
Humanlytics: An HR Dashboard Software

Muhammed Arslan
Muhammed Tayyab
Dallas Diaz
Ramya Thambabattula

Introduction

The main issue in today's work world seems to be turnover rates. It can be attributed to COVID, employee satisfaction, benefits, etc.

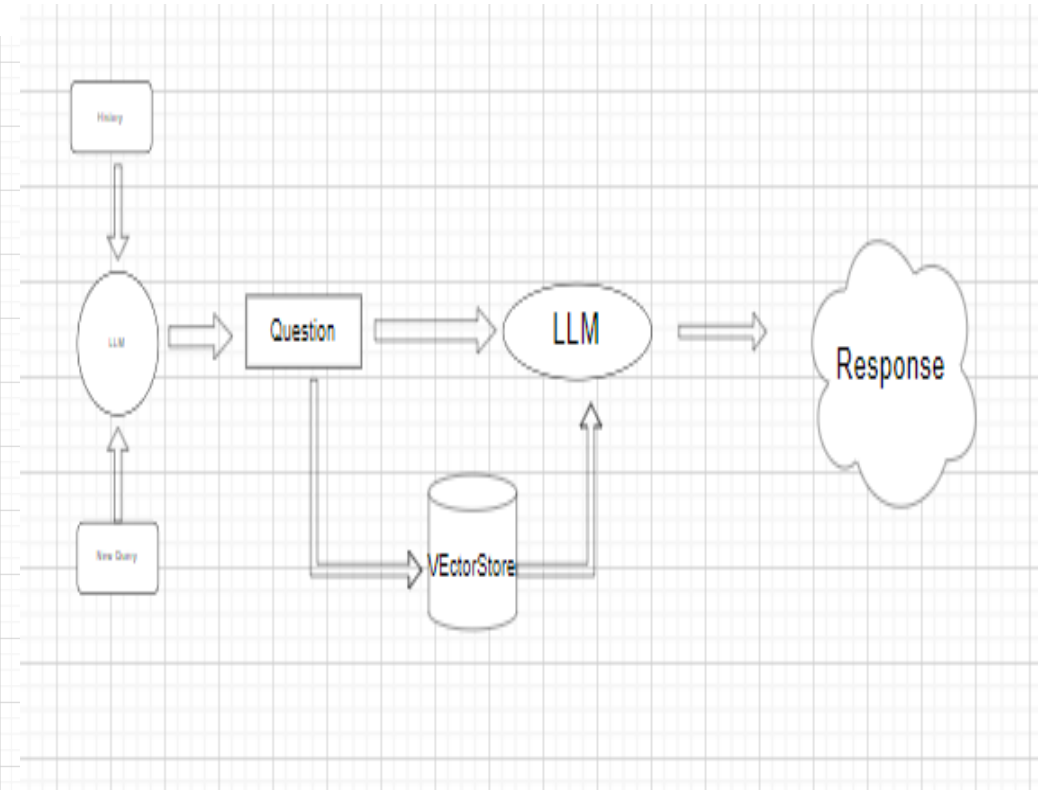
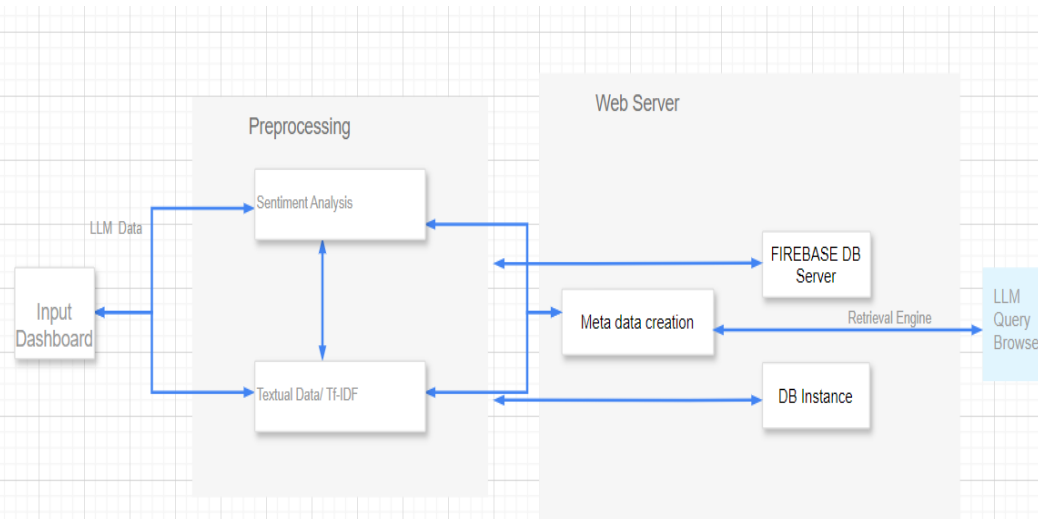
Not all HR teams within companies know these things so our product should shed more light on what is affecting their turnover rates and how they can get those down.

- **Objectives:**

- Focusing on:

- Talent acquisition
 - 1. How can we find the best of the best employees
 - Employee retention
 - 1. Looking at factors that make employees stay
 - Workforce analytics
 - Layoff Analysis
 - Sentiment on layoffs and potential suggestions



Conceptual/Model Diagram







Specific Data being used

The majority of our data for the initial information comes from Reddit and Twitter.

Listed here are the actual data sets from our google drive

 layoffs_clean.csv 

 tweets_clean.csv 

 tweets_sentiment.ipynb 

 reddit_sentiment_analysis.ipynb 

Data Preprocessing and Cleaning

Help All changes saved

+ Code + Text

RAM
Disk

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 x 9 columns

All of the data (coming from reddit and twitter) that was included in the database was retrieved using an API call and cleaned using a jupyter notebook

Once complete, we import the data into the firebase database.

Firebase Data Importer

Upload CSV file

Drag and drop file here
Limit 200MB per file • CSV

Browse files

clean_data.csv 165.1MB

Enter Firestore collection name

skills

Add Data

```

✓ [5] # Load the CSV file into a DataFrame
0s df = pd.read_csv('tweets.csv')

```

✓ 3s df

	ed	user_followers	user_friends	user_favourites	user_verified	date	text	hashtags	source
11	00	2836.0	2469.0	870	False	2022-12-20 18:00:48+00:00	today!\nhttps://t.co/eCRVPFBYiy\n\n#app...	['apply', 'hiring', 'opentowork', 'techlayoffs...]	Hootsuite Inc.
11	00	2835.0	2469.0	870	False	2022-12-20 18:00:48+00:00	today!\nhttps://t.co/eCRVPFBYiy\n\n#app...	['apply', 'hiring', 'opentowork', 'techlayoffs...]	Hootsuite Inc.
27	00	19.0	32.0	4	False	2022-12-18 00:21:03+00:00	I want to wish all @Twitter employees that got...	NaN	Twitter Web App
17	00	81.0	287.0	4484	False	2022-12-17 14:04:44+00:00	@elonmusk sir has call me to work from office....	['TwitterLayoffs']	Twitter for Android
28	00	1044.0	777.0	3742	False	2022-12-16 19:47:58+00:00	Here's what happens when you fire all your eng...	['TickTok']	Twitter for iPhone
...	

```
[ ] df.columns
```

```

Index(['user_name', 'user_location', 'user_description', 'user_created',
      'user_followers', 'user_friends', 'user_favourites', 'user_verified',
      'date', 'text', 'hashtags', 'source'],
      dtype='object')

```

```
[ ] df
```

text

0	want wish twitter employees got fired happy ho...
1	curtis415 adamparkhomenko meidastouch elon sno...
2	wkamaubell even made local news tonight one wo...
3	filed yet another case twitter today twitterla...
4	new social communication app android ios lets ...
...	...
13223	missed flight yesterday queuing 4h booked tomo...
13224	bplwijn long queues indications waiting times ...
13225	schiphol disaster security went strike summer ...
13226	need fly schiphol today arrive early allowed 4...
13227	delta yall stop booking flights schiphol wheww...

13228 rows × 1 columns



ui



	Title	Content
0	Everyone laid off in my tech company this week..	My tech company was bought by another company ...
1	No one thinks this has anything to do with lay...	I'm 7 months into my search and I've definitel...
2	This is why layoff have consequences	\nhhttps://www.cnn.com/2024/02/22/tech/att-cell...
3	Corporate America Has Abused the H1B Visa Program	The program was created to bring in experts bu...
4	I am done with tech.	This field does not bring joy but rather immen...
...
994	Laid off and insurance terminated immediately.	Is this common? I thought my family and I woul...
995	Anaconda	Approx 10% of staff laid off \n\nLaid off role...
996	Healthcare heroes no more...	NextCare reduced 538 people today. CEO apparen...
997	Tech layoffs	With news of tech layoffs coming everyday,I se...
998	Severance pay	Does anyone have experience with severance pay...

```
# Filter out short texts (less than 3 words)
df = df[df['Content'].apply(lambda x: len(x.split()) > 2)]
```

[] df

	Title	Content
0	Everyone laid off in my tech company this week..	tech company bought another company late 22 wo...
1	No one thinks this has anything to do with lay...	im 7 months search ive definitely seen dark ho...
2	This is why layoff have consequences	httpswwwcncom20240222techattcellserviceoutage...
3	Corporate America Has Abused the H1B Visa Program	program created bring experts program heavily ...
4	I am done with tech.	field bring joy rather immense stress cycle la...
...
994	Laid off and insurance terminated immediately.	common thought family would covered end month ...
995	Anaconda	approx 10 staff laid laid roles include produc...
996	Healthcare heroes no more...	nextcare reduced 538 people today ceo apparent...
997	Tech layoffs	news tech layoffs coming everydayi see lot pos...
998	Severance pay	anyone experience severance pay calc approved ...

Skills Data

Data scraped for skill matching with the job titles for recommendations.

Deploy

Best-Fit Skills for Job Title

Enter Job Title

Get Best-Fit Skills

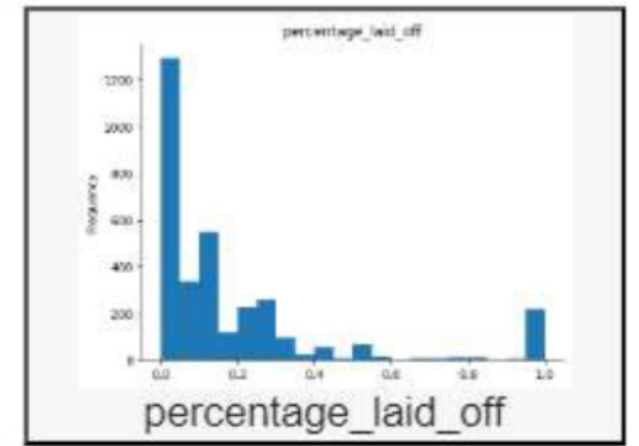
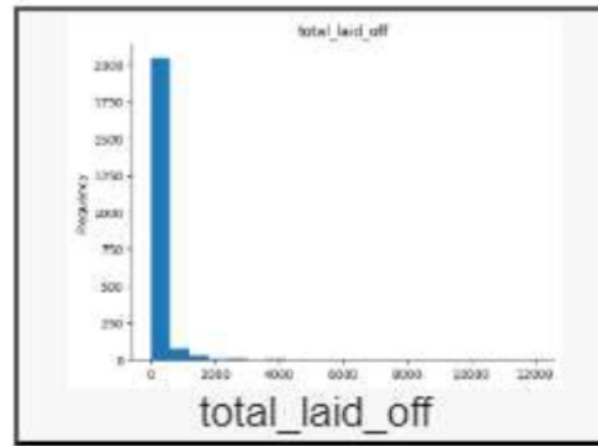
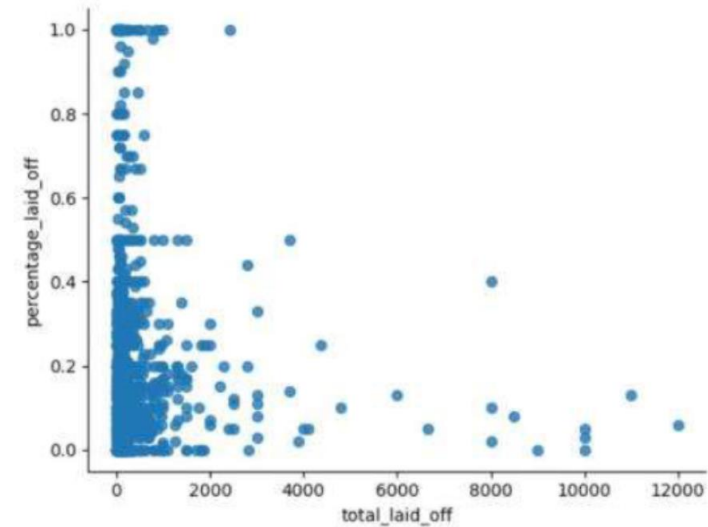
```
df = pd.read_csv(dataset)
df.head()
```

	Job Role Title	Skills Required
0	Data Engineer	Hadoop, Spark, SQL, Python, ETL, Big Data...
1	Machine Learning Engineer	Python, TensorFlow, PyTorch, Machine Learni...
2	Data Scientist	Python, R, Machine Learning, Statistical An...
3	Big Data Architect	Hadoop, Spark, Data Modeling, NoSQL, Cloud...
4	Data Analyst	SQL, Excel, Data Visualization, Statistics,...

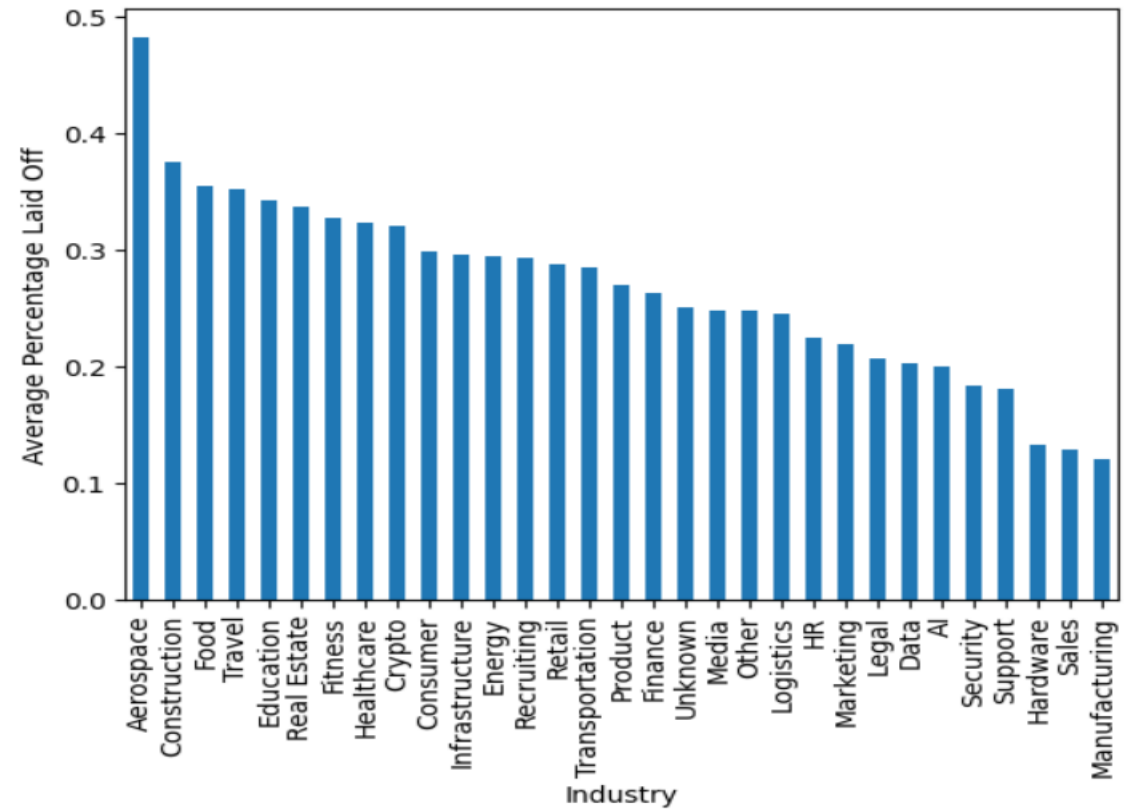
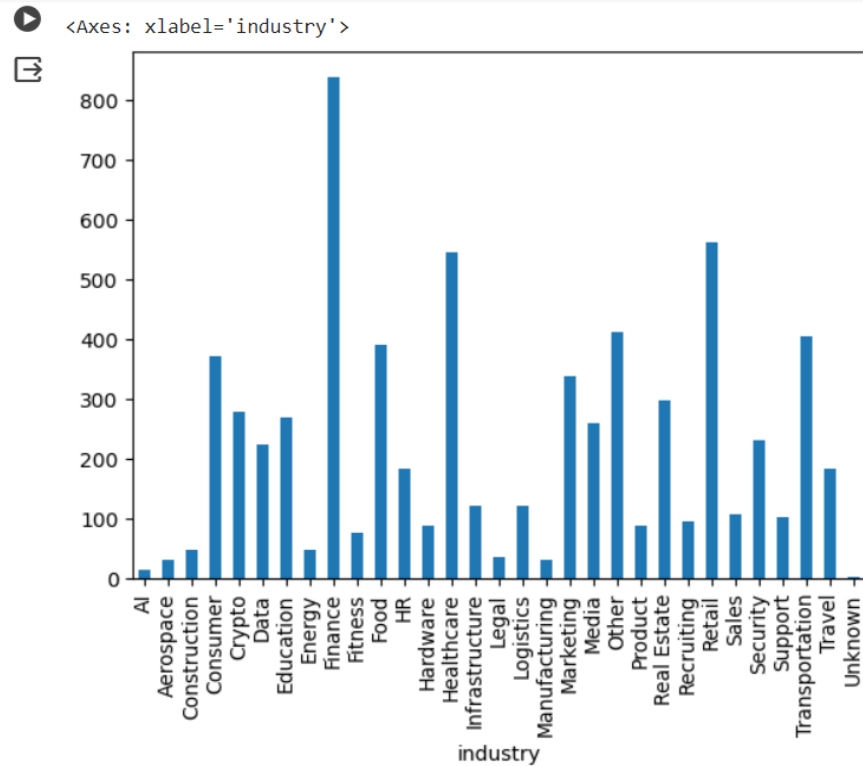
Data Analysis

Initial findings were in the form of graphs and tables before we could utilize our LLM to take the information and make recommendations

```
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)
```



Layoff Data

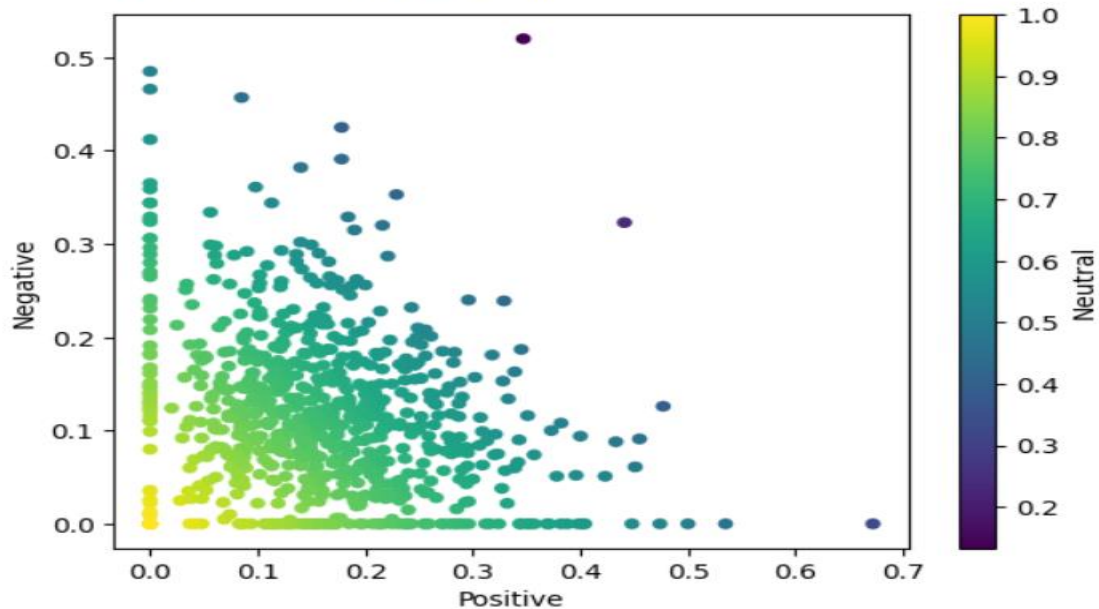


Reddit Data

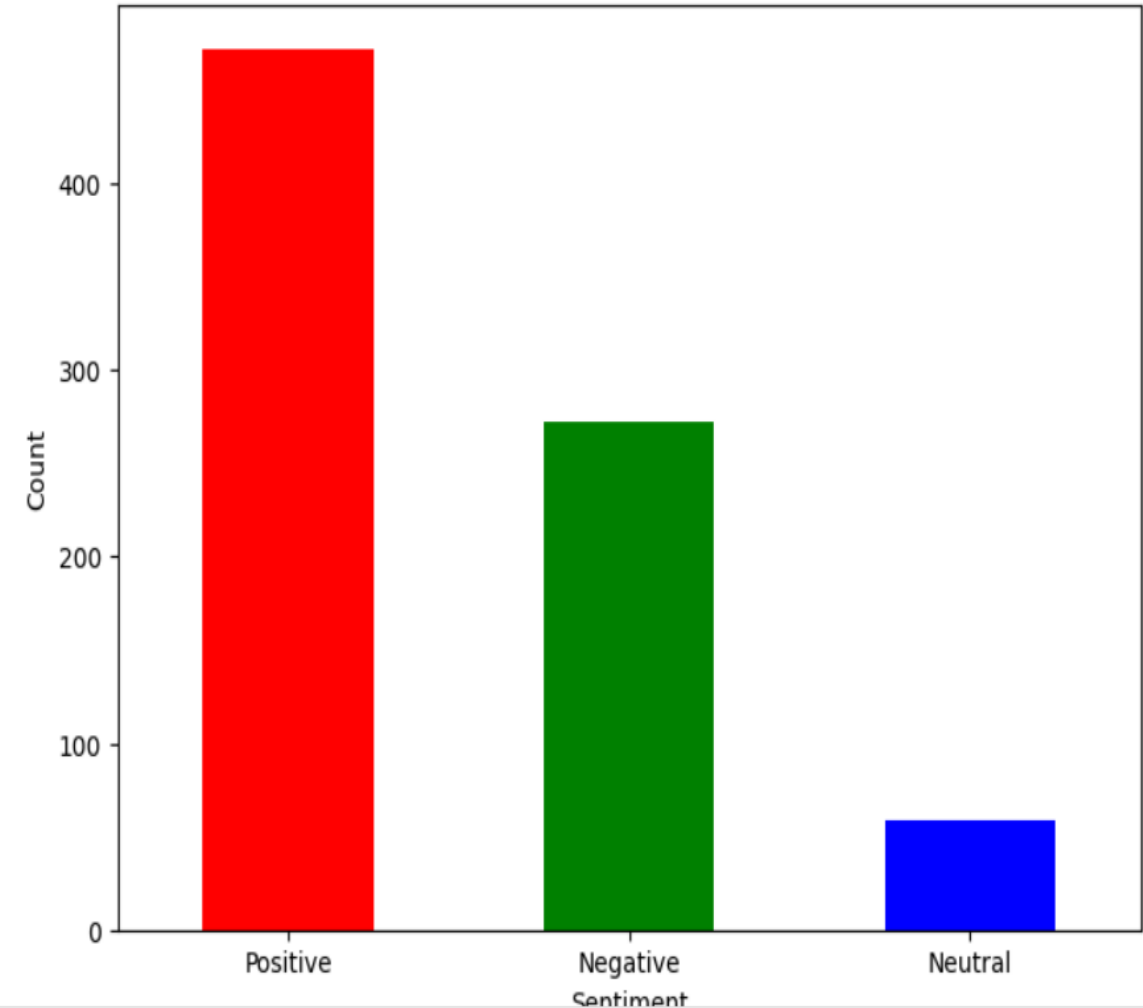
```
[ ] # Initialize the VADER sentiment analyzer
sia = SentimentIntensityAnalyzer()

[ ] # Compute sentiment scores for each submission
sentiment_scores = df['Content'].apply(lambda x: sia.polarity_scores(x))
```

<Axes: xlabel='Positive', ylabel='Negative'>



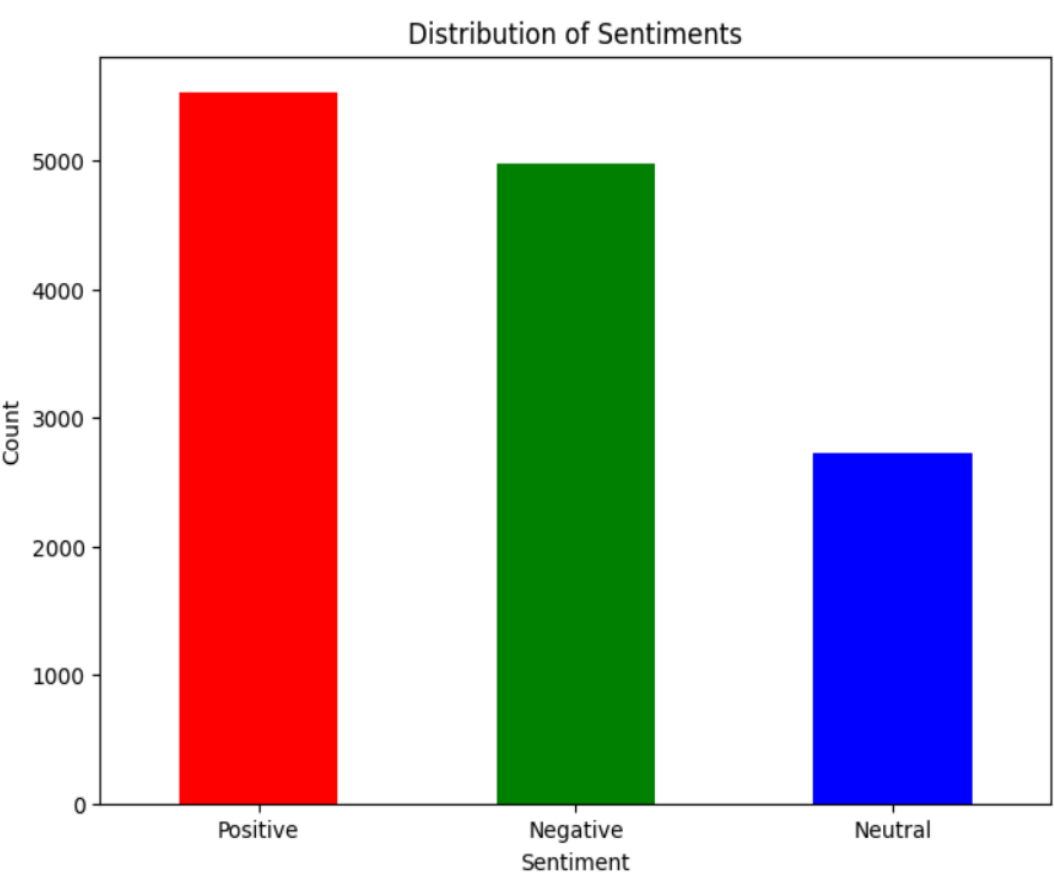
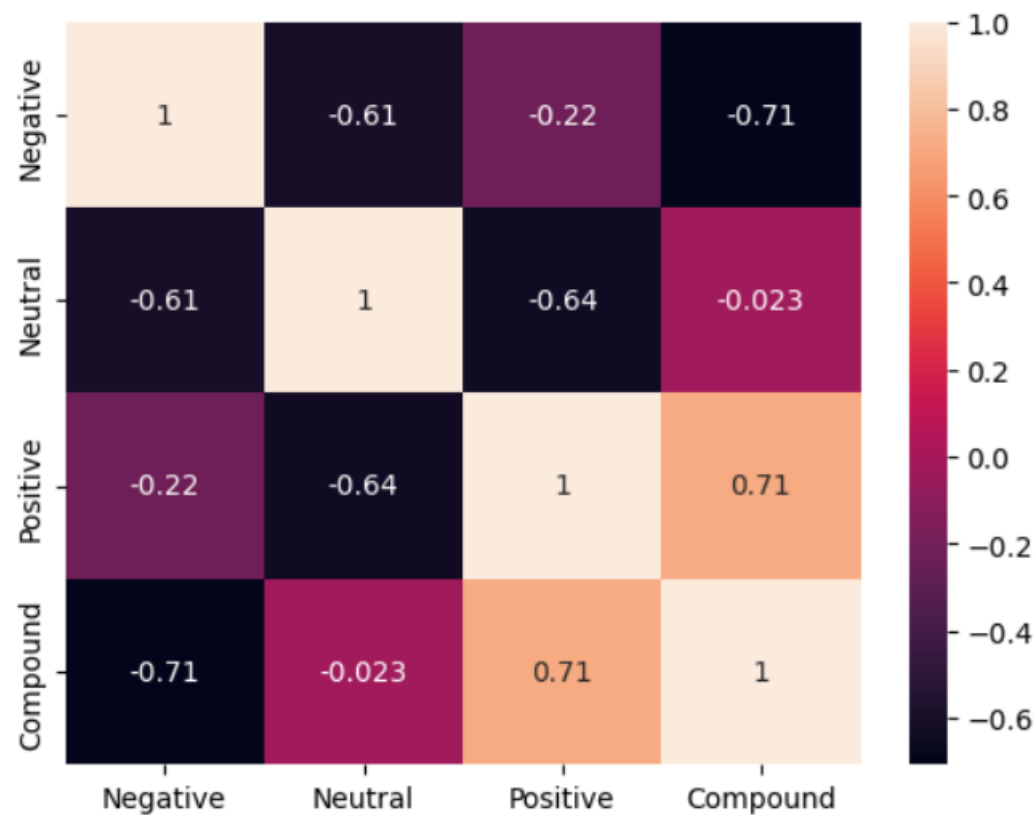
Distribution of Sentiments



Tweets Data

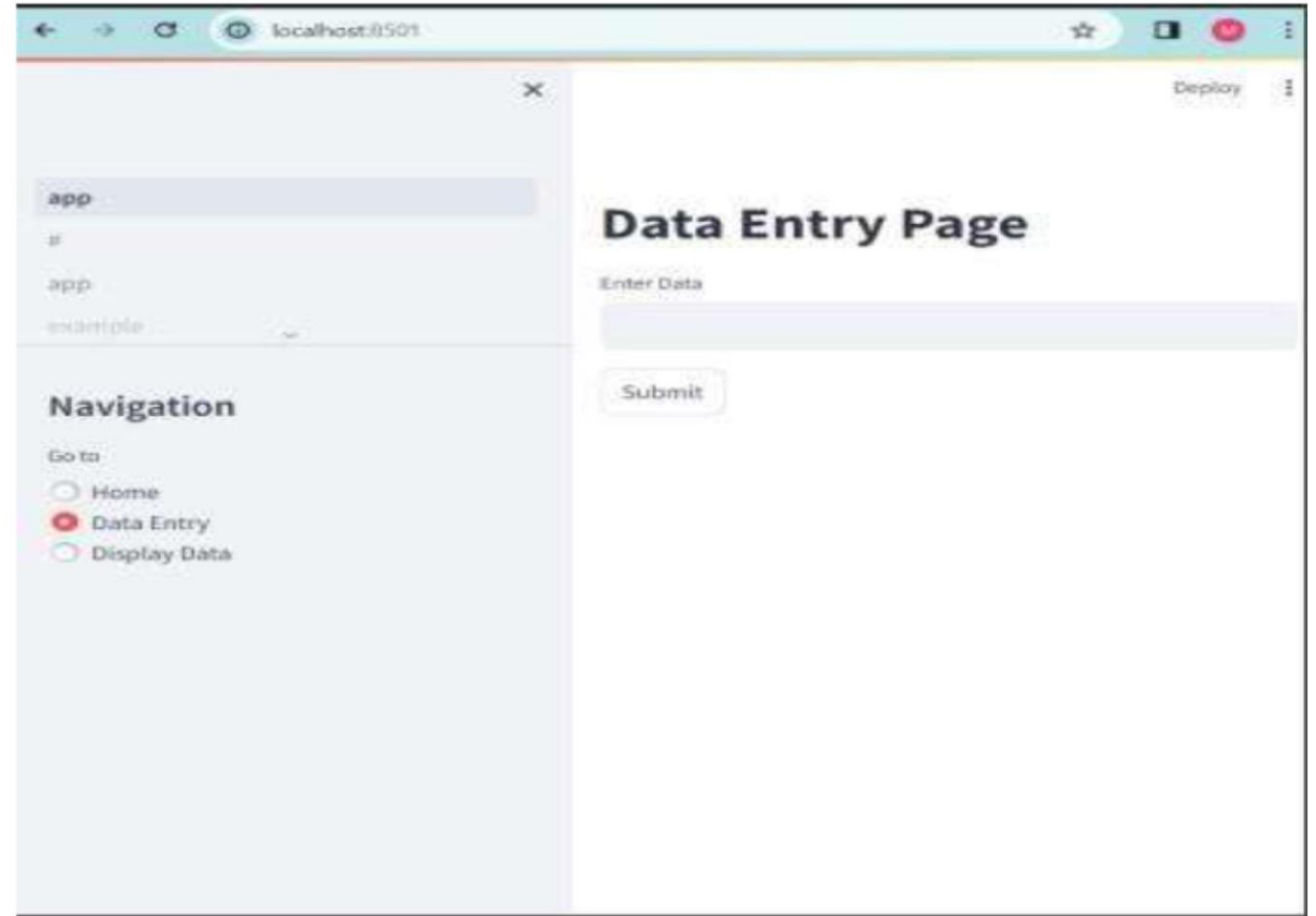
	text	Sentiment	Negative	Neutral	Positive	Compound
0	want wish twitter employees got fired happy ho...	Positive	0.122	0.441	0.437	0.8074
1	curtis415 adamparkhomenko meidastouch elon sno...	Positive	0.193	0.525	0.282	0.6908
2	wkamaubell even made local news tonight one wo...	Negative	0.291	0.709	0.000	-0.6249
3	filed yet another case twitter today twitterla...	Neutral	0.000	1.000	0.000	0.0000
4	new social communication app android ios lets ...	Neutral	0.000	1.000	0.000	0.0000

<Axes: >



Dashboard Visualization

We were also able to get our streamlit up and running to start presenting something to a user



Model Implementation

Our current updates to the model that we are using:

- Our ChatGPT model can retrieve data from our database
- It is fully connected within our triangle model
- It has been trained to make recommendations based off the information it receives.


```
capstone_project > pages > Resume.py > {} st
46 def read_docx(file):
47     text = ""
48     doc = Document(file)
49     for paragraph in doc.paragraphs:
50         text += paragraph.text
51     return text
52
53 # Function to generate analysis using GPT-3.5 API
54 def generate_analysis(resume_text, job_description):
55     # Define messages including user questions and system responses
56     messages = [
57         {"role": "system", "content": "You are a helpful assistant."},
58         {"role": "user", "content": resume_text},
59         {"role": "assistant", "content": job_description},
60     ]
61
62     response = client.chat.completions.create(
63         model="gpt-3.5-turbo", # Choose the GPT-3.5 model
64         messages=messages,
65         max_tokens=150, # Adjust based on your desired length of response
66         n=1, # Number of responses to generate
67         stop=None, # Tokens to stop generation
68         temperature=0.5, # Control the randomness of the output
69     )
70
71     return response.choices[0].message.content.strip()
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
```

PROBLEMS OUTPUT DEBUG CONSOLE TERMINAL PORTS

PS C:\Users\kckfo\OneDrive\Desktop\spring24\Capstone\capstone_project> streamlit run app.py

Resume Analyzer

Upload Resume (PDF or DOCX)

 Drag and drop file here
Limit 200MB per file

Browse files

 Muhammad_Arslan_Resume.docx 20.0KB



Enter Job Description

- Onsite: Employees who are working in Onsite roles will work primarily onsite
- This includes all production and maintenance employees, as they are essential to the development of our products

Analyze Resume

Initial Results

After the triangle model was completely connected, we were able to implement the features of the dashboard.

Current features:

- Resume Analyzer
- Skills Match
- Sentiment Analysis

Resume Analysis

Based on your experience and skills, you may be a good fit for roles such as Software Quality Assurance Engineer, Data Analyst, Machine Learning Engineer, Full Stack Developer, or IT Project Manager.

Your proficiency in programming languages such as R, Python, C++, Java, and JavaScript, along with your experience in machine learning, data processing, and modeling, make you well-suited for positions that require strong technical skills. Additionally, your experience with tools and frameworks like PySpark, TensorFlow, and MySQL can be valuable in roles that involve working with large datasets and implementing machine learning algorithms.

Your background in software engineering, database management, and web design, coupled with your experience in software quality assurance testing and agile methodologies, positions you well for roles in software

17

Initial Results

Here is the outcome of the Resume Analysis Functionality Feature:

- 1.The Dashboard takes in the input from the user
 - 2.The LLM takes the information and makes a call to the database to retrieve the information based on the request
 - 3.It finds a match and returns the output to the user on the front end
-

Next Steps

- Finalize the information that is being input into the Firebase database
- Clean up the front-end user interface and make it more user friendly and user appealing
- Future Steps:
 - Expand our network to include additional data and make additional recommendations
 - As more users interact with the dashboard, we will gather more and more information to eventually expand our recommendation features to make recommendations about other items.