HR Employees Recruitment Prediction Dashboard (Humanlytics)

Technical Report

Muhammad Arslan
School of Computing and Engineering
University of Missouri - Kansas City
Kansas City, Missouri
manbb@umsystem.edu

Dallas Diaz
School of Computing and Engineering
University of Missouri - Kansas City
Kansas City, Missouri
ddmnh@umsystem.edu

Muhammad Tayyab
School of Computing and Engineering
University of Missouri - Kansas City
Kansas City, Missouri
mt56w@umsystem.edu

Ramya kumari Thambabattula
School of Computing and Engineering
University of Missouri - Kansas City,
Kansas City, Missouri
rtvtc@umsystem.edu

ABSTRACT

This unique initiative aims to create and deploy a user-friendly dashboard for HR professionals and recruiters. It focuses on offering real-time data and analytics to streamline recruitment processes, improve user experience, and promote internal mobility via an integrated job marketplace. Additionally, the project emphasizes data security and regulatory compliance. The full project scope includes requirements gathering, dashboard design, robust development, painstaking analytics integration, rigorous testing, detailed documentation, and strategic deployment. This dashboard includes sentiment analysis to capture and understand employee emotions during organizational changes such as layoffs. The dashboard analyzes input from direct interactions and social media posts to understand the emotional repercussions of such occurrences, giving HR teams with actionable insights.

Our dedicated team of engineers, designers, and data analysts will use AI-driven analytics to comprehend complicated statistics, allowing for proactive decision-making that benefits both the corporation and the employees.The project management method also includes governance systems, better communication strategies, and a thorough framework for identifying and mitigating risks. The HR dashboard will go through formal closing processes after meeting all approval criteria and providing the appropriate training and documentation for seamless integration into HR operations. This program intends to increase operational efficiencies and data-driven tactics within HR departments while also improving the overall workplace environment by sensitively addressing employee complaints and establishing a healthy organizational culture.

1. Introduction

In response to the growing demand for more complex and efficient HR procedures, our HR Dashboard Development project will launch a unique, user-friendly dashboard tailored exclusively for HR professionals and recruiters. This project

comes at a critical moment when firms are actively looking for technological solutions to improve their recruitment procedures and overall personnel management. Our complete dashboard combines a multitude of insights gleaned from employee engagement, retention rates, performance evaluations, diversity statistics, and other sources, with the goal of transforming how firms examine and comprehend critical HR data. By centralizing HR indicators, the platform delivers a comprehensive perspective of an organization's personnel dynamics, allowing HR teams to make data-driven decisions.

Designing an interface that is intuitively aligned with user needs.Consolidating data from several sources to provide a consistent view of HR metrics.Ensure the greatest level of data security and regulatory compliance.The addition of sentiment analysis tools to our project is a significant improvement, as it addresses the sentiment surrounding layoffs during organizational restructuring. Our advanced AI tools, notably the Large Language Model (LLM), investigate the emotional consequences of such modifications using datasets containing employee input. This feature enables a comprehensive analysis of textual data in order to extract sentiments and comprehend emotional responses, with the goal of providing deep insights into the morale implications of organizational layoffs.

1.2 Target Audience

The primary target audience for our dashboard is HR departments in mid-sized businesses, who would benefit greatly from the sentiment analysis tool. This application allows HR professionals to understand the full effects of layoffs and change their policies accordingly to better manage such transitions.

1.3 Empowering HR departments.

Our dashboard enables HR departments by giving critical insights to help them negotiate layoffs sensitively and effectively. HR can use data-driven ways to address concerns proactively, building an environment of empathy and support in the workplace.

1.4 Comparative Analysis of Affected Employees

The dashboard provides comparative analysis tools for layoff-affected employees, assisting with their transition and adjustment by offering context and clarity on their experiences in comparison to their peers.

2. Methodology

The Humanalytics HR Dashboard project uses a comprehensive methodology to create a robust, user-friendly platform designed exclusively for HR professionals and recruiters. This document covers our strategy approach to various essential development components, such as frontend development, backend infrastructure, AI model integration, and accessibility and usability advancements.

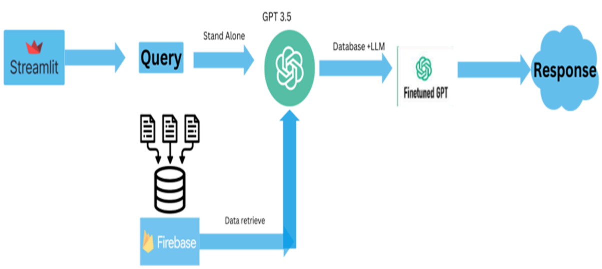


Figure 2.0 *Flow chart*

2.1 FrontEnd Infrastructure

The frontend of the Humanalytics HR Dashboard is created with Visual Studio Code (VS Code), a robust and efficient code editor that supports a wide range of computer languages and frameworks. This environment increases our developers' productivity and promotes a more efficient coding experience. We use Streamlit for interface design and user interaction, an open-source app framework that is ideal for generating interactive and visually appealing web applications. Streamlit's simplicity and flexibility allow for rapid development and prototyping, ensuring that our dashboard is not only functional but also accessible and appealing to end users.

2.2 Backend Infrastructure

The backbone of our platform's backend is Python, which is well-known for its versatility and comprehensive support for data analysis and machine learning tools. Python's scalability is critical for handling the HR dashboard's fluctuating loads and complicated data processing requirements. To store and manage data, we use Firebase, a cloud-based database noted for its real-time data synchronization and seamless connection with frontend services. Firebase provides comprehensive data management for our application, ensuring secure and efficient operations across the dashboard's various components.

2.3 AI Model Integration

To improve the dashboard's ability to read and analyze textual data, we incorporate many advanced AI models:

GPT-3.5 and GPT-4: These OpenAI models are used for their cutting-edge text creation and reasoning skills, which allow the dashboard to interpret natural language inquiries and deliver coherent and contextually appropriate responses.

LDA (Latent Dirichlet Allocation): This technique is used for topic modeling in HR data. It aids HR professionals in finding widespread themes and trends by facilitating the discovery of abstract subjects from a huge collection of documents.

XLNet, BERT, and RoBERTa: These models are used for their superior sentiment analysis and text classification. They analyze candidates' thoughts and properly categorize textual material, which is critical for determining candidate fit and improving the recruitment process.

2.4 Enhancing Accessibility and Usability

Recognizing that consumers have varying levels of technology access, we designed our dashboard to be fully responsive, ensuring that it runs smoothly on a variety of devices, including mobile phones and tablets. This responsiveness ensures that HR professionals and recruiters may access and use the dashboard's capabilities at any time and from any location, increasing the flexibility and ease of their operations.

Figure 2.4 *LLM, Streamlite and Firebase Connection*

2.5 Triangle Model

Our HR Dashboard Development project adopts a robust triangular model, integrating three key components for a comprehensive solution. For the frontend, we leverage a modern and user-friendly interface, utilizing technologies such as VS Code and Streamlit.

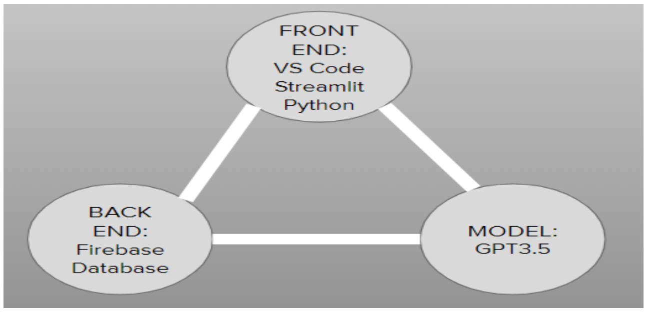


Figure 2.5 In this figure it shows all the components we are using.

3. DATASET

The acquisition of company datasets, comprising both structured and unstructured formats, stands as a pragmatic cornerstone in the development of our HR Dashboard. The structured datasets, containing detailed information such as company names, locations, employee counts, and essential HR metrics, provide a solid foundation for our analytics and visualizations. These datasets, collected from diverse industries, reflect the realistic intricacies of workforce data in various organizational settings. Simultaneously, unstructured datasets, capturing textual nuances related to HR activities, present an opportunity for our AI model, GPT-3.5, to clean valuable insights. The realistic inclusion of a range of companies, each with its unique characteristics and challenges, ensures that our dashboard is tailored to real-world scenarios, accommodating the complexities and diversities inherent in HR operations.

company	location	industry	total_laid_off	percentage_laid_off	date	stage	country	funds_raised
New Work	Hamburg	Consumer	400		2024-01-11	Post-IPO	Germany	
Playtika	Tel Aviv	Consumer	300	0.1	2024-01-11	Post-IPO	Israel	
Discord	SF Bay Area	Consumer	170	0.17	2024-01-11	Series H	United States	995
Inmobi	Bengaluru	Marketing	125	0.05	2024-01-11	Unknown	India	320
Audible	New York City	Media	100	0.05	2024-01-11	Acquired	United States	14
Sisense	New York City	Data	60	0.13	2024-01-11	Series F	United States	274
Google	SF Bay Area	Consumer	1000		2024-01-10	Post-IPO	United States	26
Beam Benefits	Columbus	Healthcare	74		2024-01-10	Series E	United States	168
Instagram	SF Bay Area	Consumer	60		2024-01-10	Acquired	United States	57
Amazon	Seattle	Retail			2024-01-10	Post-IPO	United States	108
ChargePoint	SF Bay Area	Manufacturing		0.12	2024-01-10	Unknown	United States	1400
[illegible]trix	Miami	Infrastructure		0.12	2024-01-10	Acquired	United States	20
[illegible]oFI	SF Bay Area	Finance		0.04	2024-01-10	Post-IPO	United States	3000
Twitch	SF Bay Area	Consumer	500	0.35	2024-01-09	Acquired	United States	35
Branch	Columbus	Finance	85		2024-01-09	Series C	United States	229
Nevro	SF Bay Area	Healthcare	63	0.05	2024-01-09	Post-IPO	United States	556
Uber Freight	SF Bay Area	Logistics	40		2024-01-09	Subsidiary	United States	2700
Rent the Runway	New York City	Retail	37	0.1	2024-01-09	Post-IPO	United States	526
Humane	SF Bay Area	Hardware	10	0.04	2024-01-09	Series C	United States	230
Trend Micro	Tokyo	Security		0.02	2024-01-09	Unknown	Japan	
Unity	SF Bay Area	Other	1800	0.25	2024-01-08	Post-IPO	United States	1300
NuScale Power	Corvallis	Energy	154	0.28	2024-01-08	Post-IPO	United States	470
Pitch	Berlin	Other	80	0.67	2024-01-08	Series B	Germany	138

Figure 3.0 In figure it is showing the raw data of the layoff.

user_name	user_location	user_description	user_created	user_followers	user_friends	user_favourites	user_verified	date	text	hashtags	source
BestCompaniesAZ	Arizona	We validate and ele	2009-01-11 16:28:33	2836	2469	870	FALSE	2022-12-20 18:00:48	Apply today!: https://t.co/eCRVPFE #apply #hiring #oper	['apply', 'hiring', 'ope	Hootsuite Inc.
BestCompaniesAZ	Arizona	We validate and ele	2009-01-11 16:28:33	2835	2469	870	FALSE	2022-12-20 18:00:48	Apply today!: https://t.co/eCRVPFE #apply #hiring #oper	['apply', 'hiring', 'ope	Hootsuite Inc.
CHILUDO		AMC APE MAKING	2022-03-27 17:10:55	19	32	4	FALSE	2022-12-18 00:21:03	I want to wish all @Twitter employees that		Twitter Web App
ANIL	New Delhi, India, Fro	New Delhi. Above po	2010-04-17 16:58:14	81	287	4484	FALSE	2022-12-17 14:04:44	@elonmusk sir has ca	['TwitterLayoffs']	Twitter for Android
My Head's Explodin	leswarden@mastodo	A Proud Believer in	2009-01-28 01:03:18	1044	777	3742	FALSE	2022-12-16 19:47:58	Here's what happens	['TickTok']	Twitter for iPhone
Luke Petersen		33years • born in the	2012-01-21 09:43:13	19	105	5833	FALSE	2022-12-16 09:18:57	@Curtis415 @Adaml	['Elon']	Twitter for iPhone
[illegible]W Businessworld	New Delhi	BW Businessworld is	2009-05-01 05:57:27	21539	437	5272	FALSE	2022-12-16 07:58:00	#US Judge Asks @Tv Read: https://t.co/93\ #twitter #layoffs #twit	['US']	TweetDeck Web App
[illegible]z-e-Amor	NorCal	ELL teacher 30+ yea	2010-07-17 23:36:39	395	1186	106249	FALSE	2022-12-16 01:13:18	@wkamaubell he eve	['TwitterLayoffs']	Twitter for iPhone
Lisa Bloom	Los Angeles	Trial lawyer fighting	2008-08-26 14:09:47	133631	952	3481	TRUE	2022-12-16 00:51:02	We just filed yet ano #TwitterLayoffs	['TwitterLayoffs']	Twitter Web App
The White Rabbit	Naples, Italy	U.S. Navy Vet. No Pa	2022-11-26 22:33:27	276	526	1731	FALSE	2022-12-15 17:17:44	We got stick-up kids, corrupt cops, and cra		Twitter for iPhone
KevinB	Arizona	Internet made simple	2022-11-28 20:00:18	6	54	4	FALSE	2022-12-15 01:56:22	@elonmusk Hey papa elon @elonmusk, wi #twitterlayoffs #thisshouldntbelegal https://		Twitter Web App
Abdullah	San Francisco, CA	CEO T one Social M	2021-11-27 18:19:16	27	423	12	FALSE	2022-12-14 18:51:07	I own a new social communication app Now on Android and IOS Let's start a new company and launch the #TwitterLayoffs #TwitchStreamers #twitchaffiliate #smallstreamer #closetoaffiliat		Twitter for iPhone
Mario Di Maggio	Bristol	Senior science comi	2021-09-10 20:00:12	98	50	3530	FALSE	2022-12-14 17:33:25	Erm @elonmusk, acc	['LoveWhereYouWor	Twitter Web App
Samfiru Tumarkin LL	Ontario, BC and Albe	One of Canada's larg	2009-07-16 20:34:41	1846	1056	2341	FALSE	2022-12-14 16:15:11	The lawsuit by forme Understand your #rig $TWTR https://t.co/W0ynY4h	['Twitter']	Agorapulse app

Figure 3.0.1 In figure it is showing the raw data we have collected from the employees tweets.

3.1 DATA PRE-PROCESSING

For better visualization and analysis we extract some information from the data set and also rename some columns we have performed these steps shown below.

s Help All changes saved Comment Share

+ Code + Text RAM Disk

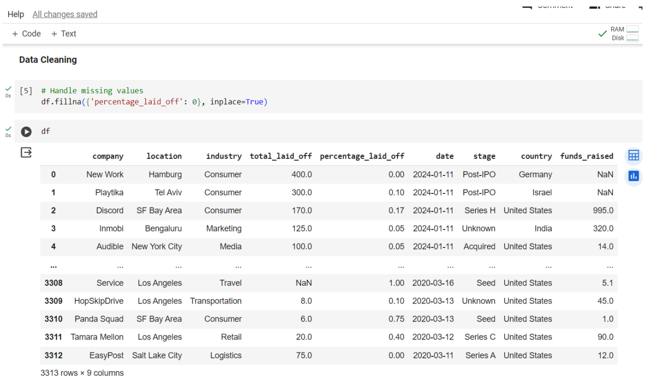
```
[1] import pandas as pd

[2] # Load the dataset
    df = pd.read_csv('/content/drive/MyDrive/capstone/layoffs.csv')

df.head()
```

	company	location	industry	total_laid_off	percentage_laid_off	date	stage	country	funds_raised
0	New Work	Hamburg	Consumer	400.0	NaN	2024-01-11	Post-IPO	Germany	NaN
1	Playtika	Tel Aviv	Consumer	300.0	0.10	2024-01-11	Post-IPO	Israel	NaN
2	Discord	SF Bay Area	Consumer	170.0	0.17	2024-01-11	Series H	United States	995.0
3	Inmobi	Bengaluru	Marketing	125.0	0.05	2024-01-11	Unknown	India	320.0
4	Audible	New York City	Media	100.0	0.05	2024-01-11	Acquired	United States	14.0

Next steps: View recommended plots

Figure 3.1.1 Dataset loading

In figure 3.1.1 we perform the Column Removal. This approach simplifies the dataset by removing attributes that are irrelevant to the investigation, boosting model performance, and increasing interpretability.



Help All changes saved

+ Code + Text

Data Cleaning

```
[5] # Handle missing values
df.fillna({'percentage_laid_off': 0}, inplace=True)
```

```
df
```

	company	location	industry	total_laid_off	percentage_laid_off	date	stage	country	funds_raised
0	New Work	Hamburg	Consumer	400.0	0.00	2024-01-11	Post-IPO	Germany	NaN
1	Playtika	Tel Aviv	Consumer	300.0	0.10	2024-01-11	Post-IPO	Israel	NaN
2	Discord	SF Bay Area	Consumer	170.0	0.17	2024-01-11	Series H	United States	995.0
3	Inmobi	Bengaluru	Marketing	125.0	0.05	2024-01-11	Unknown	India	320.0
4	Audible	New York City	Media	100.0	0.05	2024-01-11	Acquired	United States	14.0
...	...	...	...	...	...	...	...	...	...
3308	Service	Los Angeles	Travel	NaN	1.00	2020-03-16	Seed	United States	5.1
3309	HopSkipDrive	Los Angeles	Transportation	8.0	0.10	2020-03-13	Unknown	United States	45.0
3310	Panda Squad	SF Bay Area	Consumer	6.0	0.75	2020-03-13	Seed	United States	1.0
3311	Tamara Mellon	Los Angeles	Retail	20.0	0.40	2020-03-12	Series C	United States	90.0
3312	EasyPost	Salt Lake City	Logistics	75.0	0.00	2020-03-11	Series A	United States	12.0

3313 rows × 9 columns

Figure 3.1.2 *Data Cleaning*

In figure 3.1.2 we perform the Column renaming. This clarifies the dataset and assists users in understanding the purpose and substance of each attribute.

+ Code + Text

```
[8] # Convert date columns to datetime format
df['date'] = pd.to_datetime(df['date'])
```

```
[9] # Check for duplicates
df.drop_duplicates(inplace=True)
```

```
[10] df
```

	company	location	industry	total_laid_off	percentage_laid_off	date	stage	country	funds_raised
0	New Work	Hamburg	Consumer	400.0	0.00	2024-01-11	Post-IPO	Germany	NaN
1	Playtika	Tel Aviv	Consumer	300.0	0.10	2024-01-11	Post-IPO	Israel	NaN
2	Discord	SF Bay Area	Consumer	170.0	0.17	2024-01-11	Series H	United States	995.0

Figure 3.1.3 Remove null values

In figure 3.1.3 we are dealing with *NaN* or *null* values in the dataset. To maintain data quality and analytical accuracy, common procedures include imputation (replacing missing values with estimates), eliminating rows or columns with missing values, or specific treatment based on the context.

Data Type Identification: Recognizing the data type (e.g., integer, float, string) of each column is critical for selecting appropriate data transformation techniques, dealing with categorical variables, and ensuring compatibility with analysis or modeling algorithms as shown in figure 6.

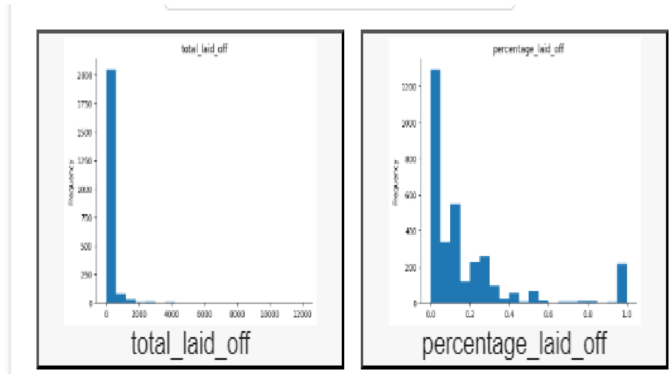


Figure 3.1.4 *Total and Percentage Layoff Visualization*

3.2 Data Analysis:

In our HR Dashboard, we've implemented two insightful graphs to provide a comprehensive view of workforce dynamics. The first graph elegantly illustrates the percentage of layoffs, offering a visual representation of the proportion of employees affected. This metric is crucial for understanding the relative impact on different departments or organizational levels. The second graph, focusing on the total number of layoffs, offers a tangible figure, providing a quantitative perspective on the extent of workforce changes. Together, these graphs empower HR professionals with valuable insights into workforce restructuring, facilitating strategic decision-making and proactive management of organizational changes.

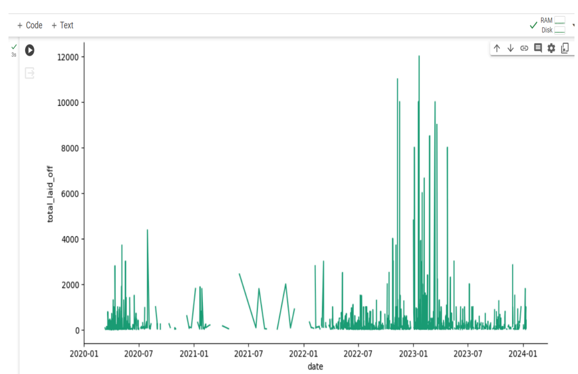


Figure 3.2.1 *Layoffs Visualization*

In figure 3.2.1 the graph depicting the total layoffs from 2020 to 2024 serves as a dynamic temporal visualization, offering a chronological overview of workforce changes over the specified period. This graph provides HR professionals with a comprehensive historical perspective, enabling them to identify trends, patterns, and potential correlations with external factors such as economic fluctuations or organizational shifts. The plotted data facilitates the analysis of long-term workforce dynamics, helping stakeholders make informed decisions based on historical context. This graphical representation proves instrumental in understanding the trajectory of workforce changes, supporting strategic planning, and fostering a proactive approach to human resource management.A pivotal moment in the pandemic, as numerous U.S. school districts initiated the transition to online learning. This shift in educational delivery transformed the daily routines and mobility patterns of students, parents, and educators.

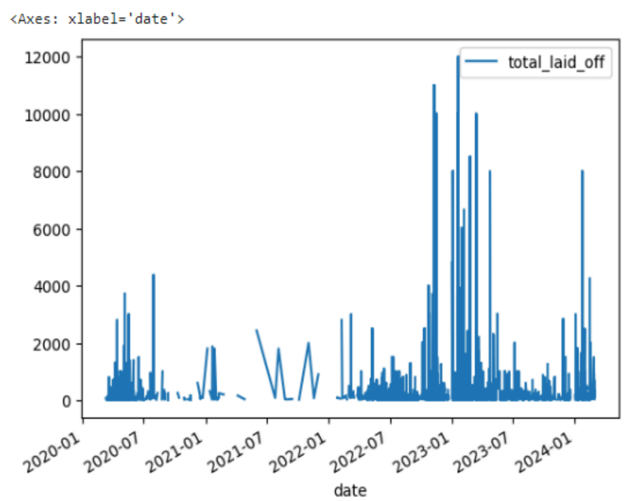


Figure 3.2.2 *Total Layoff Visualization from 2020 to 2024*

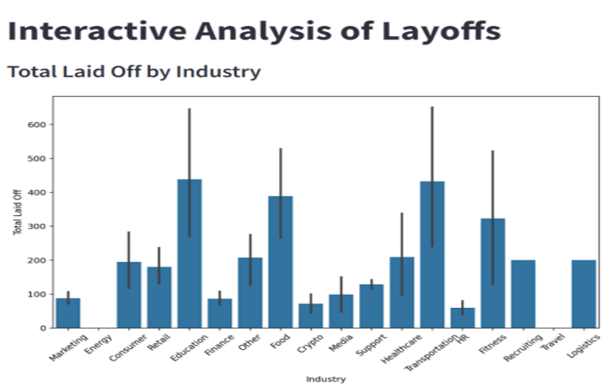


Figure 3.2.3 Interactive Analysis of Layoffs on Industry

The image displays an interactive analysis graph that depicts the impact of layoffs across several businesses, offering a clear visual indication of how different sectors have been affected. The x-axis of the graph is divided into numerous main businesses, including marketing, energy, retail, education, food, and transportation. Each category represents a separate industry, highlighting the specific focal areas where layoffs were quantitatively studied.

The graph's y-axis shows the overall number of layoffs, providing a comparison statistic for evaluating the severity and distribution of layoffs within each industry. This enables users to immediately identify which industries have experienced the most employment cuts, making the graph a valuable tool for examining patterns and implications across many economic sectors.

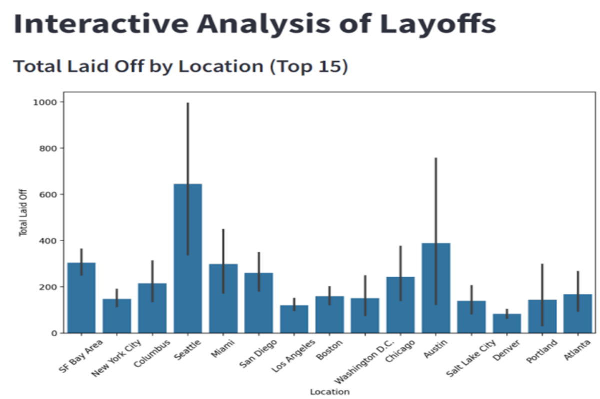


Figure 3.2.4 Total Laid off by Location (Top 15)

This bar graph displays a study of layoffs segregated by geographic location across various states in the United States. The x-axis of the graph displays the names of 15 different states, offering a geographical breakdown of where layoffs have occurred most frequently. This classification enables readers to quickly compare the impact of layoffs across areas.

On the y-axis, the graph counts the total number of layoffs, allowing for a direct comparison of the magnitude of workforce losses among the states shown. This vertical measurement provides strong insights into the magnitude of layoffs, with bigger bars suggesting more layoffs within those specific states.

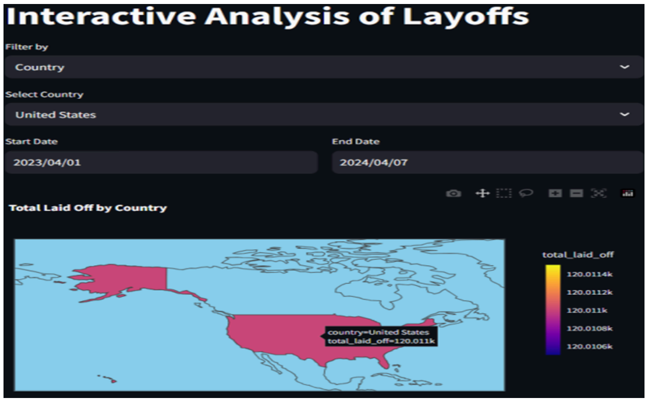


Figure 3.2.5 *User interface for Interactive Layoff Analysis*

This interactive nature increases user engagement by allowing viewers to interact with the data, such as changing the country name, clicking on specific industry segments to drill down for more detailed data, or filtering the display to show only the most affected industries. This interactivity not only makes the data more accessible, but also more understandable, since users may change the visual representation to better grasp the subtleties of the layoff environment. We have also added map visualization.

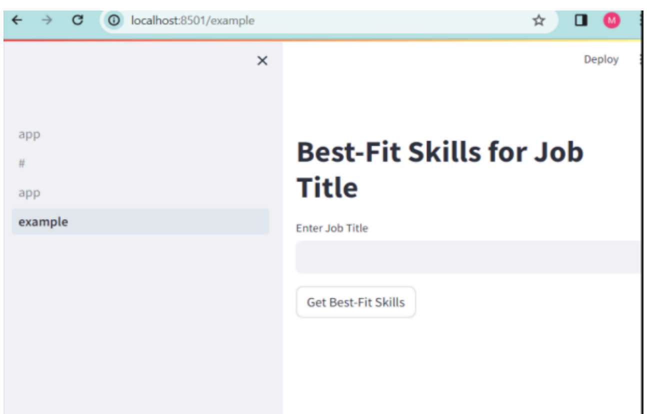


Figure 3.2.6 *User Interface for Best_fit skills*

The integration of a user-friendly front-end feature, where applicants input their skills to receive tailored job recommendations, marks a significant advancement in our HR Dashboard. This innovative interface not only streamlines the job application process but also enhances the user experience by providing personalized and relevant job matches. Applicants can enter their skill sets, and the system, powered by advanced algorithms, swiftly matches their qualifications with available job opportunities. This dynamic front-end functionality not only empowers applicants to find roles that align with their expertise but also facilitates a more efficient and targeted recruitment process for HR professionals. This user-centric approach reinforces the dashboard's commitment to improving accessibility and engagement, ultimately fostering a more seamless connection between job seekers and potential employers.

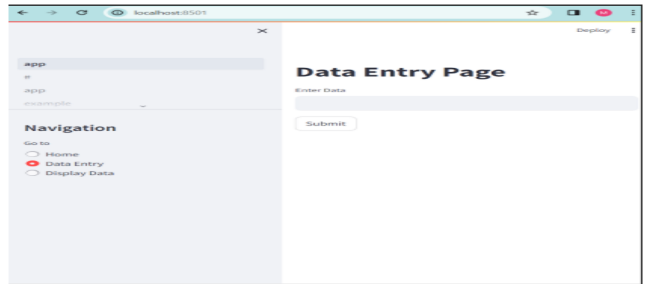


Figure 3.2.7 Data Entry Page

In figure 3.2.7 it showed the incorporation of a data entry box on the front end, seamlessly connected to our backend database using Firebase, enhances the functionality and versatility of our HR Dashboard. This feature allows for the easy and efficient input of new data directly from the user interface, eliminating the need for complex data entry processes. The entered data, encompassing various HR metrics and information, is securely stored in our Firebase backend, ensuring real-time accessibility and centralized management.

3.3-Model-Implementation:

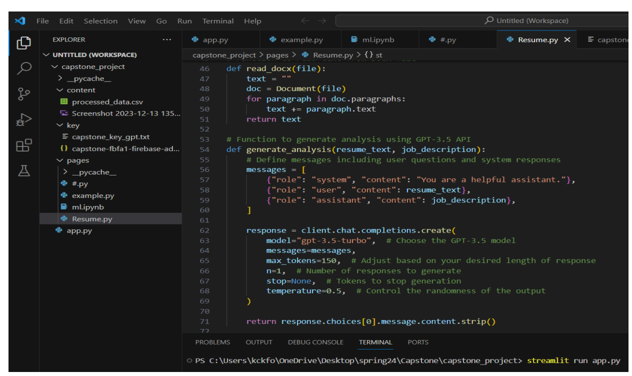


Figure 3.3.0 *GPT-3.5 Code* Analyses

In figure 3.3.0 the function uses the GPT-3.5 API to generate textual analyses depending on user input. It allows you to customize settings like response length, amount of responses, and randomization control. Furthermore, it contains facilities for defining user inquiries and system responses, allowing flexibility in generating specialized analysis.

Figure 3.3.1: *UI Resume Classification*

In figure 3.3.1 the integration of a resume analysis function in our HR Dashboard significantly improves user experience and recruitment efficiency. By allowing applicants to submit their resumes, our system uses sophisticated algorithms to extract relevant skills and qualifications. This research offers specific employment recommendations based on each applicant's expertise and career goals. This unique feature not only speeds the application process, but also provides HR managers with vital information about candidate profiles. Our dashboard enables a more targeted recruitment process by utilizing AI-driven resume analysis, matching individuals with opportunities that best suit their strengths.

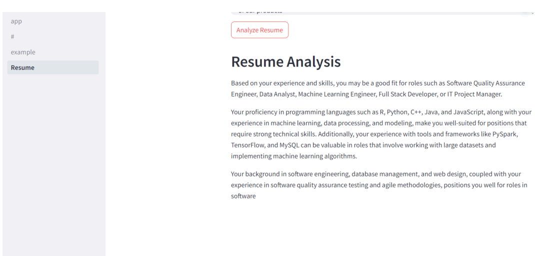


Figure 3.3.2 : *Resume Classification Report*

In figure 3.3.2 the GPT-3.5 model offers a significant leap in candidate evaluation. This tool provides users with comprehensive insights on their resumes, including relevant talents, experiences, and qualifications found using powerful natural language processing. Our system uses the capabilities of GPT-3.5 to provide thorough analysis and comments on resumes, allowing users to better understand their strengths and areas for improvement.

The inclusion of the Firebase data importer functionality in our HR Dashboard improves data management and accessibility for HR professionals. This feature allows users to simply enter additional address or skill data straight into our backend Firebase database via a user-friendly interface. This automated approach eliminates the need for manual data entry and guarantees that the database is updated in real time. We ensure data integrity while allowing easy integration with the dashboard's analytics and visualization features by utilizing Firebase's scalable and secure storage solution.

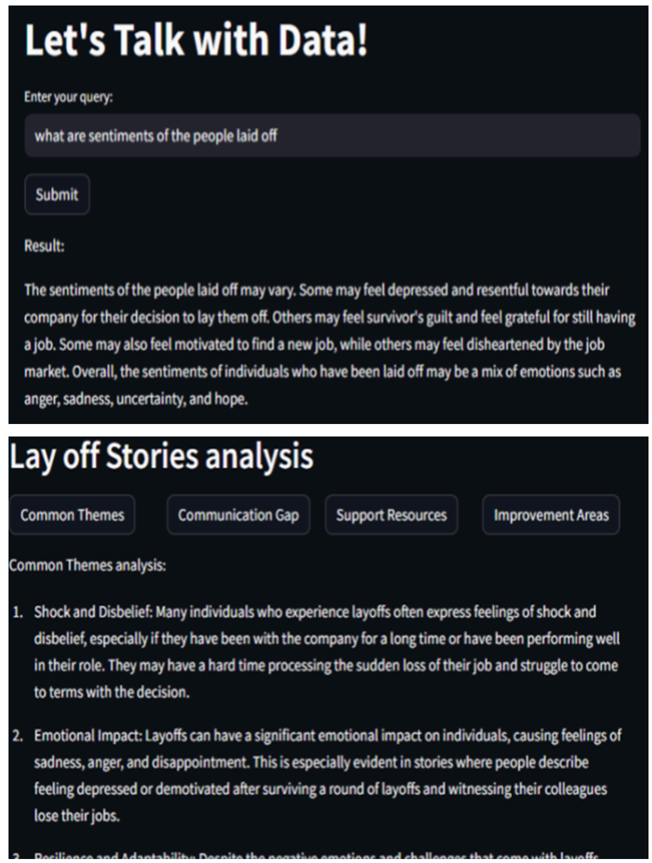


Figure 3.3.3 *Story Analysis for Layoff*

The inclusion of a conceptual diagram displaying the LLM (Long-Short Term Memory) model utilized in our system demonstrates the complex process of handling incoming requests and generating responses. This model serves as the foundation for our system's natural language processing capabilities, allowing it to understand and interpret user queries effectively. Incoming requests, such as job lay off stories analysis or talk with data, are handled by the LLM model/langchain which evaluates the context and extracts pertinent data. As a result, the model produces accurate and contextually appropriate responses, ensuring a smooth connection between people and the system. This is the technology behind our HR Dashboard, demonstrating our dedication to employing advanced AI models to enhance the user experience and give significant insights in the field of human resources.

4. Results (LDA & Comparison GPT 3.5, 4 Bert, RoBERTa)

```
[6] ground_truth = [
    {"question": "What major event occurred in late 2022 involving a tech company?", "answer": "A tech company bought another company and worked on
    {"question": "What was the expectation after the integration of the two companies?", "answer": "The expectation was for full integration of hea
    {"question": "What happened to senior management, including the former CEO, after the integration?", "answer": "Senior management, including th
    {"question": "What was the unexpected news delivered via email from the CEO?", "answer": "The email stated that due to the current business cli
    {"question": "How did the layoffs affect the entire team?", "answer": "The layoffs affected everyone, including leaders, leaving them feeling b
    {"question": "How long has the person been searching for a new job?", "answer": "The person has been searching for a new job for seven months."
    {"question": "What significant event happened in the tech industry that the person read about on CNN?", "answer": "There was an AT&T outage lik
    {"question": "What concerns does the person have about certain programs in the tech industry?", "answer": "The person is concerned about progra
    {"question": "How does the person plan to cope with the stress of unemployment?", "answer": "The person plans to simplify life, possibly by mov
    {"question": "What lesson did the person learn from their career in software quality assurance?", "answer": "The person learned to conduct thei
    {"question": "How did the layoffs affect the person financially and emotionally?", "answer": "The person experienced financial strain due to be
    {"question": "What advice did the person receive after being laid off?", "answer": "The person received advice to continue networking, learn ne
    {"question": "How did the layoffs affect the person's perspective on loyalty to a company?", "answer": "The person shifted from a perspective o
    {"question": "What advice did the person offer to others facing layoffs?", "answer": "The person advised others to stay proactive, continue lea
    {"question": "What was the person's experience after being laid off from their job in tech?", "answer": "The person faced challenges in finding
    {"question": "How did the layoff affect the person's financial situation?", "answer": "The layoff significantly reduced the person's income, le
    {"question": "What were some strategies the person used to survive financially after the layoff?", "answer": "The person utilized state program
    {"question": "How did the layoff impact the person's mental health?", "answer": "The layoff took a toll on the person's mental health, leading
    {"question": "What advice did the person offer to others facing layoffs?", "answer": "The person advised others to take action quickly, seek su
    {"question": "How did the person's perspective on job security change after being laid off?", "answer": "The person no longer viewed job securi
    ]
```

Figure 4.0 Responses generator for layoffs stories

The image depicts a full display of an analytical evaluation performed utilizing LLM (Large Language Models) such as GPT-3.5 and GPT-4, with a focus on their performance in generating responses to queries drawn from layoff stories. The experiment involves constructing a dataset of 20 questions based on real-world circumstances (ground truth) related to layoffs, to which the LLMs responded. These responses were then evaluated for semantic similarity using spaCy's word similarity metrics in Python, a quantitative approach that compares the AI-generated text to the human-generated answers.

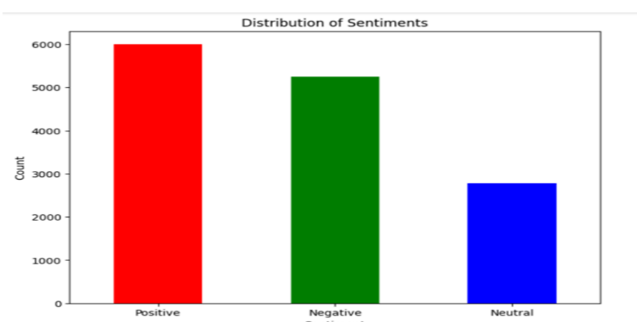


Figure 4.1 *Distribution of sentiments*

The image shows a bar graph that visually depicts the distribution of feelings taken from text data about layoffs. The x-axis divides sentiments into three categories: positive, negative, and neutral, allowing users to rapidly determine the overall emotional tone represented by the dataset. The y-axis measures the frequency of each sentiment, giving a tally of how many times each sentiment category appears in the examined texts.

The bar graph clearly indicates the prevalence of each sentiment type, with negative sentiment most likely dominating the graph, representing common emotional reactions to layoffs such as disappointment, irritation, and fear. Neutral sentiments may reflect more unclear or conflicted emotions, and positive sentiments, while possibly less common, may show instances of hope or acceptance among impacted persons.

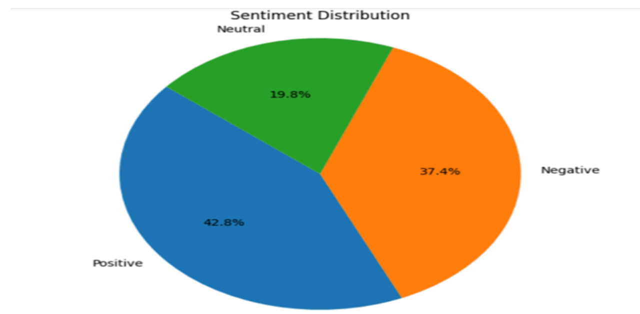


Figure 4.2 *Pie Sentiment Distribution Analysis*

The figure 4.2 includes a pie graphic that depicts the sentiment distribution derived from text data about layoffs. This figure efficiently divides down the proportions of positive, negative, and neutral attitudes conveyed in the studied texts, with each segment color-coded for easy identification.

The pie chart shows the percentage that each sentiment category contributes to the overall dataset. Typically, the Negative segment would take up a large portion of the chart, representing common emotional reactions to layoffs such as despair and frustration. The Neutral category includes replies that indicate neither direct negative nor positive, maybe indicating ambivalence or a guarded attitude toward layoffs. The Positive part, which is likely the shortest, depicts situations in which people express positive or hopeful feelings despite the circumstances.

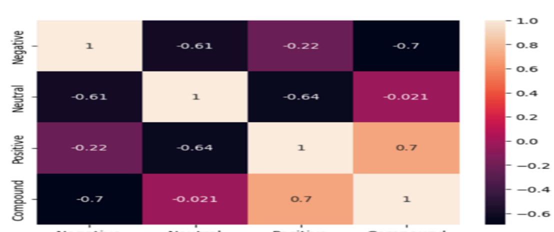


Figure 4.3: *Layoffs* sentiment analysis

The graphic depicts the findings of a sentiment analysis performed on layoff-related text data, as shown by correlation coefficients that quantitatively quantify the emotional tones and themes conveyed by employees affected by layoffs. This study reveals a primarily negative sentiment across the sample, indicating widespread feelings of disappointment, dissatisfaction, and despair among employees.

The graph shows that negativity increases when employees explain the direct effects of layoffs on their careers and financial stability, implying that these factors are important triggers for negative emotional responses. Furthermore, the data reveals significant indicators of uncertainty and worry about the future, reflecting employees' fears about their capacity to acquire new job possibilities and effectively navigate the problems that come following layoff.

```
from matplotlib import pyplot as plt
df.plot(kind='scatter', x='total_laid_off', y='percentage_laid_off', s=32, alpha=.8)
plt.gca().spines[['top', 'right',]].set_visible(False)
```

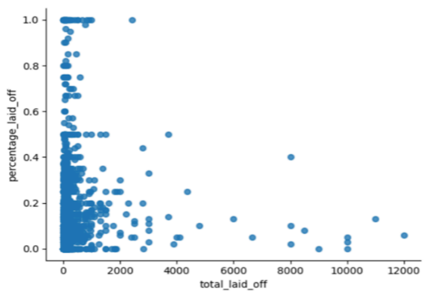


Figure 4.4 *Layoffs scatter plot visualization*

The data points are scattered across the plot, revealing a pattern where higher values of "total_laid_off" generally correspond to lower values of "percentage_laid_off". The plot demonstrates an overall negative correlation between the two variables. The code snippet provided in the image specifies the plot type as a scatter plot ('kind='scatter"), sets the x and y variables, adjusts the marker size (s=32), and sets the transparency (alpha=0.8). Additionally, it hides the top and right spines of the plot using the set_visible(False) method.

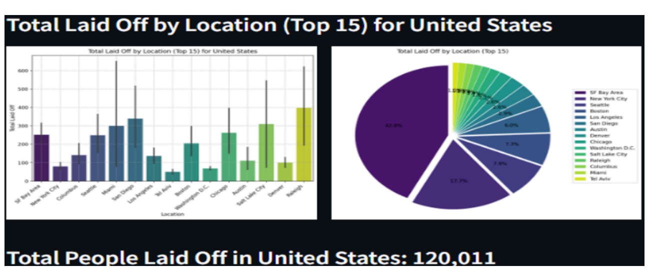


Figure 4.5 *United States total laid of Visualization*

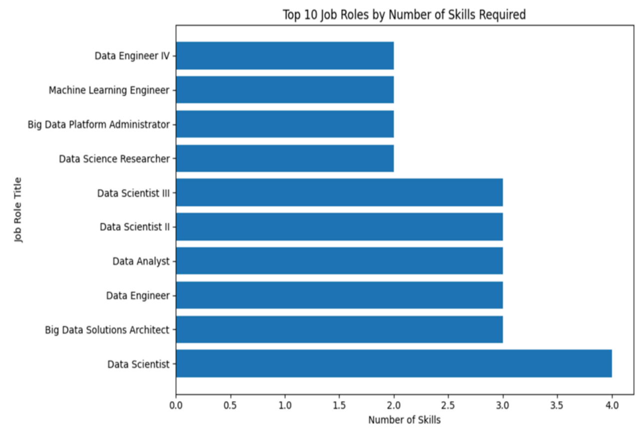


Figure 4.6 *Skill Analysis*

The top 10 job roles are ranked by the number of skills required for each role. At the top of the list are "Data Engineer IV," "Machine Learning Engineer," and "Big Data Platform Administrator," indicating they require the highest number of skills. Other roles listed include "Data Science Researcher," "Data Scientist III," "Data Scientist II," "Data Analyst," "Data Engineer," "Big Data Solutions Architect," and "Data Scientist." This visualization offers valuable insight into the complexity and skill requirements associated with various data-related job roles.

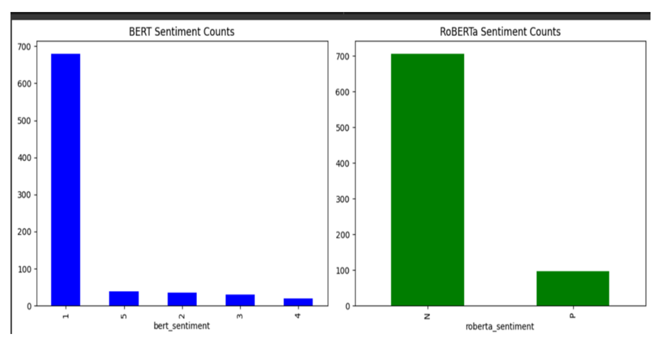


Figure 4.7 *Bert* sentiment *validation*

In this image The bar graphs present a visual comparison of the sentiment analysis results obtained from two different models, BERT and RoBERTa, when applied to a dataset of layoff-related text items. Each graph divides the sentiments into categories such as "positive," "neutral," and "negative."

The left graph shows the sentiment counts as examined by the BERT model, while the right graph shows the sentiment counts as analyzed by the RoBERTa model. Notably, there are significant disparities in how each model perceived the sentiments. For example, the BERT model appears to detect a higher number of extreme feelings (either highly positive or very negative), whereas the RoBERTa model classifies more attitudes as "Neutral."

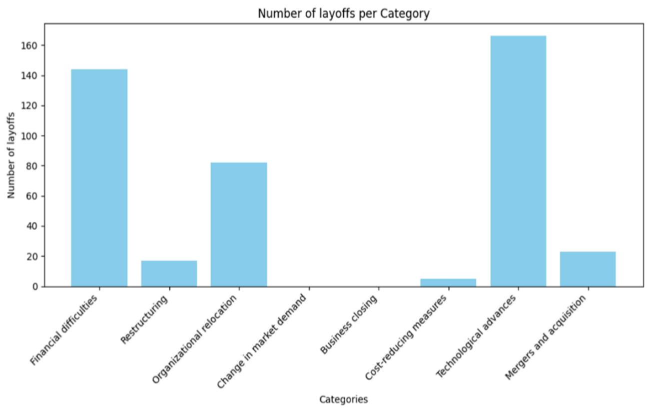


Figure 4.8: *Layoffs category Visualization*

In this Image The bar graph headed "Number of layoffs per Category" depicts the reasons for employee layoffs in several categories. The x-axis categorizes the reasons for layoffs, and the y-axis quantifies them. Notably, the categories "financial difficulties" and "technological advances" have the greatest numbers, indicating a considerable impact on layoffs. "Restructuring" also has a large number, indicating significant organizational changes. In contrast, "organizational relocation/change in market demand" and "mergers and acquisition" are connected with fewer layoffs, while "business closing/cost-cutting measures" shows moderate figures. This graph effectively emphasizes the primary variables contributing to layoffs inside firms, providing significant insights for identifying patterns in staff cutbacks.

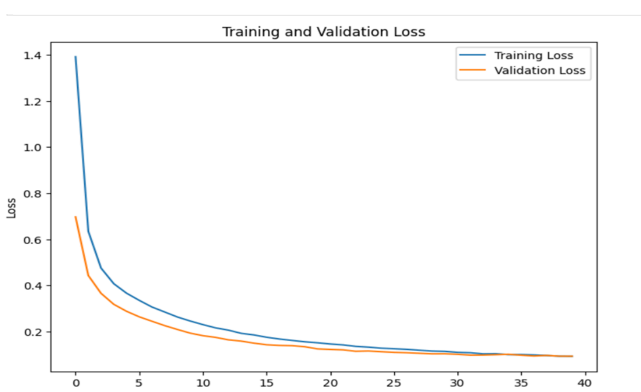


Figure 4.9 *Layoffs Training and testing validation*

The line plot depicts the training and validation loss curves throughout the training of a machine learning model. The y-axis represents the loss value, while the x-axis represents the epoch number. Initially, both the training and validation losses are high, but they gradually decrease as the training proceeds. The validation loss curve closely tracks the training loss curve, suggesting that the model is neither overfitting nor underfitting. This plot serves as a valuable tool for monitoring the model's convergence during training and identifying potential issues such as overfitting or underfitting.

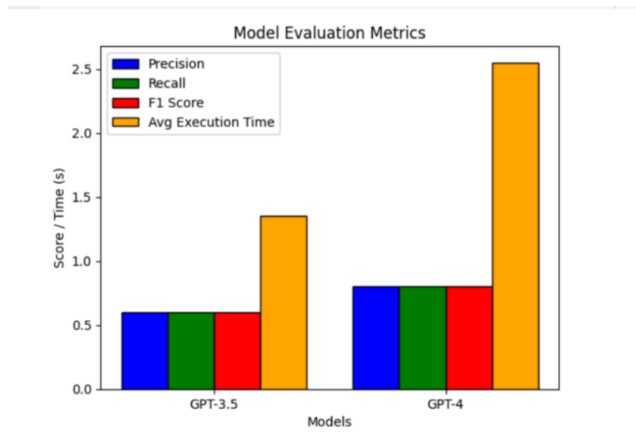


Figure 4.10 Model Evaluation Metrics

The graphic depicts a detailed comparison of two complex language models, GPT-3.5 and GPT-4, utilizing a Model Evaluation Matrix to visually illustrate their performance in important measures such as Precision, Recall, F1 Score, and Average Execution Time. These metrics critically evaluate each model's capacity to process and reply to complicated layoff-related queries. Precision shows the accuracy of the models' positive predictions, recall measures the models' capacity to catch all relevant occurrences, and the F1 Score is a balanced statistic that combines precision and recall. Furthermore, the Average Execution Time statistic measures each model's efficiency, showcasing the speed with which responses are generated.

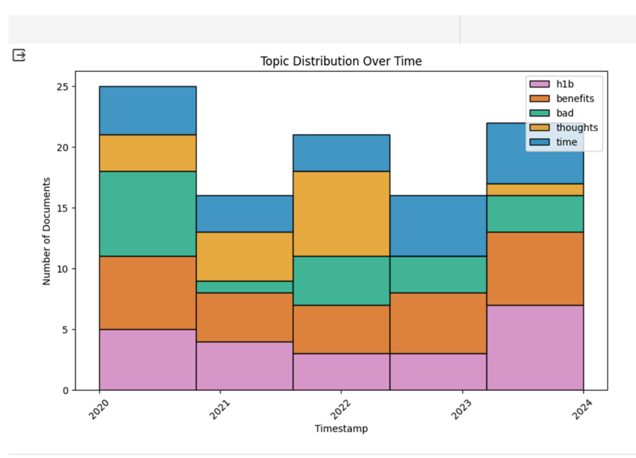


Figure 4.11 Topic Distribution Over Time

The chart illustrates the distribution of various topics over different timestamps spanning from 2020 to 2024. The topics depicted include "h1b," "benefits," "bad," "thoughts," and "time." The distribution of these topics fluctuates across the years, with some topics being more prominent in certain years than others.

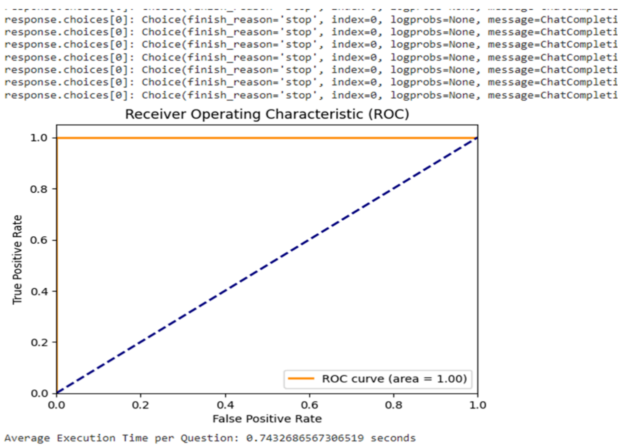


Figure 4.12 *ROC* classification

The ROC curve illustrates the true positive rate plotted against the false positive rate across different threshold settings. Here, the area under the ROC curve is 1.0, suggesting flawless classification performance.

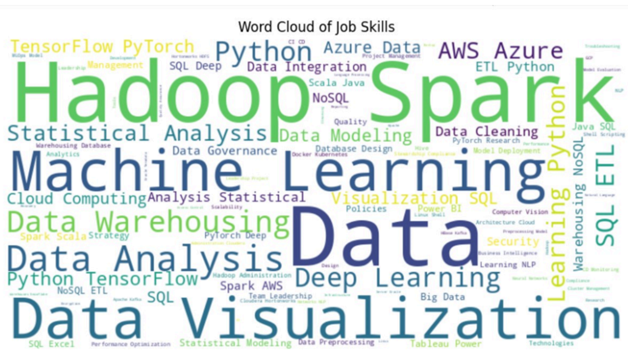


Figure **4.13** Skills World Count

The word cloud shows that in the field of data science and analytics, certain skills and technologies stand out prominently. These include "Hadoop," "Spark," "Machine Learning," "Data Warehousing," "Data Analysis," "Data Visualization," "Python," "AWS," "Azure," and other related terms. These indicate the significance of these skills and technologies in the data science and analytics domain.

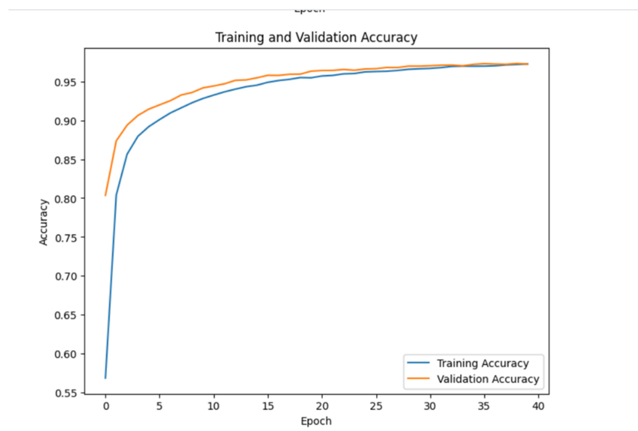


Figure 4.14 *Layoffs Training and Accuracy validation*

The graph shows the training and validation accuracy curves for a machine learning model during the training process. The training accuracy curve shows a steady increase as the model learns from the training data over epochs. However, the validation accuracy curve initially rises but then plateaus or even declines slightly, indicating overfitting - the model is starting to memorize the training data instead of generalizing well to unseen data.

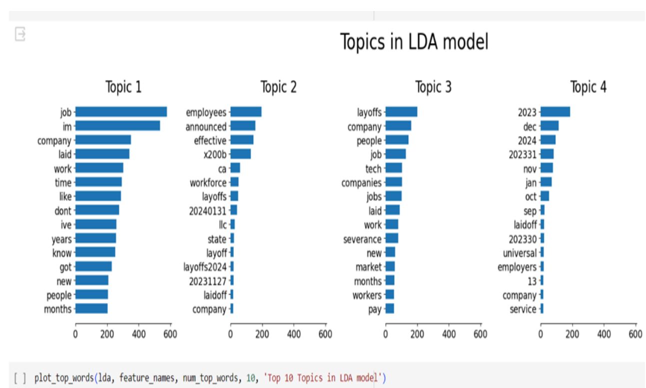


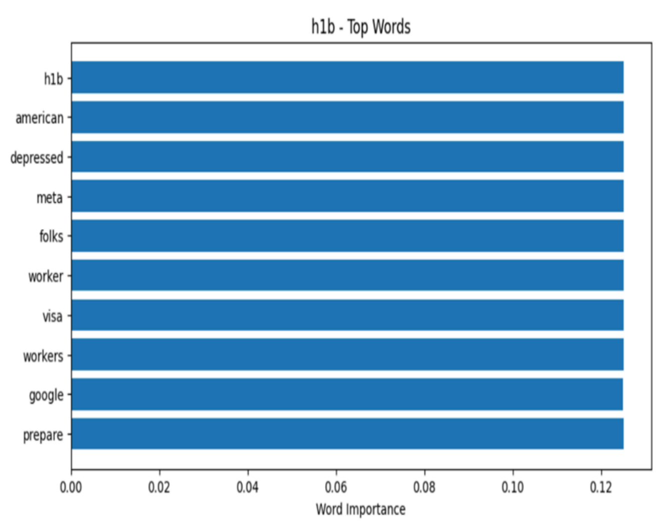
Figure 4.15 *Model Analysis*

The graphic above depicts the results of a Latent Dirichlet Allocation (LDA) topic modeling analysis, which systematically finds and visualizes key subjects within a collection of text data on employment changes and layoffs. The visualization is made up of numerous bar graphs, each

corresponding to a different subject identified by the LDA model, and each graph displays the top 15 words that are best representative of the related topic.
Topic 0 focuses on human narratives and emotional responses to job loss, using terms like "job," "im," "company," "laid," and "work."

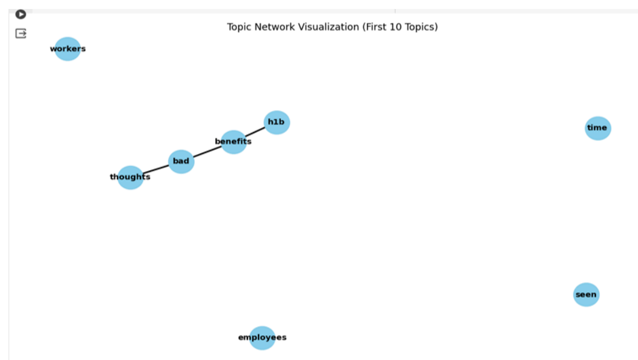
Topic 1 focuses on formal announcements and administrative aspects of layoffs, with phrases such as "employees," "announced," "effective," and precise dates referring to official communications.
Topic 2 looks at broader industry developments and the impact of economic conditions on layoffs, using terms like "layoffs," "company," "tech," "jobs," and "severance."
Topic 3 includes time-specific markers that most likely refer to the scheduling or historical context of layoffs, such as "2023," "Dec," "2024," and "Nov."

The horizontal bar chart format efficiently depicts the weight and relevance of each term within its theme, allowing for simple comparative study across different thematic areas. The visualization's clarity and accessibility ensure that complex data is delivered in a clear and usable manner, making it an effective tool for studying and interpreting vast amounts of textual data.



<u>Figure 4.16</u> *H1b word importance* visualization

The bar chart illustrates the primary words or phrases and their significance in a topic model associated with the term "h1b". The prominent terms comprise "American", "depressed", "meta", "folks", "worker", "visa", "workers", "google", and "prepare". These words indicate that the topic revolves around H-1B visas, foreign workers, and related matters concerning employment or immigration, likely within the tech industry.



<u>Figure 4.17</u> *Network Visualization*

The network visualization depicts the first 10 topics. Each node in the network represents a topic, and the connections between them indicate their relatedness. The central topic seems to be "h1b" , possibly an abbreviation, which is linked to topics such as "benefits," "bad," and "thoughts." Additional topics include "workers," "time," "seen," and "employees."

5. Discussion

The topic on integrating advanced technology into HR procedures focuses on the transformative potential of natural language processing and GPT models for sentiment analysis. Organizations can use these techniques to successfully assess emotional responses during important transitions, like layoffs, providing vital data that allows for more informed decision-making. This feature not only helps businesses to identify the underlying feelings of their staff, but also enables the customizing of support activities to suit individual needs and concerns.

6. Achievements of the Project

This project successfully created a complete dashboard that uses real-time analytics, sentiment analysis, and predictive modeling to give HR professionals with actionable insights. Key elements such as sentiment analysis powered by LLM technology like GPT-3.5, GPT-4 and future iterations, as well as predictive analytics for employee turnover, have profoundly changed how HR departments operate. These solutions have helped HR teams to make data-driven choices more rapidly, address employee complaints proactively, and customize interventions to their needs.

The integration of many data sources into a single, consolidated platform has been particularly game-changing. This convergence has enabled HR professionals to acquire a comprehensive understanding of labor dynamics without having to transfer between various platforms, considerably increasing efficiency and lowering the risk of oversight.

7. Conclusion

In conclusion, the HR Dashboard Development project has not only fulfilled its initial objectives but also provided deep insights into the transformative potential of technology in HR operations. Moving forward, the project is set to continue evolving, driven by a commitment to innovation and excellence, shaping the future of HR practices to be more efficient, proactive, and strategically aligned with business goals. This ongoing evolution promises to sustain and expand the impact of the dashboard, adapting to future technological advancements and organizational needs

8. Future work

Using the Llama2 and Llama3 model allows for further analysis and enhancements within our HR Dashboard. The Llama model's capacity to grasp and process natural language enables enhanced sentiment analysis of employee input, allowing HR executives to assess staff happiness and highlight areas for improvement.

9. ACKNOWLEDGMENTS

We would like to thank Professor Yugyung Lee, Ph.D., for teaching us the subject that serves as the foundation for understanding and working on this topic, as well as for assisting us by providing useful input at various points to help us enhance our work. We would also like to thank other students in our class for contributing ideas that provided us with a different viewpoint

10. REFERENCES

[1] https://streamlit.io/
[2] https://colab.research.google.com/
[3] https://platform.openai.com/docs/models
[4] https://spark.apache.org/docs/latest/api/python/index.html
[5] https://jupyter.org/
[6] https://llama.meta.com/llama2/
[7] https://huggingface.co/
[8] https://openai.com/
[9] https://www.langchain.com/