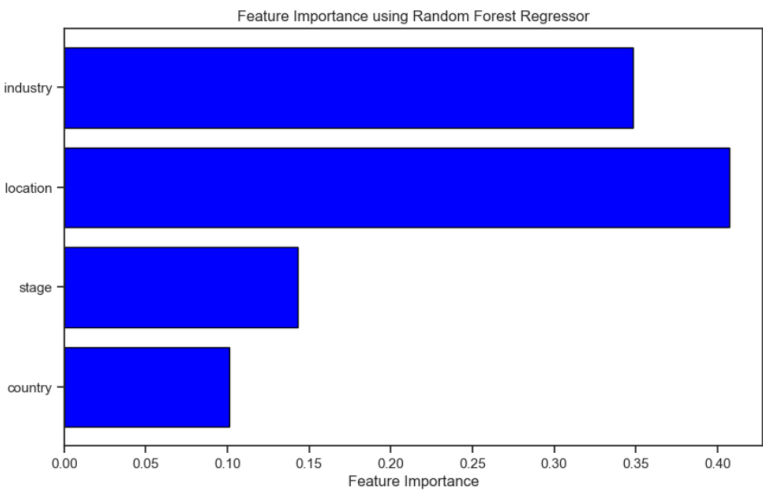


Figure 4: Feature importance.

3.2.3.Distribution of Predicted Layoffs

As Figure 5, the distribution appears to be right-skewed, with a majority of predictions falling into the lower range of layoffs, and a long tail extending to higher numbers. This suggests that while most companies are predicted to have fewer layoffs, there is a subset with significantly more, indicating potential outliers or severe cases of downsizing. Understanding this distribution is the key for organizations to allocate resources effectively and devise contingency plans to address the varied impacts of economic pressures on employment stability.

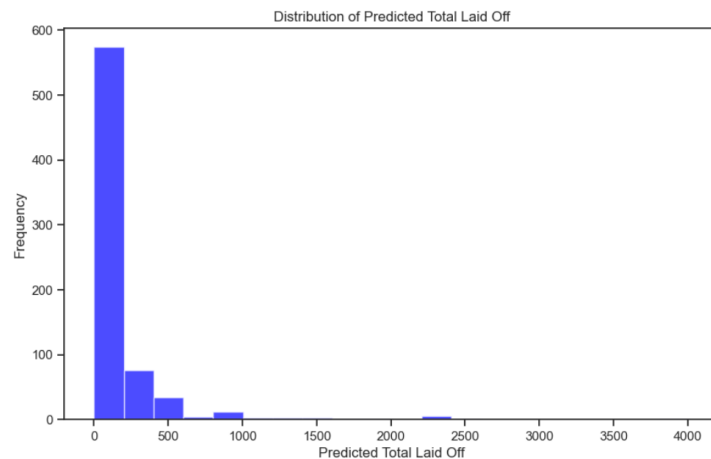


Figure 5: Distribution of predicted layoffs.

4. Considerations and Suggestions

Enterprise strategic planning should incorporate several key considerations to ensure its effectiveness. Here are the main points.

4.1. Geographic Location and Redundancy Risk

Considering the significant influence of 'location' on layoffs, enterprises must meticulously evaluate how their geographic placement affects redundancy risks. Strategic planning should integrate this assessment, enabling companies in high-risk areas to adopt agile HR strategies that can swiftly adapt to economic shifts and workforce needs [9]. This proactive approach can mitigate the adverse impacts of layoffs and enhance organizational resilience.

4.2. Industry-Specific Strategies

Different industries are affected by external factors, such as epidemics, to varying degrees. Therefore, enterprises should closely monitor industry dynamics and devise countermeasures tailored to their industry's specific characteristics [10].

4.3. Financing Stage Considerations

The financial stage of an enterprise plays a significant role in its decision-making process. Enterprises at different financing stages may encounter diverse financial pressures and market expectations. Understanding the implications of these stages on layoffs can aid in better managing capital and human resources.

4.4. International Perspective

Multinational enterprises must take into account the policies, economic conditions, and market responses of various countries to devise differentiated management strategies [11, 12]. This ensures their adaptability and responsiveness in a global context.

4.5. Data-Driven Decision-Making

Enterprises should leverage data analysis to guide their layoff decisions, ensuring transparency and rationality in the process. Data-driven approaches provide a solid foundation for informed and strategic decisions.

4.6. Role of Policymakers

Policymakers have a crucial role in addressing the economic and social impacts of layoffs. They should formulate appropriate policies to support affected enterprises and employees, minimizing the negative effects of such events.

5. Conclusion

The application of decision tree ensemble methods in this study has not only offered a robust analytical framework for understanding the complexities of post-COVID-19 layoffs but also highlighted the pivotal role of location and industry in predicting such workforce adjustments. By incorporating machine learning techniques, our research has unearthed patterns that extend the current academic discourse, providing actionable intelligence for businesses and policymakers. The insights gleaned from this analysis advocate for a more nuanced approach to managing layoffs, considering the geographic and sector-specific vulnerabilities. As we progress, these findings aim to inform evidence-based strategies, ultimately shaping a more resilient and adaptive workforce landscape in the face of global crises.

References

- [1] Takahashi, Y., Schmidt, L., Milbradt, K., et al. (2016). *Layoff risk, the welfare cost of business cycles, and monetary policy*. In 2016 Meeting Papers. Society for Economic Dynamics.
- [2] Murphy, I. K. J. (2003). *The effect of industrial diversity on state unemployment rate and per capita income*. The Annals of Regional Science.
- [3] Senci, M., Svabova, L., & Klietnik, T. (2024). *The anatomy of unemployment: Determinants during and after the COVID-19 crisis*. Management Dynamics in the Knowledge Economy, 12(1), 86–100.
- [4] Arano, K. G., & Srinivasan, A. K. (2023). *Geography of unemployment in the Midwest during the COVID-19 pandemic: A spatial analysis*. Journal of Economic Insight, 49(2), 1–36.
- [5] Zhang, Y., & Liao, J. (2006). *Analysis on the layoff theory and empirical research of foreign enterprises*. The Foreign Economy and Management, 28(10), 7.
- [6] Crane, L. D., Green, E., Harnish, M., McClennan, W., Soto, P. E., Vrankovich, B., & Williams, J. (2024). *Tracking real-time layoffs with SEC filings: A preliminary investigation*. U.S. Federal Reserve Board's Finance & Economic Discussion Series, 1–28.
- [7] Kim, K.-M., Kim, J.-H., Rhee, H.-S., & Youn, B.-Y. (2023). *Development of a prediction model for the depression level of the elderly in low-income households: Using decision trees, logistic regression, neural networks, and random forest*. Scientific Reports, 13(1), 1–10.
- [8] Park, J., Cho, J., Gang, A. C., Lee, H. W., & Pedersen, P. M. (2024). *Machine learning prediction of factors affecting Major League Baseball (MLB) game attendance: Algorithm comparisons and macroeconomic factor of unemployment*. International Journal of Sports Marketing and Sponsorship, 2, 382–395.
- [9] Sajjadi, S., Kammeyer-Mueller, J., & Benson, A. (2023). *Who is leaving and why? The dynamics of high-quality human capital outflows*. Academy of Management Journal, 66(6), 1929–1953.
- [10] Tallon, K. (2024). *Leadership tips: How to conduct a layoff the right way*. Carrier Management, 11(1), 11–14.
- [11] Snowball, J. D., & Gouws, A. (2023). *The impact of COVID-19 on the cultural and creative industries: Determinants of vulnerability and estimated recovery times*. Cultural Trends, 3, 207–230.
- [12] O'Sullivan, O., Houston, A., Ladlow, P., Barker Davies, R. M., Chamley, R., Bennett, A. N., & Holdsworth, D. A. (2023). *Factors influencing medium- and long-term occupational impact following COVID-19*. Occupational Medicine (Oxford, England).