Qualetics
DATA MACHINES

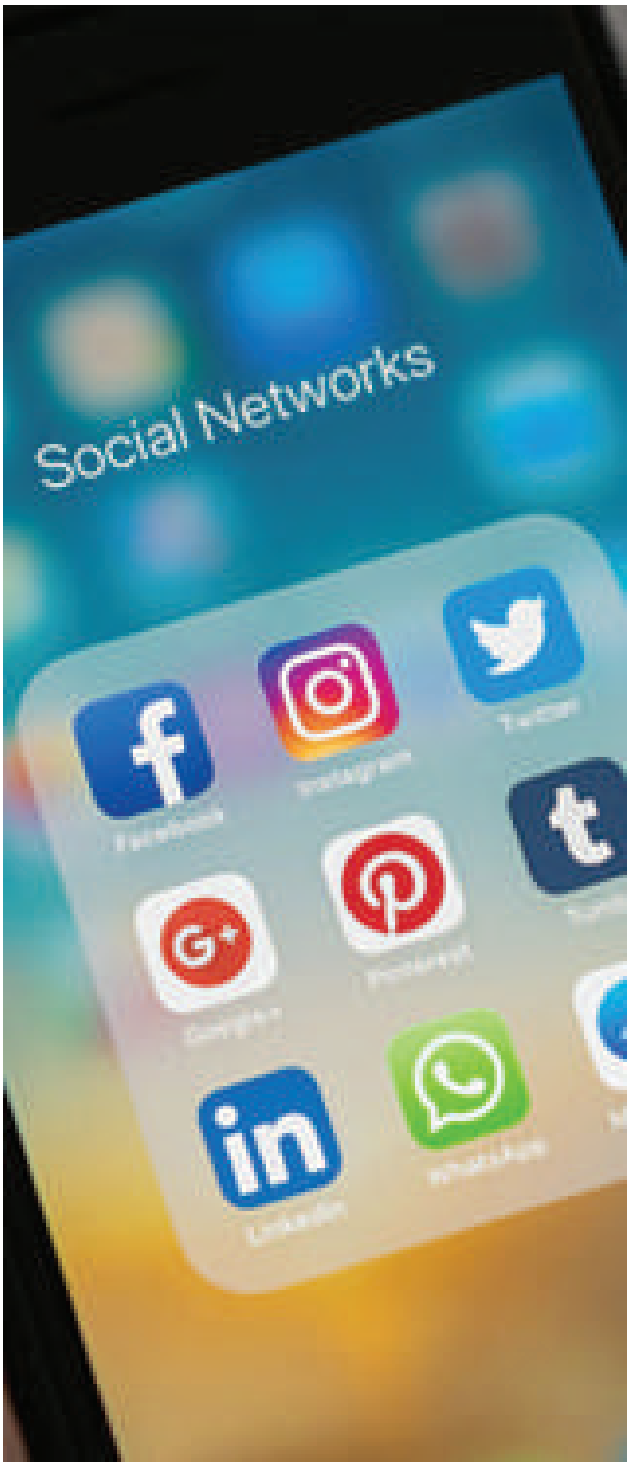
Use Case

Analyzing Customer Support on Social Media

Using Machine Learning and Deep Learning

Introduction

Social customer support is the practice of providing customer service on social media sites like facebook, twitter, etc. These social platforms have evolved to become important and valid for customers soliciting and receiving customer support through these platforms.



There are around 3.5 billion users globally who are actively using various social media platforms. A business, to be successful, should meet its customers wherever they are. Therefore, businesses venturing into social media is no more an option but a must step as they help grow businesses by boosting their brand awareness, marketing their products or services, and listening to their customers.

Artificial Intelligence technologies like Machine Learning, Deep Learning, etc prove to be powerful tools to optimally leverage social media. Sentiment analysis uses NLP and Machine Learning and helps in collecting feedback on a new product or design from social media and customer support. Image recognition helps businesses when their customers upload images of products on social media. Chatbots help companies by automating customer service without requiring human interaction.

> "A brand is no longer what we tell the consumer it is. It is what consumers tell each other it is."
>
> – Mark Schaefer, leading Marketing expert & Keynote Speaker

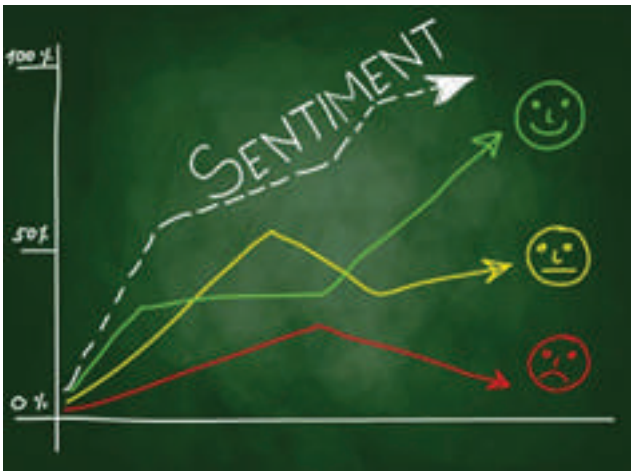
Problem Statement

Social media is about interacting with your customers in real-time and getting feedback from them. Hence, understanding and interpreting the right intent of the customer is critical. Getting the right information to the right customers through social media is challenging. Understanding the high volume of unstructured information generated by social media channels helps businesses by providing deep insights into customer behavior and preferences.

Therefore, we tried to analyze the queries raised by customers on a particular social media platform (Twitter) and come up with possible insights on which sentiment analysis can be performed.

About the Dataset

The Dataset consists of over 7,94,299 queries sent by 5,97,075 customers to 108 different companies. These queries are related to issues of customer support services a particular customer is facing. The dataset contains data on the queries raised by the customers, their IDs, and the time when they have sent that query. Similarly, it includes the reply from the company recorded in the timestamp. To view the dataset, click here.



Assumptions:

This data could be regarded as a Sentiment Analysis problem.

Solution provided

We worked with a complete dataset and considered a sample of it whenever required to get insights for a particular company. After studying the data carefully, we noted the most important columns for our analysis i.e., the queries raised by the customers and the response given by the customer support service of the respective company.

As social media has become one of the biggest platforms for the expansion and branding of businesses, we decided to make use of Machine Learning and Deep Learning technologies and analyze the social customer support.

We then tried to clean the dataset with the help of NLTK to find a compound score for most occurring words. NLTK is a set of libraries and programs that work with human language data for applying in statistical NLP. The NLTK consists of text processing libraries for tokenization, parsing, classification, stemming, tagging, and semantic reasoning. This will help us in understanding customer activities, opinions, and feedback.

Technologies Used

Data Preprocessing

Deep Learning

Machine Learning

Exploratory Data Analysis (EDA)

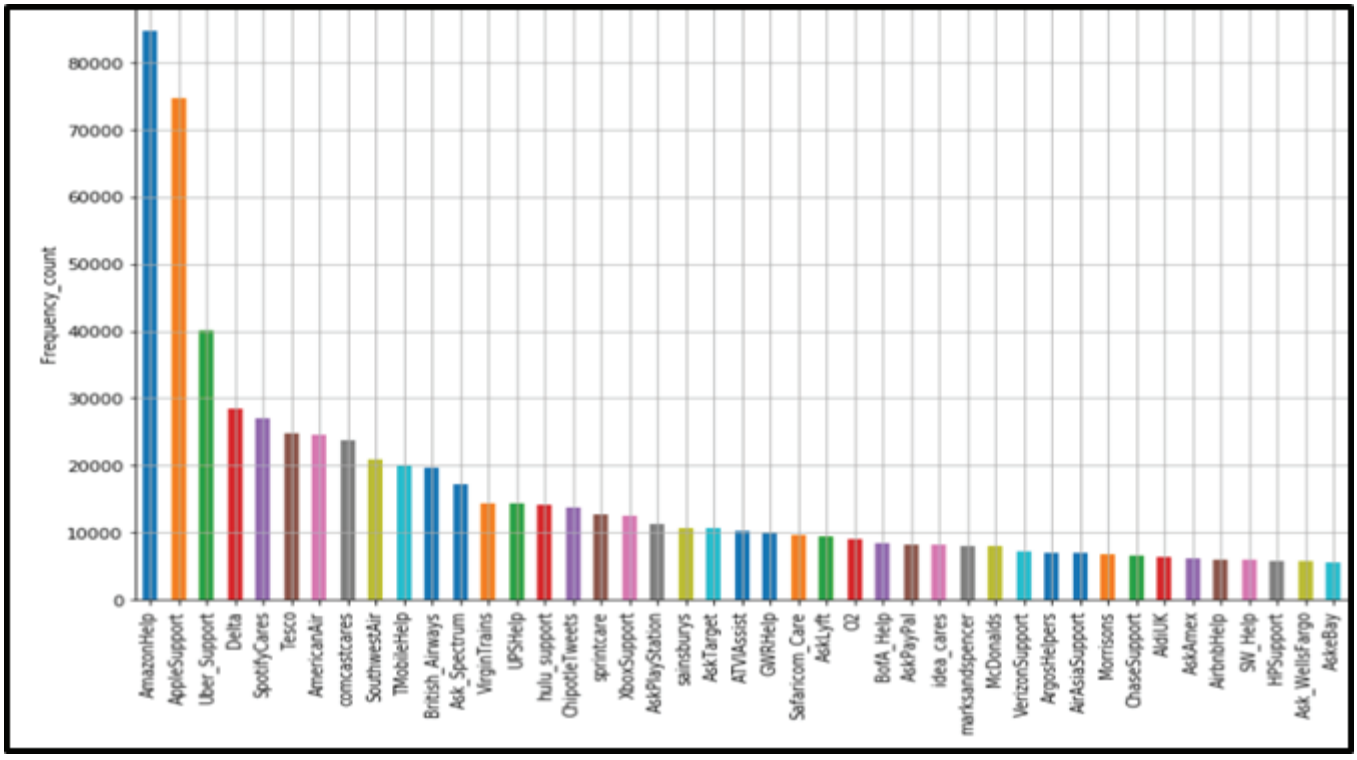
Natural Language Tool Kit (NLTK)

Our Process

- We started our study by analyzing and cleaning the data. This involves checking for null values, dropping unnecessary columns, and transforming the unstructured data to structured data. We dropped the columns that are not crucial in our analysis. (e.g. ids).

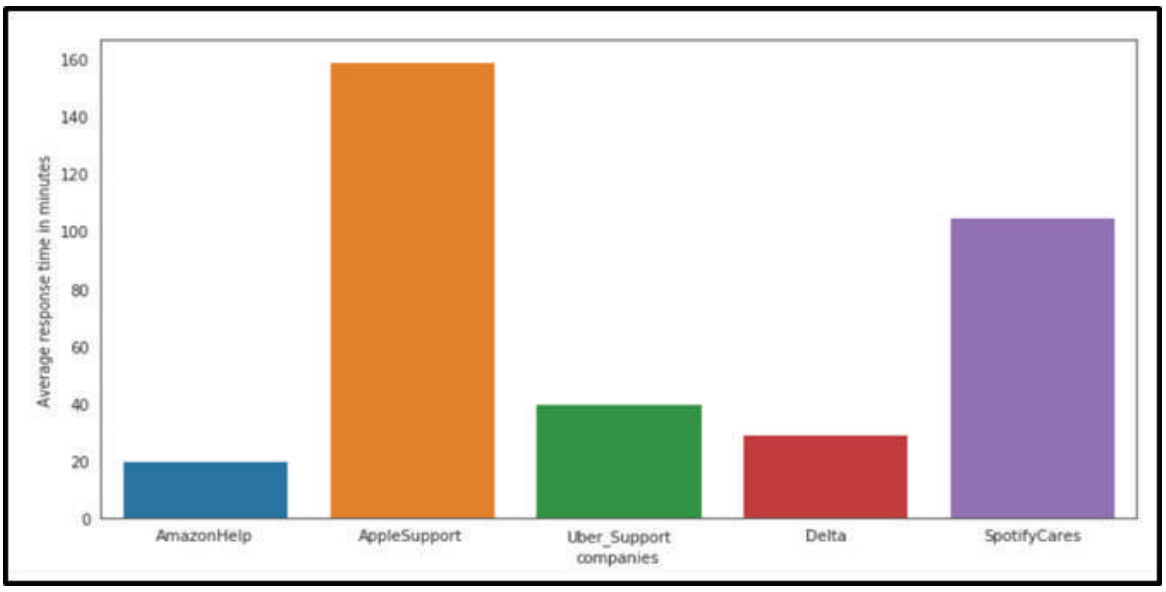
- We then performed EDA to extract insights from the data. We extracted the most frequently used words in the queries.
- Later, we identified emojis sent in the queries. Emojis have to be analyzed to ascertain the true intent of the customer and inspect the presence of sarcasm in the query.
- We then performed a cluster analysis of the data. We created a group of similar queries/data.
- Finally, we performed model building to identify the group for future queries.

Number of queries to each company is shown below



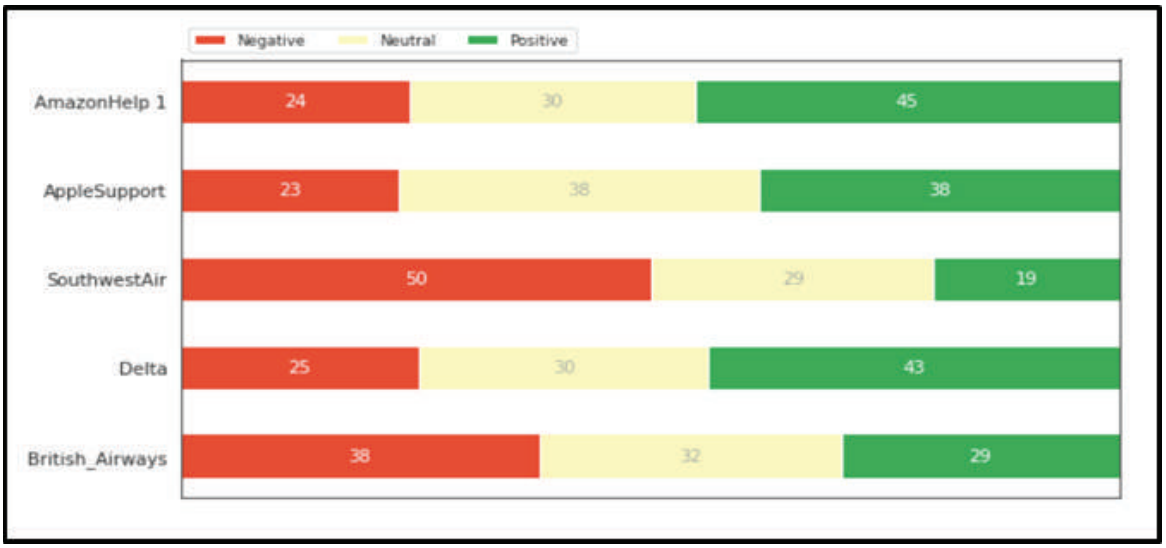
Inference - Looking at the plot, it can be inferred that most of the queries are directed towards AmazonHelp followed by the AppleSupport, Uber_Support, Delta and so on.

Time taken by companies to respond to the query



Inference - This plot gives us the average time taken by the social customer support service team of the particular company to respond to the customers and their queries.

Sentiment Analysis of Customer queries



Inference - The above plot depicts the percentage of negative, neutral and positive sentences sent by the customers in their queries to the respective company. We can see that Southwest Air has the most number of negative sentences and AmazonHelp has the most number of positive sentences given by their respective customers.

Emoji Detection on AmazonHelp

Considering AmazonHelp as an instance, we are displaying the most frequently used emojis in each query raised by the customers.

	word	abs_freq	wtd_freq	rel_value	name	group	sub_group
0		535	535	1.0	pouting face	Smileys & Emotion	face-negative
1		228	228	1.0	face with rolling eyes	Smileys & Emotion	face-neutral-skeptical
2		155	155	1.0	thinking face	Smileys & Emotion	face-hand
3		140	140	1.0	unamused face	Smileys & Emotion	face-neutral-skeptical
4		113	113	1.0	middle finger	People & Body	hand-single-finger
5		99	99	1.0	face with steam from nose	Smileys & Emotion	face-negative
6		88	88	1.0	face with tears of joy	Smileys & Emotion	face-smiling
7		81	81	1.0	loudly crying face	Smileys & Emotion	face-concerned
8		80	80	1.0	face with symbols on mouth	Smileys & Emotion	face-negative
9		72	72	1.0	expressionless face	Smileys & Emotion	face-neutral-skeptical
10		69	69	1.0	weary face	Smileys & Emotion	face-concerned
11		56	56	1.0	upside-down face	Smileys & Emotion	face-smiling
12		46	46	1.0	litter in bin sign	Symbols	transport-sign
13		43	43	1.0	middle finger: medium-dark skin tone	People & Body	hand-single-finger
14		38	38	1.0	angry face	Smileys & Emotion	face-negative
15		37	37	1.0	speaking head	People & Body	person-symbol

Word Cloud

The word cloud is a data visualization technique used to represent text data. The size of each word indicates its importance or frequency of its appearance. It helps to understand the most frequently used words. Word clouds are usually used to analyze data generated from social media platforms.

We tried to identify the most frequently used words by the customers in their queries for the respective company. We took the top 5 companies for which most of the queries were raised -

AmazonHelp

Apple Support

Uber Support

Delta

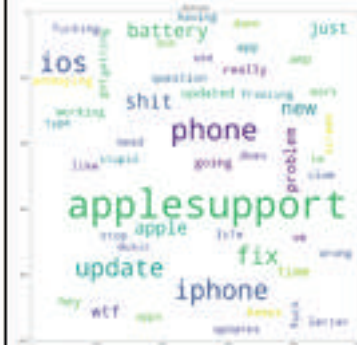
SpotifyCares

We have 3 word clouds namely positive, negative and neutral for each of the above-mentioned companies.

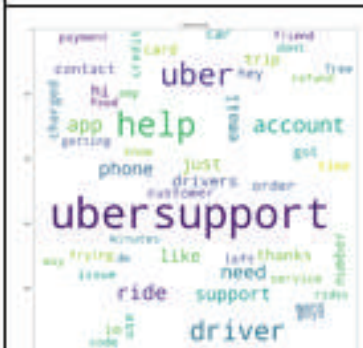
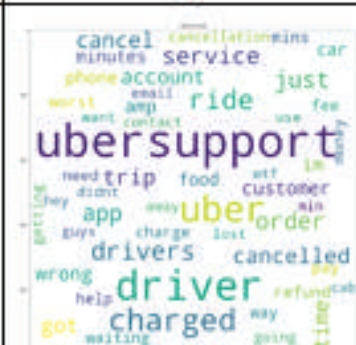
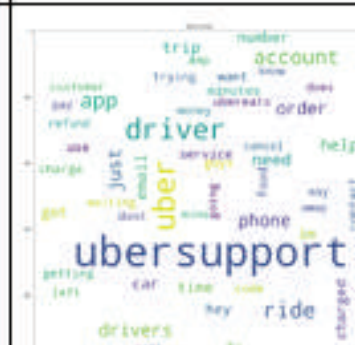
AmazonHelp

Positive	Negative	Neutral
		

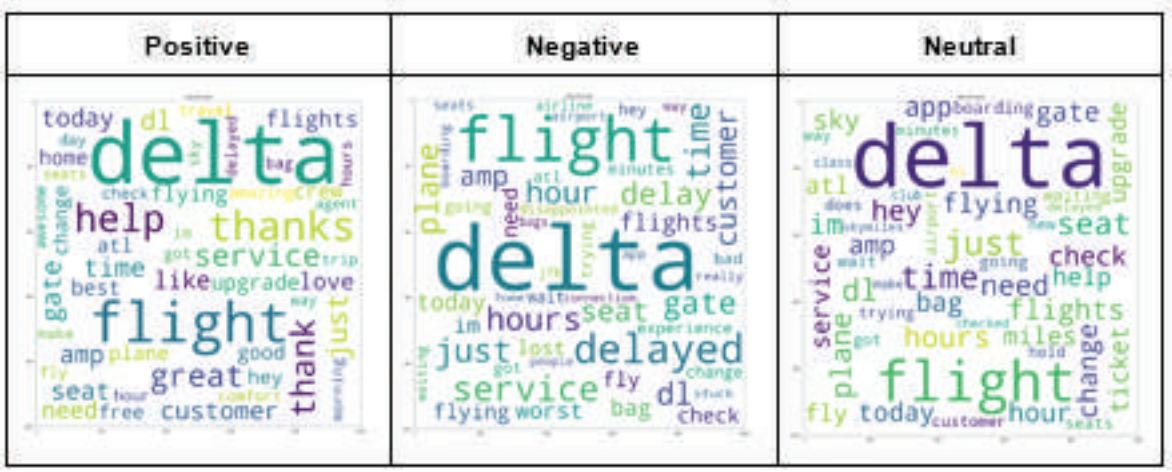
Apple Support

Positive	Negative	Neutral
		

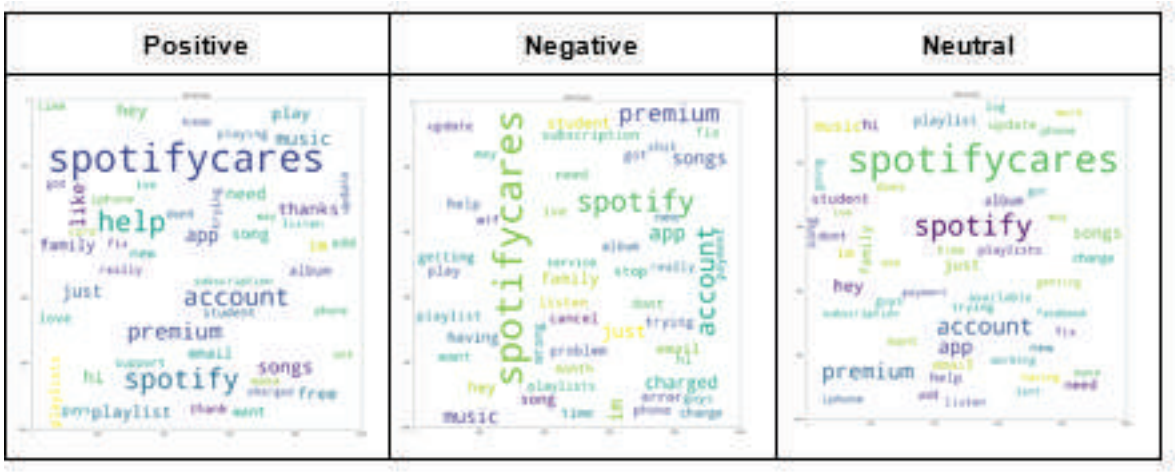
Uber Support

Postive	Negative	Neutral
		

Delta



Spotify Cares

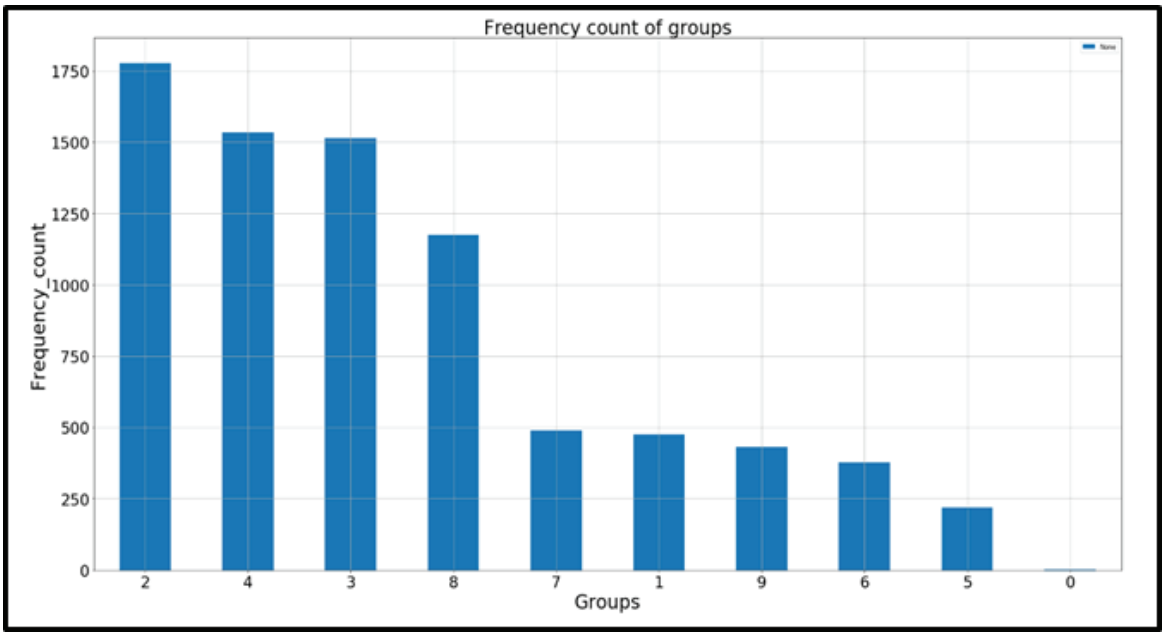


Companies and number of queries raised

116421~VirginTrains	91
317609~airtel_care	76
188231~Delta	57
121239~LondonMidland	56
121325~GWRHelp	55
118423~airtel_care	49
531342~comcastcares	47
119784~Kimpton	39
130848~airtel_care	37
133091~Safaricom_Care	37
125464~ChipotleTweets	34
130560~VirginTrains	33
407042~Safaricom_Care	32
127850~AskeBay	30
217320~sprintcare	30

The above image tells us the total number of queries raised for a particular id to a given company.

Cluster Analysis (Frequency count for a group of similar text)



Inference - The above plot is the cluster analysis of McDonalds, here taken for illustration purposes. We are displaying the group having the most frequency, after clustering, based on the similarity of the text that appeared in the queries.

K-Mean Algorithm - K-means is an unsupervised and iterative learning algorithm that is mostly used for clustering i.e making the clusters(groups) based on similarity or distance between two things. Using this technique we created 10 groups based on the similarity in the text used in queries, giving each group labels from 0 to 9. This changed the unlabeled data to labeled data.

Convolution 1-D Algorithm - Convolution 1-D (CNN) is a very effective algorithm of deep learning networks to derive interesting features from shorter (fixed-length) segments of the data set. After performing the clustering technique, we got the labeled data that was used to create a CNN-1D model. The model successfully classifies the group for the new unlabeled queries.

Result -

Convolution 1-D Classifier

Accuracy achieved - 76%

About Qualetics Data Machines

Qualetics Data Machines is a venture-funded startup established in July 2018, in the state of New Jersey in the USA. The company was founded and is led by the CEO, Sumanth Vakada, a software industry veteran with 20-years of experience in building modern software applications across various business domains. Qualetics currently has its core team based in the USA and an extended Application Development and Data Science team based in Hyderabad, India.

Our mission is to enable organizations to make an easier transition into the fields of Data Analytics and Artificial Intelligence. We aim to achieve this more efficiently than an organization might expect, were it to invest in the manpower and technology to build such complex platforms. For a universal need such as Data Science and AI, a dedicated focus is absolutely critical and Qualetics empowers our clients with the knowledge that allows them to apply in their core business offerings.

Key Features

Data Ingestion Platform

Visual Insights Delivery

License Pre-Existing Models

API based Intelligence Delivery

Integration with Apps

Develop Custom Models & Insights

For Inquiries, please contact:

Mike Fowler

Chief Commercial Officer

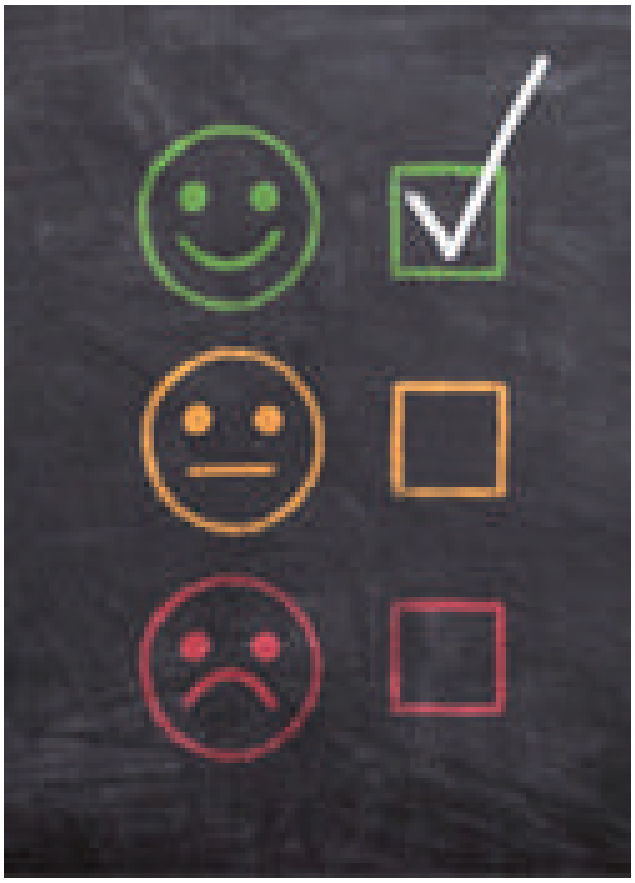
mike@qualetics.com

630.715.4540

www.qualetics.com

Qualetics as a Solution

This analysis helps us understand the customer queries of a company and the average time taken by the customer support service team to address the query. From the analysis, we can tell that the response time should be improved either by giving preference to that group of customers who are using negative words or emojis in their queries, or to that group of customers from which the company is getting maximum profits. Optimization in the social customer support service team can be achieved by giving the right group of queries to the agents who have experience handling similar queries. This will help the companies to provide satisfactory services to the customers i.e improving the brand and business growth of the company.



Often, social media interactions fail to impress customers and don't live up to their expectations. Comments, messages, and any other such interactions create a lot of noise. Qualetics provides solutions that enable you to stay in touch with customers. It delivers real value to businesses by ensuring an ROI with social media. It enables integrating social channels into call centers, aligning the customer service teams for a better information sharing process. Qualetics enables an easy and seamless integration of our solution into your applications or analytics solutions. It also provides the flexibility for immediate adaptations to be applied to meet the dynamically changing consumer needs.

Qualetics Intelligence Platform provides the following data intelligence services-

- Data Extraction and Processing Pipelines
- Data Analysis
- Visualization and Integration Platform.

Data Extraction and Processing Pipelines

Our proprietary architecture built using some of the industry-standard protocols and tools like MQTT, Kafka, NoSQL data stores allows us to set up a data pipeline from the source to the analytical data store with relative ease. This allows us to apply pre-designed learning models to data as it is being generated.

Data Analysis

In addition to pre-designed and pre-trained models, having a continuous data stream that can pick up new data points or variations allows us to refine and tweak the models over time or develop new models altogether. Our expert Data Science team can periodically monitor the efficiency and performance of the models and update them over time

Visualization and Integration Platform

The results of the analysis are delivered through our proprietary Visualization and Integration platform. The platform can be integrated into your Products, Systems, and Processes by Single Sign-On (SSO) or APIs.

Along with features such as continuous monitoring and automated alerts, we can automate an otherwise manual and tedious process such as data analysis and help you build intelligent frameworks from your most valuable resource, your data.