Bellarmine University

**Technology Industry Layoffs: Data Science Capstone Project Proposal**

Kerry Effinger

Data Science 450

Dr. Robert Kelley

January 17, 2023

I. Executive Summary

In this report, layoffs by many companies in the technology industry were studied to find similarities and patterns. The results of this analysis will then be used to construct a predictive analysis to forecast how large a layoff is likely to occur based on the profile of a tech company.

II. Project Idea

With news of companies laying off employees in technology positions constantly appearing in headlines, it seems as if there is a never-ending trend of tech layoffs. To investigate the characteristics of companies engaging in these layoffs and the frequency with which they occur, data was collected on these layoffs as well as some information about the company that made the decision to let these people go. The characteristics of these companies will be studied to find patterns and to create a model that will predict whether a company in the technology industry will conduct a mass layoff and how large it will be. When creating the predictive analysis, the model can be formulated with some of the data and then tested for accuracy with the rest of the data to see how well it predicts layoffs. Finally, the results and findings will be summarized visually in a dashboard that will update as the data changes.

III. Background

The website, layoffs.fyi, is a simple website created by startup founder, Roger Lee. It has tracked companies that have laid off employees (in small numbers or very large) since the beginning of the COVID-19 pandemic in March of 2020. While the layoffs in the past several months are likely not a result of the initial shocks of the pandemic, the website still tracks the layoffs. The website only has one page, and almost the entire architecture is an interactive spreadsheet from Airtable. The spreadsheet allows users to see over 1200 entries, with each row representing a different layoff event. Some companies have several rows, as they have implemented multiple rounds of layoffs. The other columns in the table include additional information about the company, like its country, industry, date the layoff occurred, the number of employees laid off (if information is available), the source from which the information comes, the business stage of the company, and the amount of money the company raised with the layoffs. The data is updated daily, as new rows appear almost every day. Some rows, however, are entered days or weeks after the initial layoff occurs, while others are reported almost instantly. Without knowing the exact way the data is collected, it is possible that the creator has some program to scrape the web for layoff events which are automatically populated into the Airtable. There is also a form on the website where anyone can enter data to be added to the spreadsheet. This implies that the data in the table is supplemented with data from web scraping as well as manually entered information from website viewers and users.

With all of this data compiled in one place, it should be convenient to have updated data for the project. However, there is not a simple way of downloading the data as it already appears in the online spreadsheet, so web scraping will be necessary to download it all. While the data includes several columns about each company and its layoff, the goal for this project will be to have the size of the layoff as the dependent variable and the other characteristics as the independent variables; location, industry, business stage and amount of money raised. If necessary, data about the revenue or financial situation of each company will be compiled as well from another source to be added as another independent variable. The goal is to be able to predict whether a company will have a layoff event in the future and approximately how large it will be based on its characteristics as studied in this analysis.

IV. Preliminary ArchitectureDiagram

Description automatically generated

As mentioned, the data for this project will primarily come from the service layoffs.fyi. To acquire the data in a format that can be worked with, web scraping will be utilized with the package Beautiful Soup from Python. This is the first step in the diagram above; downloading data from layoffs.fyi with web scraping in Python. By accessing the HTML code for the website, the spreadsheet of data can be downloaded. From there, the data will be loaded and stored in a database in MySQL. Since the data is updated regularly, web scraping will be used periodically to update the data in the MySQL database, so it remains up to date. While the data is stored in MySQL, it can then be downloaded to a notebook to be edited and manipulated with Python. The Python notebook is where exploratory data analysis and the majority of the model formulation will occur. Finally, the layoff data will be linked to a data visualization tool such as Tableau to visualize all the data from the project and show the results of the analysis.

V. Modeling

The data from this study will be used for a predictive analytics model on how large a layoff event is predicted to occur from a company given various features of that company. Based on data that has been collected over the course of nearly three years and is continually updated, a model will be constructed with the goal of being able to look at a current company in the industry and estimate how many employees (if any) are predicted to be laid off. Since there are over 1200 entries and counting, there should be enough data to create the model with about half the data, then test its accuracy with the remaining data in predicting the size of layoffs from companies. Hopefully a linear regression will be a good fit for the data, but if not, then other regression techniques will be utilized.

VI. Conclusion

The goal of this project is to uncover the common attributes that companies who have conducted major layoffs have. The data from past layoffs will be used to create a model to predict future layoffs. The technology layoffs are relevant to data scientists or new college graduates with STEM-related degrees as they can affect all jobs in technology companies as well as technology jobs in non-tech companies. This changes the job market landscape for those seeking employment. The results from this study of historical layoffs can potentially be used to provide insight on some warning signs within a company that layoffs may be coming. Perhaps this insight can be used by executives and managers to intervene in the problem before people’s jobs and livelihoods are affected.

1. References

Lee, Roger. “Tech Layoff Tracker and Startup Layoff Lists.” *Layoffs.fyi*, 9 Jan. 2023, <https://layoffs.fyi/>.