Self-Aware Contextual Behavior Analysis for Service Quality Assurance Over Social Networks

Deepanshi, Jaypee Institute of Information Technology, India

Adwitiya Sinha, Jaypee Institute of Information Technology, India

ABSTRACT

Social media allows people to share their ideology through an efficient channel of communication. The social dialogues carry sentiment in expression regarding a particular social profile, trend, or topic. In this research, the authors have collected real-time user comments and feedback from Twitter portals of two food delivery services. This is followed by the extraction of the most prevalent contexts using natural language analytics. Further, the proposed algorithmic framework is used to generate a signed social network to analyze the product-centric behavioral sentiment. Analysis of sentiment with the fine-grained level about contexts gave a broader view to evaluate and perform contextual predictions. Customer behavior is analyzed, and the outcome is received in terms of positive and negative contexts. The results from the social behavioral model predicted the positive and negative contextual sentiments of customers, which can be further used to help in deciding future strategies and assuring service quality for better customer satisfaction.

KEYWORDS

Artificial Intelligence, Behaviour Prediction, Context-Aware Computing, Natural Language Processing, Social Network Analysis, Twitter-Centric Model

1. INTRODUCTION

The social network is a highly complex and interconnected structure of online communities that offers an interactive platform to exchange ideas and ideologies. The socially interacting networks are regarded as complex not only due to the number of profiles (nodes) and links (edges) but because of the existence of profile-centric features (aspects) and language associated (Ji et. al., 2020; Lu et. al., 2014). Such association requires context-based language analysis to distinguishes the profiles as well as the nature of links connecting those online profiles. For instance, profile name, date, time, posts, etc. are regarded as multiple aspects in the social network. Social network analysis mainly targets the nodes and edges linking in any specific context. It focuses on the type of connection between the nodes and the relationship defining how the nodes are interconnected with each other. We use nodes to identify contexts, however, one may also use contexts to identify nodes. Analyzing the contextual information helps to gain better insights that can be used for measuring, monitoring, and evaluating contents to improve the operability of any organization.

Behavioral analysis, also generally referred to as sentimental analysis, combines the usage of text processing and mining techniques to thoroughly evaluate the subjective mood of users about extracted contexts (Pradhan et. al., 2020; Kumar et. al., 2020). It represents the process of analyzing

people's sentiments or attitudes towards any product, policy, topic, or event. Any shift in the online opinion of people in different contexts directly affects the performance parameters of large-scale social, political, or profit organizations. Behavioral analysis is one of the most buzzed filed in social network analysis. With the increase in social networking sites, e-commerce sites, and online platforms for expressing user opinions like Facebook, Twitter, etc. competition among the products and policy has increased a lot. User satisfaction and user loyalty is the utmost concern of the policy and product makers for assuring the quality of service. Business analysts, customer support personnel, human resource managers, company directors, and other stakeholders make use of sentiment analysis to understand the problems and challenges being presently confronted. User feedback has been taken very seriously for the quality improvement of the product and policy. This could further assist in creating space for designing a wide range of feasible solutions for organizational progress. Behavioral analysis is applied to user feedback, comments, reviews, or any other text message to identify the polarity of the message as positive, negative, or neutral. Such analysis is generically performed on the document level as a whole, thereby resulting in overall polarity which leads to very high sparsity in the result. To improve this problem, our research is focussed on contextual behavioral analysis (Zhao et. al., 2018).

One of the significances of analyzing user sentiment is to perform segregate textual content into smaller parts with polarity. This performs classification of the associated sentiment of text as positive, negative, or neutral. There are several graphical ways of representing the sentimental aspects, of which signed graphs provide richer visualization. A signed network belongs to the network theory that assists in modeling real-world phenomena. A signed edge in the social network graph represents an edge with a positive, negative, or neutral sign. Social networks offer a rich blend of positive and negative connections. This network architecture helps in identifying the pattern of interaction that is only feasible by analyzing the data with finer granularity. In the present day scenario, besides understanding the thought process of people regarding newly launched products, existing policies, etc.; it is equally important to analyze feedbacks for each of the associated features and classify whether the experience is positive, negative, or neutral. Some of the significant terms related to our research are highlighted as follows:

- *Aspect* – For experimentation, aspects are Twitter attributes (features) that are extracted from Twitter using Twitter API analysis is done using these aspects only
- *Context* – These are the words which are most frequently used in the significance of experience, on whose basis the user sentiment from the tweets, could be further evaluated
- *Behavior* – It denotes the sentiment value negative, positive, or neutral value of the tweet chunk respective to the context words
- *Signed Network* – It is an advanced way of representing the signed data which is a mixture of a friendly and antagonistic relationship based on sentiment in the network

Signed Social Network is a type of social network where nodes are the entities and the edges form the social connections. Signed means whether the connection is positive and is represented by green, negative edge is represented by red, and neutral is represented by grey.

The context is the most talked-about noun words that are extracted from the tweeted text and user behavior is analyzed concerning the context word, based on which signed network is constructed. A real-time dataset is extracted from Twitter based on profile-based information (i.e. Twitter user handle). It is easier to analyze tweets because they are short in length due to character limitation on Twitter, hence people are inclined to precisely put their views. The words used in tweets are of high sentiment values, rather than posts from any other medium, e.g. Facebook. Our research goal is to propose a generic model that can be used for any purpose (policy, product, campaign, etc.) or profile (e.g. individual, organization, etc.) to capture the temporal dynamics presented in the real-time data. For instance, Zomato launched ZomatoPremierLeague mostly referred to as ZPL with the Indian

Premier League (IPL) matches. People were posting positive tweets in large numbers, thereby making #ZomatoPremierLeague a popular trend on Twitter, which can be used to strategize plans for launching ZomatoCup with World Cup matches.

2. RELATED WORK

The literature survey encompasses several research studies from which have inspired towards contributing in a better and effective way. All the shortcomings and benefits of all these papers have collectively led to improvements and innovations made in our proposed work. This survey is a quick summary of the Context-based behavioral analysis using a signed network.

A self-regulating and unsupervised approach is designed for sentiment analysis techniques (M. et al, 2018). The mechanism mainly consists of two essential functions contextual analysis and unsupervised ensemble learning in which sentiment lexicon, SentiWordNet, is employed. Another research had proposed lexicon-based methods for opinion mining includes the identification of sentiment lexicon polarities and predict the overall sentiment of a text by aggregation of scores (Muhammad et. al., 2016). Lexicon based methods are used for labeled data which is hard to obtain if we work on real-time data. This technique reduces the semantic difference between the polarities of terms by lexicon methods and new contextual polarity. Sentiment analysis is the most popular area for text mining and opinion mining as the internet is so extensively used nowadays. A rich overview of different aspect extraction technique and approaches are illustrated in (Rana et. al., 2016). In yet another research, aspect level sentiment analysis has been applied to process fine-grain sentiment information which can be useful for applications in different fields, for business planning, product quality enhancement strategy, etc. This study aims to find concept centric sentiment for aspect level semantics (Schouten et. al, 2016).

Sentiment Analysis of online text is the demand of different organizations. product-based sentiment analysis aware users about the positives and negatives of the product. The authors in (Bahrainian et. al., 2013) introduced a target-oriented hybrid method for sentiment summarization on Twitter with their many other additional functionality and features. In yet another research, sentiment analysis used text evaluation techniques for finding the sentiment using some natural processing approaches (Adarsh et. al., 2018). opinion extraction helps in business development and perceiving the user needs. This study calculates the score of words and classifies them as negative, positive, or neutral. Aspect-based sentiment categorizing method for tweet sentiment classification is proposed in (Lek et. al., 2013). The authors have used aspects, sentiment words, and polarity pairs to evaluate the aspects and improved the classification method. According to research conducted in (Liu, 2012), the sentiment extraction can be formally based on the quadruple $\left(s;g;h;t\right)$, where s denotes the sentiment, g denotes the target object about which the sentiment is expressed, h denotes the holder (user expressing the sentiment), and t denotes the time at which the sentiment was expressed. In previous studies, most studies focus only on sentiment and the target object $\left(s,g\right)$. The target can be a product or policy or entity, or it can be its aspect that can be characteristic of that entity or product or policy. Sentiment analysis has been examined mostly at three levels of categorization. The sentiment is categorized on either the document level, the sentence level, or the entity or aspect level. In sentiment analysis opinion can be stated as judgment or belief not founded on certainty or proof.

In (Leskovec et. al., 2010), the authors have innovated a model for analyzing interconnections among users by identifying triads in static social data, extracted Epinions, Slashdot, Wikipedia. The edge between the users of social media indicate the relations among the users, Links in social media is a blend of positive (friendly) and negative (antagonistic) interconnection. In disparity to most of the research in the area of social networks has fixated mostly on positive links among people, in this paper, we get to know how the interconnection between positive and negative interconnection change

the arrangements of social networks. An extension of this research on the signed network analyses social network links to validate social relationships, which could either be positive or negative. By analyzing the links, one can predict the sign of a link with high precision. Social media is a complex connection between negative and positive ties, forming a large and complex structure, which is mostly elusive. Authors in (Teixeira et. al., 2017) have studied the change in the perception of users by considering the behavior of others in the same social network. They have studied how peer influence systems work and to what degree one can balance the social network. Signed Social Network have positive, negative or neutral behavior about any product or policy or individual which evolve with time, user change and grow their social setting which in turn effect sign in the social network (Gross and Blasius, 2008; Weller et. al., 2014).

The research conducted in (Yang et. al., 2007) reveals that the real-world system can be converted into a social signed network with signed links between them that can be either negative or positive. Previous research mainly focussed on positive relations which were not very suitable in the context of the signed network. Hence, the authors introduced an algorithm that studies both the factors, including the sign and density of the connections, which makes it more competitive for both the signed network and the traditional network. The algorithm uses an agent-based heuristic approach that gives optimal solutions and benchmark results in various experiments. Identification of communities in a social signed network using the designed algorithm yielded better speed and clustering techniques. Another research done in the field (Tran et. al., 2018) shows how the trend of online shopping has increased, online reviews play an important role as they help other customers to know about the pros and cons of the product and help the company to know the feedback and improve the product quality or service. It has become a task to summarize the information present in the reviews. In this paper, they introduced a novel system for sentiment summarization. It has two steps, first that use frequency, polarity, and strength of the review aspect and second is to the summarization of the reviews automatically using various artificial intelligence techniques of natural language processing. The system automatically incorporates the new reviews into the system and update it from time to time.

In yet another study (Zhu et. al., 2018) the authors have stated opinion mining through polling in a traditional way for better customer satisfaction along with feedback analysis of products. Studies highlighted the opinion polling without using conventional polling techniques and relied only on customer reviews available in the form of textual feedback. Aspects are extracted from the reviews using the bootstrapping method and then opinion mining is done by breaking the sentence into multiple sentences with a single aspect each and finally using the polling algorithm results are evaluated. The method is applied to unlabelled data and does not need any labeled data for training. The proposed system is not domain-specific or language-specific it can be applied to any product or policy. E-commerce's rapid growth has led to an increase in online purchases day today. While shopping online customer satisfaction is of utmost importance to keep the buyer a regular customer of the vendor. In another recent study (Sangeetha et. al., 2017), authors have used reviews and feedback given by the users about any specific product or policy that is taken as a positive insight to enhance the quality of manufacturing and service. Aspects are extracted and prioritized. The proposed system has four main steps, the first being extracting the aspects, followed by categorizing or grouping the aspects and predicting the score of the sentiment. The final stage performs a ranking of the aspects based on opinion. The importance of the aspect is determined by the aspect ranking score.

3. PROPOSED MODEL & DESIGN FRAMEWORK

In the following segment, the model assumptions are discussed along with the dataset and its features. It also highlights module descriptions, libraries, and a software application used for the implementation.

3.1. Data Source

Streaming data is extracted using Twitter API for the Twitter handle of Zomato India i.e. @ZomatoIN and @swiggy_in having 10,000 tweets. Twitter provides various features, also called aspects, which are used for analysis that are extracted from Twitter using Twitter API. For instance - tweet text, date, time, users mentioned, location, followers, screen name, verified status, etc. Only English language tweets are considered for further processing in our model. In the following segment, our proposed strategy and algorithm for the context-based signed network is explained.

3.2. Proposed Methodology

Our proposed methodology is designed to build a self-aware computational model for predicting the contextual behaviors of the customers using natural language processing. In the following section, the proposed methodology used for data extraction and data pre-processing is highlighted.

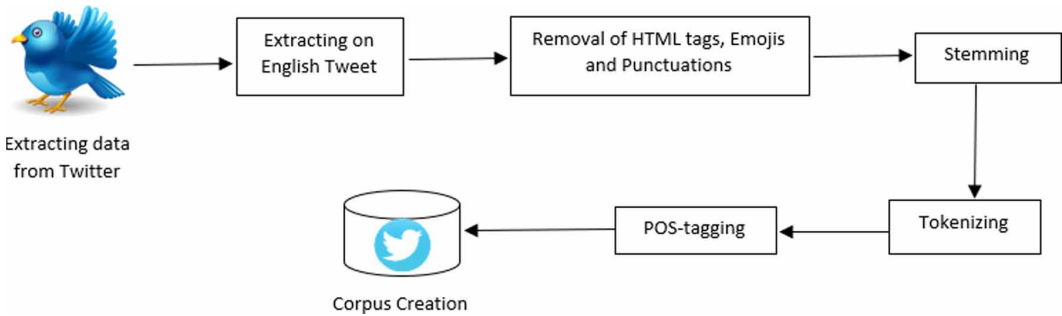Figure 1. Data Extraction and Pre-Processing

Fig. 1 represents data extraction and pre-processing. Twitter is a social networking microblogging website that users use to post and communicate through messages called tweets. The amount of information Twitter has is extremely large and unstructured written in common dialect. Data is extracted from Twitter using the Twitter streaming API. Real-time tweets are streamed from Twitter. For accessing Twitter API, OAuth is used, OAuth provides authorization access to the Twitter API in a simple and standard technique from the web, mobile, and desktop application. Real-time tweets are steamed with the aid of the Twitter streaming API. To fetch data from Twitter, the user should have a Twitter account. To authentication the account, a developer profile is created through which the consumer key and consumer secret and the access key and access token are generated. By using these credentials Twitter API can be accessed to acquire the data. In our model, data can be extracted by using both screen names or hashtags according to the requirement. Tweets retrieved from Twitter contain information in a very unstructured way which is very hard to process as the language used in tweets is mostly very lament. Tweets contain many emojis, special symbols, HTTP links, and many punctuation marks. So, all the emojis, all types of special symbols, all the links, and punctuation marks are removed to clean the tweets so that the further processing of tweets can be done easily. Tweets are then tokenized, which splits the stream of text into meaningful words or elements called tokens. Tokenized tweets' results are further processed stemming that is considered an effective way of reducing all the words with the same origin and in the case of the prefix. The POS tagger states some specific tags for all the words in a sentence (table 1). The processing of data is further performed by using a corpus created by POS tagging.

Table 1. List of POS Tags and Description

S.No.	Symbol	Description
1.	JJ	Adjective
2.	JJR	Adjective, comparative
3.	JJS	Adjective, superlative
4.	NN	Noun, singular or mass
5.	NNS	Noun, plural
6.	NNP	A proper noun, singular
7.	NNPS	A proper noun, plural
8.	PRP	Personal pronoun
9.	PRP$	Possessive pronoun
10.	RB	Adverb
11.	RBR	Adverb, comparative
12.	RBS	Adverb, superlative
13.	UH	Interjection
14.	VB	Verb, the base form
15.	VBD	Verb, past tense
16.	VBG	Verb, gerund or present participle
17.	VBN	Verb, past participle
18.	VBP	Verb, non-3rd person singular present
19.	VBZ	Verb
20.	PP	Possessive Pronoun

3.3. Algorithm for Context-Based Signed Network Creation

In this section, the proposed algorithm is highlighted, which explains context extraction, chunking of tweets, deriving sentiments, and graphical representation (Algorithm 1). Table 2 illustrates the list of symbols used in the algorithm.

Corpus of data created by the POS tagging is used for three main processes in our model, which are detailed as follows:

● **Context Extraction**

In our model, sentiments are derived and evaluated according to the context terms. Most of the context terms in a tweet are nouns, we needed to extract the noun words from tweets, which are further divided into four categories. Table 3 lists all the POS-Tags that are further used for context extraction.

All the context words of the noun category from all the tweets are extracted using Term Frequency–Inverse Document Frequency (TF-IDF), which is a statistical way of calculating the frequency of a word in any document to express its importance. The number of times a word is present in a document, the more is its frequency. It is one of the simplest, efficient, and extensively used term weighting scheme, and is expressed in equation (1).

Table 2. List of Symbols Used

S.No.	Symbol	Description
1.	$\Omega_{handle}\left(t\right)$	Pre-processed tweet dataset extracted for user $handle$ at time t
2.	tag	Tags assigned by pos-tag function
3.	δ	Set of tags that will be used for context extraction
4.	X	Extracted context words
5.	$\bar{X}$	Highest occurring top 50 context words
6.	θ	Chunks of tweet
7.	ψ	A sentiment value of a chunk
8.	C_{pos}	Positive tweet classification
9.	C_{neg}	Negative tweet classification
10.	C_{net}	Neutral tweet classification
11.	α_{color}	Color set according to tweet category classification
12.	$G_{signed}\left(S,T,\alpha_{color}\right)$	Signed graph

$$TFIDF\left(t,d,D\right) = TF\left(t,d\right) \times IDF\left(t,D\right) \qquad (1)$$

where,

$$tf\left(t,d\right) = \log\left(1 + freq\left(t,d\right)\right) \qquad (2)$$

$$idf\left(t,D\right) = \log\left(\frac{N}{count\left(d \in D : f \in d\right)}\right) \qquad (3)$$

Here, t refers to the number of terms, d denotes each document in the collection of documents, represented by D. So, after applying TF-IDF we get all the context words in the order of their frequency.

Weight Frequency Inverse Document Frequency (WFIDF): It is a proposed solution to the drawback of using term frequency which is the assumption that the count of the number of appearances

Algorithm 1. Context-Based Signed Network Creation

Input: $\Omega_{handle}\left(t\right)$
Procedure:
1: for each tweet in $\Omega_{handle}\left(t\right)$:
2:     $tag \leftarrow$ pos tagging ($\Omega_{handle}\left(t\right)$)
3: $\delta = \{NN, NNP, NNPS\}$
4: if tag is in δ:
5: $X \leftarrow$ extract context words ($\Omega_{handle}\left(t\right)$, δ)
6: $\bar{X} \leftarrow$ cleaned top 50 context words $\left(X\right)$
7: end-for
8: for each tweet in $\Omega_{handle}\left(t\right)$
9: $\theta \leftarrow$ chunk $\left(\Omega_{handle}\left(t\right)\right)$
10: for x in $\bar{X}$
11: if x is in θ :
12: $\psi \leftarrow$ extract (sentiment value)
13: end-for
14: end-for
15: C_{pos} , C_{neg} , C_{net} $:= \phi$
16: $\alpha_{color} = \left\{\alpha_{green}, \alpha_{red}, \alpha_{grey}\right\}$
17: for each ψ in θ :
18: if $\psi > 0$:
19: $C_{pos} := C_{pos} \cup \left(S, T, \alpha_{green}\right)$
20: if $\psi < 0$:
21: $C_{neg} := C_{neg} \cup \left(S, T, \alpha_{red}\right)$
22: if $\psi = 0$:
23: $C_{net} := C_{net} \cup \left(S, T, \alpha_{grey}\right)$
24: end-for
25: generate graph $G_{signed}\left(S, T, \alpha_{color}\right)$

Output: $G_{signed}\left(S, T, \alpha_{color}\right)$

of a term in a document is equal to the count of the significance of a single occurrence. Equation (4) is the mathematical expression of WFIDF.

Table 3. List of POS-tag for context extraction

S.No.	Symbol	Description
a.	NN	Singular Noun
b.	NNS	Plural Noun
c.	NNP	Singular Proper Noun
d.	NNPS	Plural Proper Noun

$$wfidf = \begin{cases} 1 + logtf_{i.j}.idf_{i,} & if\ tf_{i.j} > 0 \\ 0, & otherwise \end{cases} \quad (4)$$

where tf_{ij} is the term normalization of term *i*, and idf_i is the inverse document frequency of term i. All the words of one or two characters are cleaned from the top words and for further processing, we choose the top 50 context words in whose reference we will do the sentiment analysis. Top context word does not include the search screen name or hashtag that is used for fetching the data as that word or phrase will be included in every tweet as all the tweets are extracted in that particular reference only.

- **Chunking of Tweets**

Chunking is a process of breaking singular bits of data and then grouping them in a particular pattern. By grouping the data, we can improve the measure of data we can recollect. In chunking, we select a subset of the token into a hopefully meaningful chunk. In our model, the main focus is on doing noun phrase chunking. Chunking is applied to the pre-processed data corpus. Each chunk of the data must contain one or more noun words because context words extracted are noun words. We aim to group words in a way all the words related to that context or noun words come in a single chunk so that we can find the sentiment about that context word. In our model, we applied grammar for chunking tweets into meaningful sub-parts, called chunks. For our experimentation, a chunk is expressed as follows:

$$chunk = a\,?\,b\,?\,c\,?\,d\,?\,e\,?\,f\,?\,g\,?\,h\,?\,i\,? \quad (5)$$

where the symbols $a \ldots i$ are highlighted with relevant description and expression in the following table 4.

The grammar rule state that the chunk may contain optional determiner or possessive pronoun or optional any form of adverb or optional any form of the verb or optional any number of the adjective or optional one or more proper noun or optional one or more singular noun or optional one or more singular proper noun or optional one or more plural proper noun. The above grammar chunks the tweets. After chunking, all the chunks of tweets are stored with the respective tweet id. Sentiment analysis of all the chunks is performed using the Textblob library. Sentiment analysis through Textblob returns a sentiment in the form of (polarity, subjectivity) the polarity lies between the range -1 to 1. The sentiment value of all the chunks is stored with the respective chunks for further processing. Now, the main task is matching context words to the noun words in the tweet chunk. If the context word lies in the tweet chunk then we take the sentiment value of the tweets. Every word of the context words file that contain the top 50 pre-processed context words is matched with all the possible chunks

Table 4. List of Symbols used for Chunking of Tweets

S. No.	Variable	Expression	Description
1.	a	<DT|PP\$>	optional determiner or possessive pronoun
2.	b	<RB.?>	any form of the adverb
3.	c	<VB.?>	any form of the verb
4.	d	<JJ. *>	any number of adjectives
5.	e	<NNP>	one or more proper noun
6.	f	<NN. *>	one or more singular noun
7.	g	<NNP>	one or more singular proper noun
8.	h	<NNPS>	one or more plural proper noun
9.	i	< NNS>	one or more plural proper noun

made from all the tweets. If the tweet chunks that contain that specific context word are collected in a file and their respective sentiment value is extracted that was stored concerning the chunk sentiment. The average of all the sentiment scores of all the tweet chunks with the particular context word is computed and we get the overall sentiment score about that context. The same process is repeated with all the context words being extracted.

- **Graphical Representation**

Representation of data done through graphical techniques to determine the optimal output is called Graphical representation. It is one of the most important and powerful tools used for data interpretation. Graphical representation helps in summarizing characteristics of data adequately and knowledgeably giving us a clear picture at the first glance only. The simplicity of graph theory enables any complex structure to be represented easily with visual interpretation. The graph provides more information about the shape of the distribution, the relation between the different variables and trends. Graphical techniques make it very easy to understand the overall data in every interactive way with due details. After computing the overall sentiment value of the context and the individual value of the chunks, the main task is to represent the outcomes. Data is represented in a graphical form, which produces quality figures such as histogram, bar graph, pie chart scatterplot, etc. depending upon the requirement of the user. Overall sentiment analysis of the data is done by using a pie chart and the context word extracted is shown by using a bar graph. For network visualization, we have used Gephi, an open-source graph, and network analysis software. It can deal with a large amount of data for large networks and provide a variety of functions for representing the graph network in a highly interactive way to deliver the best possible outlook of the data with minimum effort. As we are doing sentiment analysis based on context, therefore we have represented the data in the form of a signed network.

Fig. 2 represents the Hierarchical Signed Social Network, Level 0 which is represented by orange nodes is the Twitter handle that is used for extracting data. Data can be fetched by using any handle, screen name, hashtags, or keyword. Level 1 is represented by the blue nodes in the context words extracted from the data. It represents the context that is most talked about. Level 2, represented by the green node, is the tweet ids in which respective context words are found. The edge between the

context word and tweet Ids and screen name represents the link between tweet ids and context words and screen names. The color of the edge depends upon the sign on tweet sentiment value. In signed network edges contain the sign according to the sentiment value. The sign edges are not shown by putting signs on the edges but by the color of the edges. If the tweet sentiment value about respective context words is negative then the color of the edge is red from level 2 to level 1. If the tweet sentiment value about respective context words is positive then the color of the edge is green from level 2 to level 1. If the tweet sentiment value about respective context words is neutral then the color of the edge remains grey from level 2 to level 1. If the overall sentiment value of the context word is negative then the color of the edge is red from level 1 to level 0. If the overall sentiment value of the context word is positive then the color of the edge is green from level 1 to level 0. If the overall sentiment value of the context word is neutral then the color of the edge remains grey from level 1 to level 0.

Figure 2. Hierarchical Signed Social Network
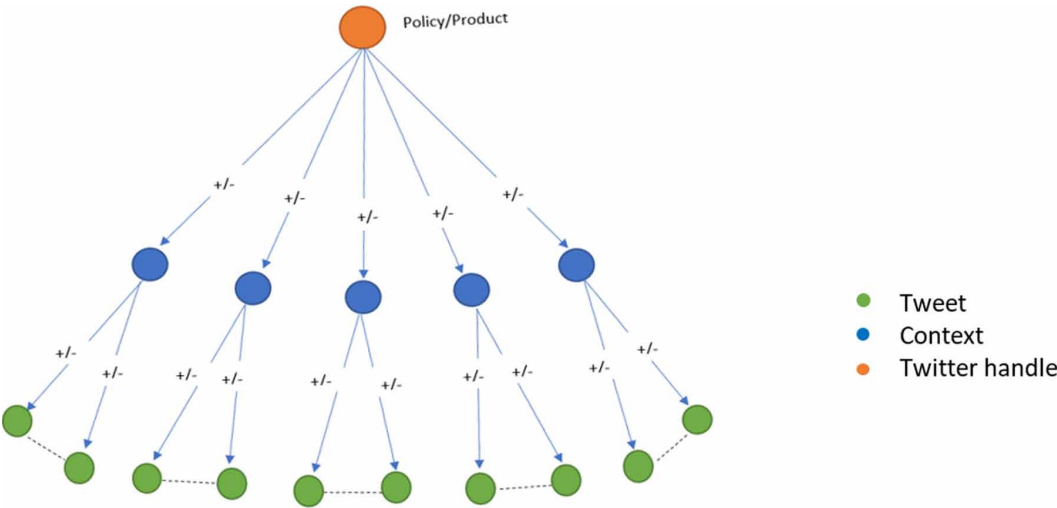
4. RESULTS ANALYSIS

The Twitter handles used for fetching the live data include @ZomatoIN and @swiggy_in, respectively.

4.1. General Sentiment Analysis

The following results show outcomes on applying sentiment analysis on the dataset and then classifying tweets into the positive, negative, or neutral tweet.

Fig. 3- 4 shows the sentiment analysis for Zomato and Swiggy respectively. It is the process of identifying the user's polarity of positive, negative, or neutral behavior on any given topic. We have used the screen name @ZomatoIN and collected 8,000 latest tweets and @swiggy_in and collected 6000 latest tweets respectively. Sentiment analysis of tweet text is conducted after pre-processing by using the python library Textblob. It shows the sentiment analysis of the tweets based on their chunk sentiment value. The tweet is divided into chunks, whose sentiment analysis is computed individually and the average of the sentiment values of the chunk is evaluated as a new overall sentiment of the tweet.

This provided 26.2% of positive, 49.0% of neutral, and 24.8% of negative tweets for Zomato and 30.3% of positive, 41.8% of neutral, and 27.9% of negative tweets for Swiggy. Fig. 5- 6 shows the classification of tweets as per generating sources using which the tweets are posted. These sources
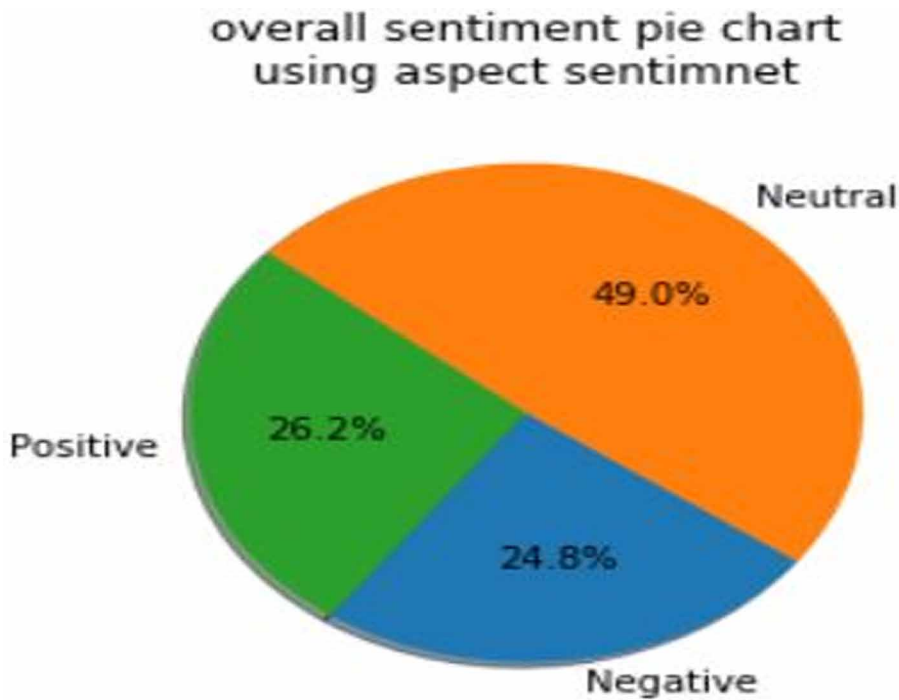
Figure 3. Sentiment Analysis for Zomato
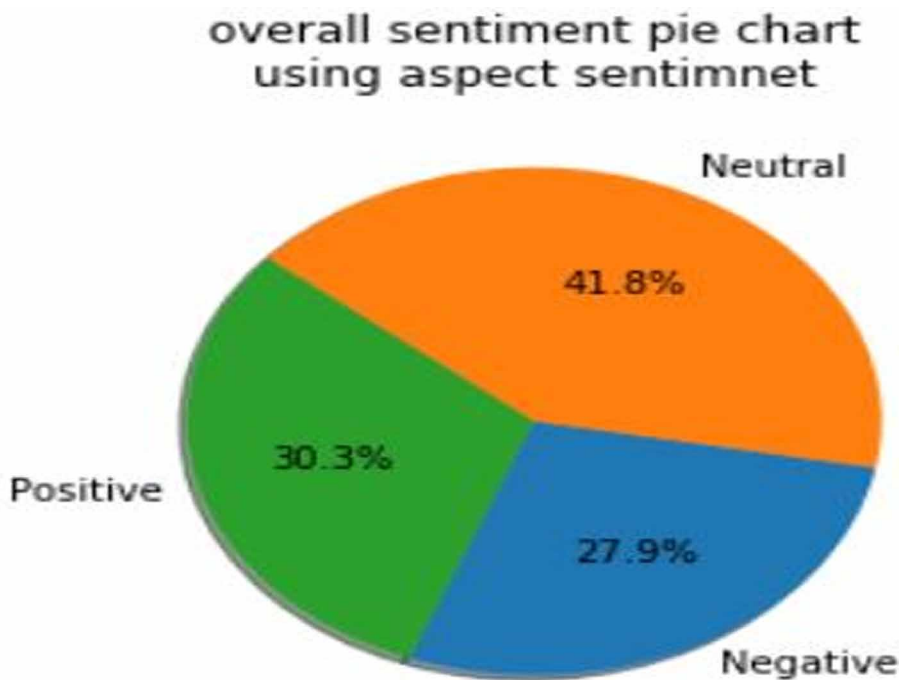

Figure 4. Sentiment Analysis for Swiggy

Figure 5. Classification of Zomato Tweets as per Generating Sources
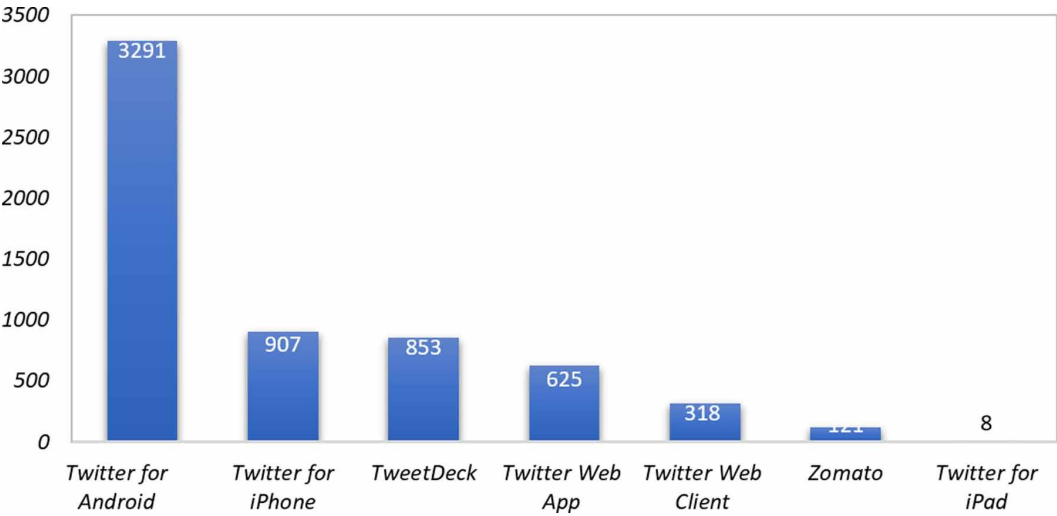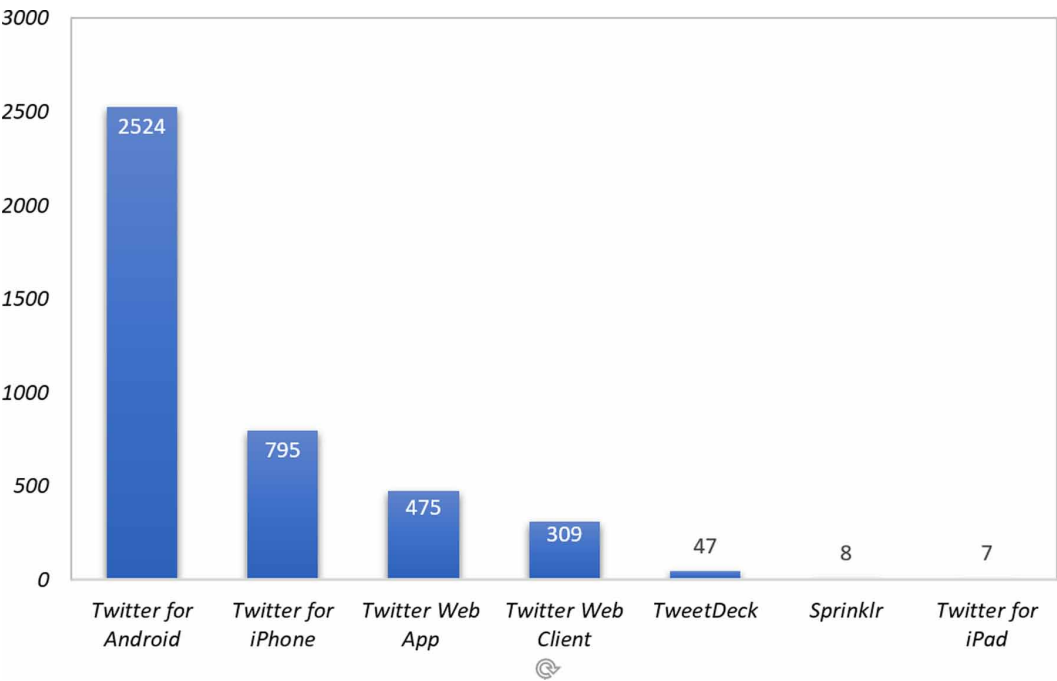

Figure 6. Classification of Swiggy Tweets as per Generating Sources

include, Twitter for Android, Twitter for iPhone, TweetDeck, Twitter Web App, Twitter Web Client, Twitter for iPad of which maximum tweets are posted by using Twitter Android and lowest tweets are posted by using iPad.

4.2. Contextual Analysis

In the following section, the results of the context words are highlighted that are extracted from the dataset for contextual analysis.

Fig. 7-8 shows the top context words extracted from the tweets from Zomato and Swiggy respectively. These context words are the most frequent noun words extracted from the dataset. The X-axis shows the words that act as context words and Y-axis denotes the frequency of word occurrence. But these context words contain the screenname or hashtags used for extracting data. It may also contain one or two-letter words that could be noun words but do not necessarily contribute to any context. Hence, such words need to be cleaned. So, such words are cleaned out of the top 50 context words extracted from the dataset. 10 words are dropped in the context pre-processing giving us the top 40 Context words that can be used for further processing.

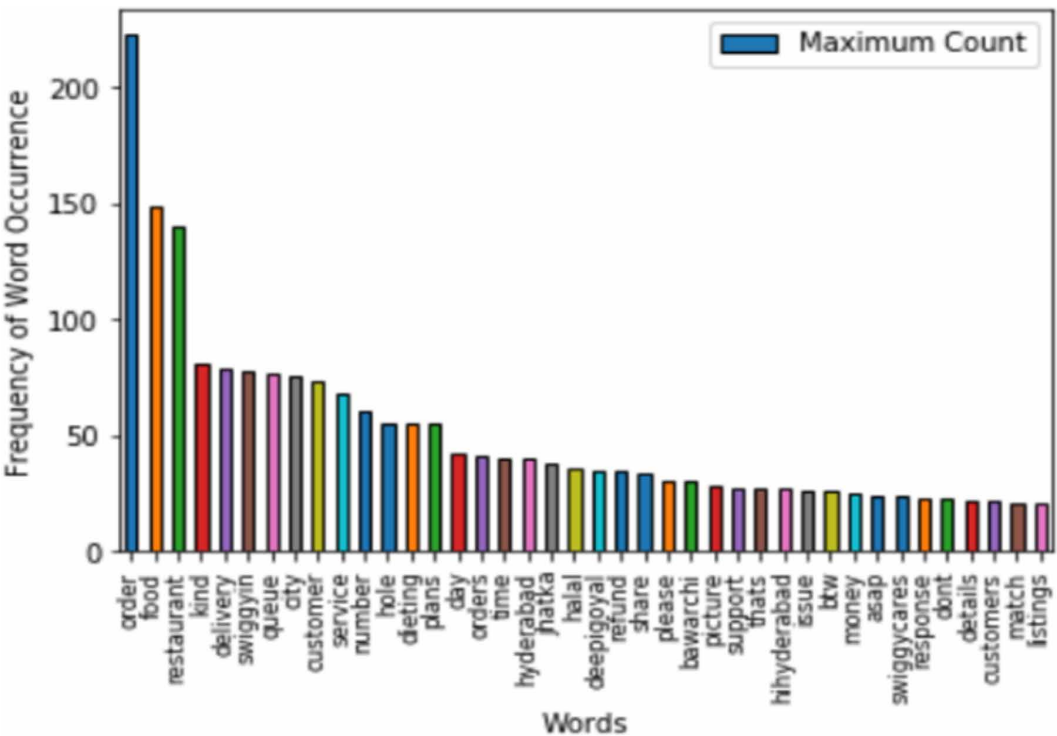Figure 7. Top Context Words in Zomato

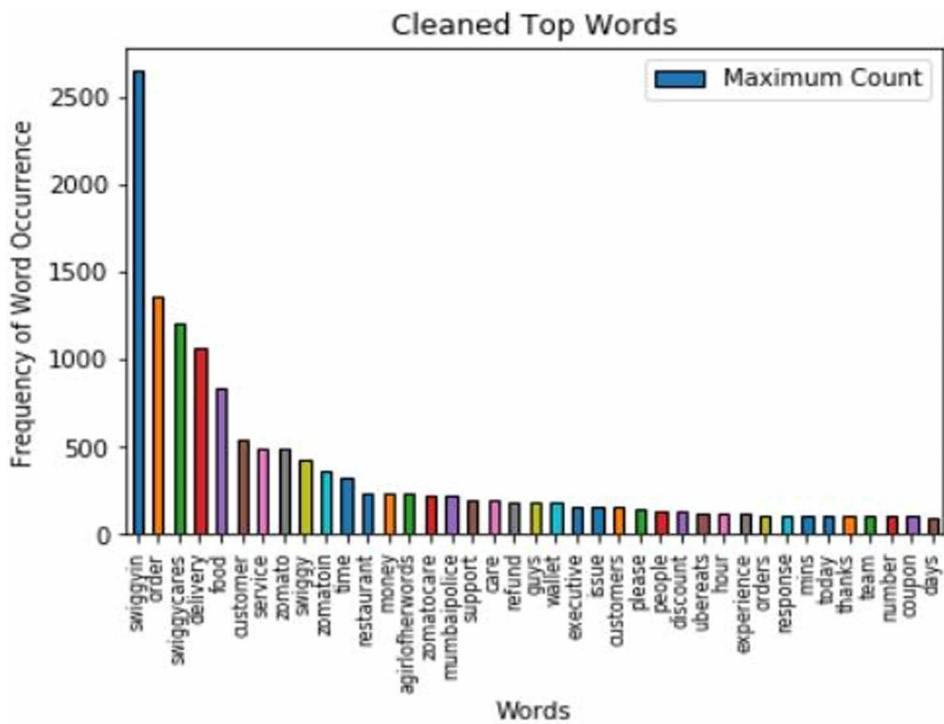**Figure 8. Cleaned Top Context Words in Swiggy**

Fig. 9 -10 represents the network of context words for Zomato and Swiggy. The yellow nodes are the context words and the red node represents the master Twitter handle for which the data is extracted. The edge between the context word and Twitter handle represents the link between Twitter handle and context words.

Figure 9. Network graph of Top 50 Context Words in Zomato

Figure 10. Network graph of Top 50 Context Words in Swiggy
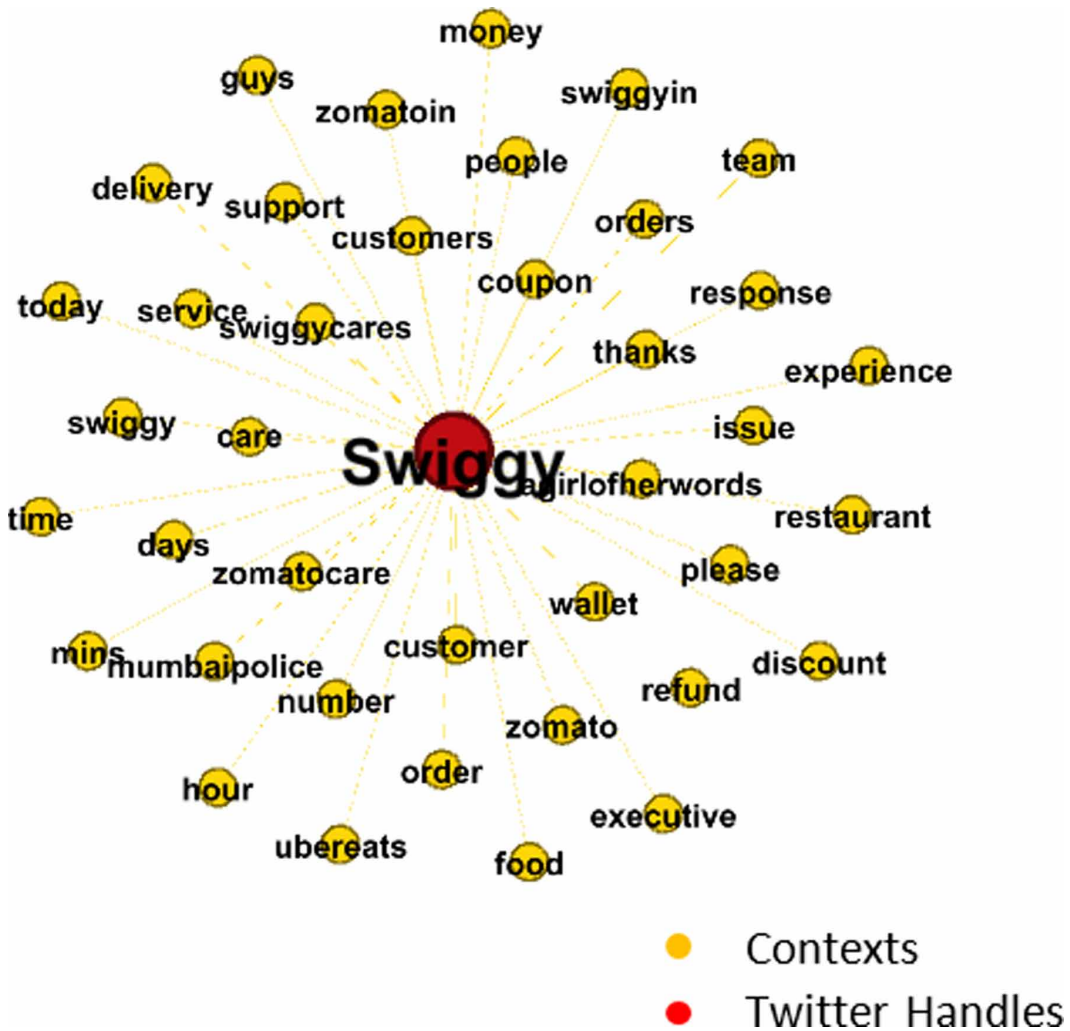
4.3. Context Networks

In the following section, we are representing the results of the social network which contain the context words and tweet ids. Fig. 11-12 represents the network of cleaned context words and the tweet Ids for Swiggy and Zomato respectively.

The yellow nodes are the context words and the blue nodes are the tweet ids. Grey edges between context word and tweet ids represent links. Tweet ids are connected to one or more than one context words depending upon the number of context words present in the tweet. The size of the context word which is represented by the yellow nodes is depending upon the degree of the node. The greater the degree of a node, the greater the size of the node. Labels on context word nodes are shown. It is a simple directed graph in which context nodes are the target and tweet ids are the source.

4.4. Signed Social Network

In the following section, the results are highlighted in the form of a signed social network, between the context words and the tweet ids with signed edges.

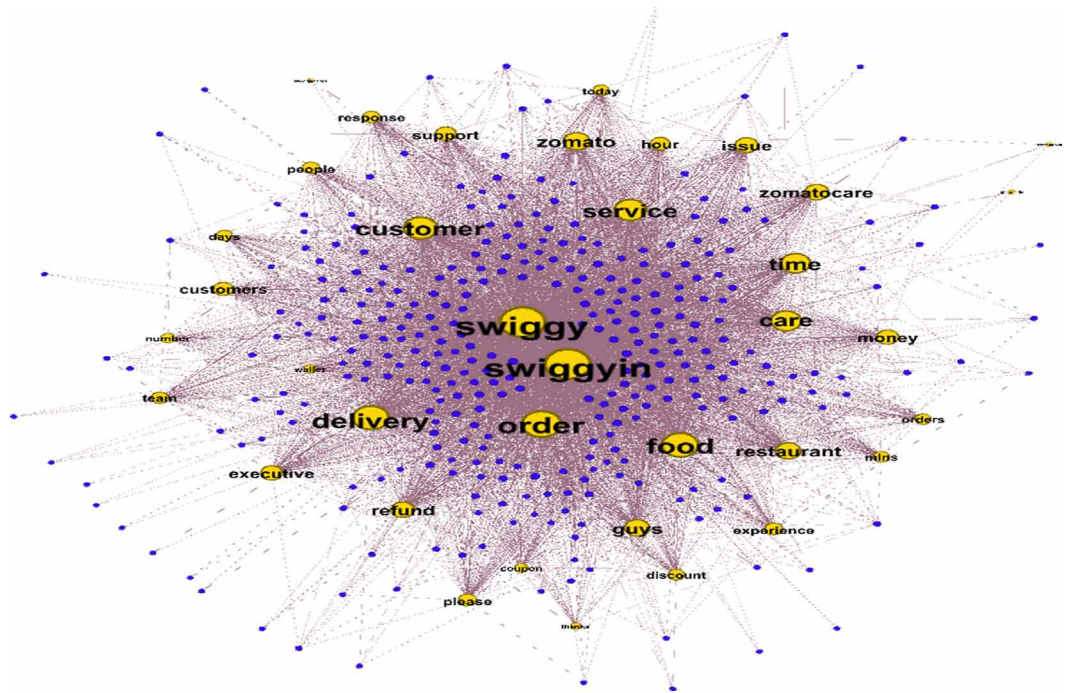**Figure 11. Network graph of Top Context Words in Swiggy**

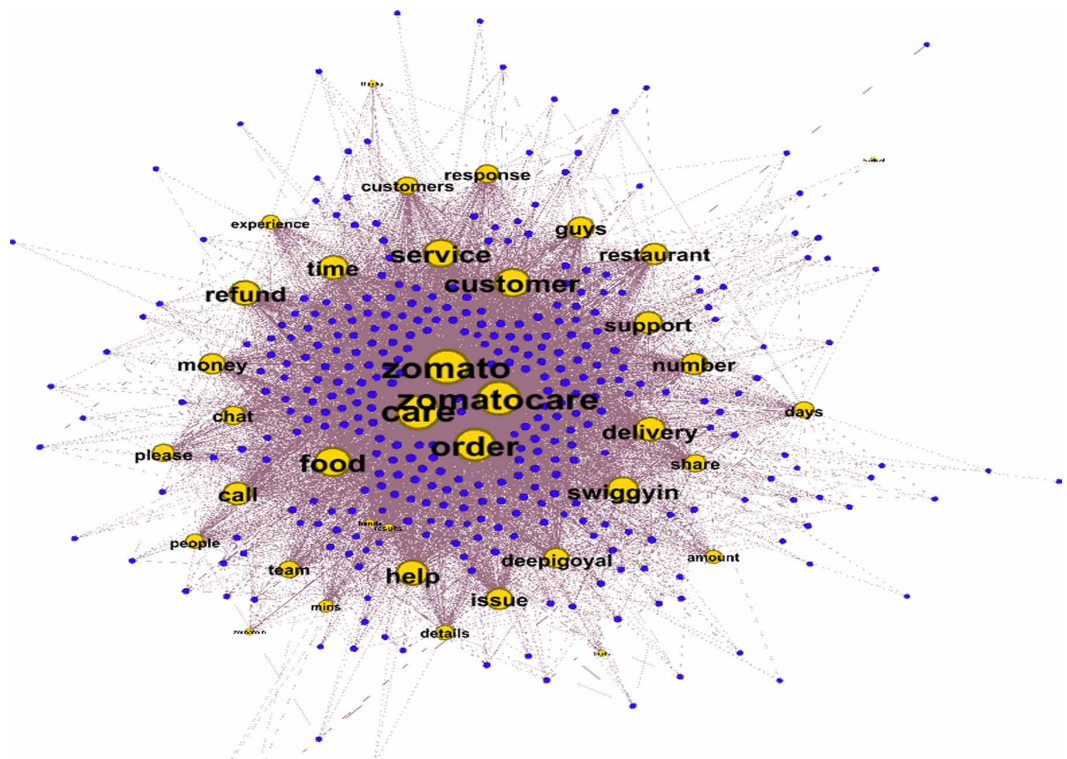**Figure 12. Network graph of Top Context Words in Zomato**

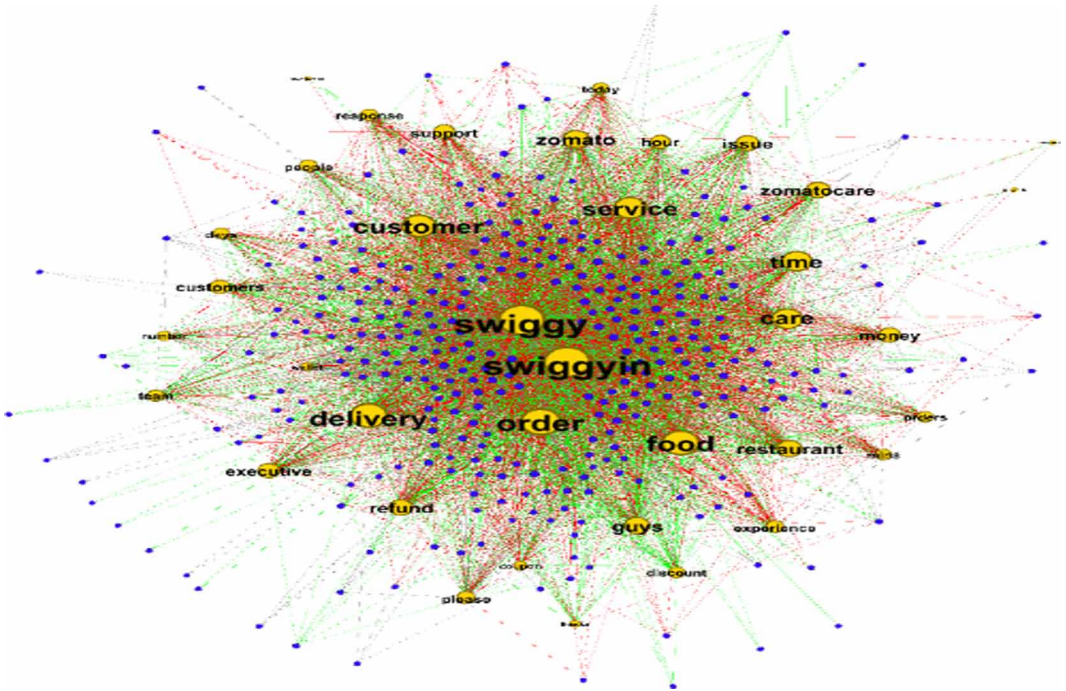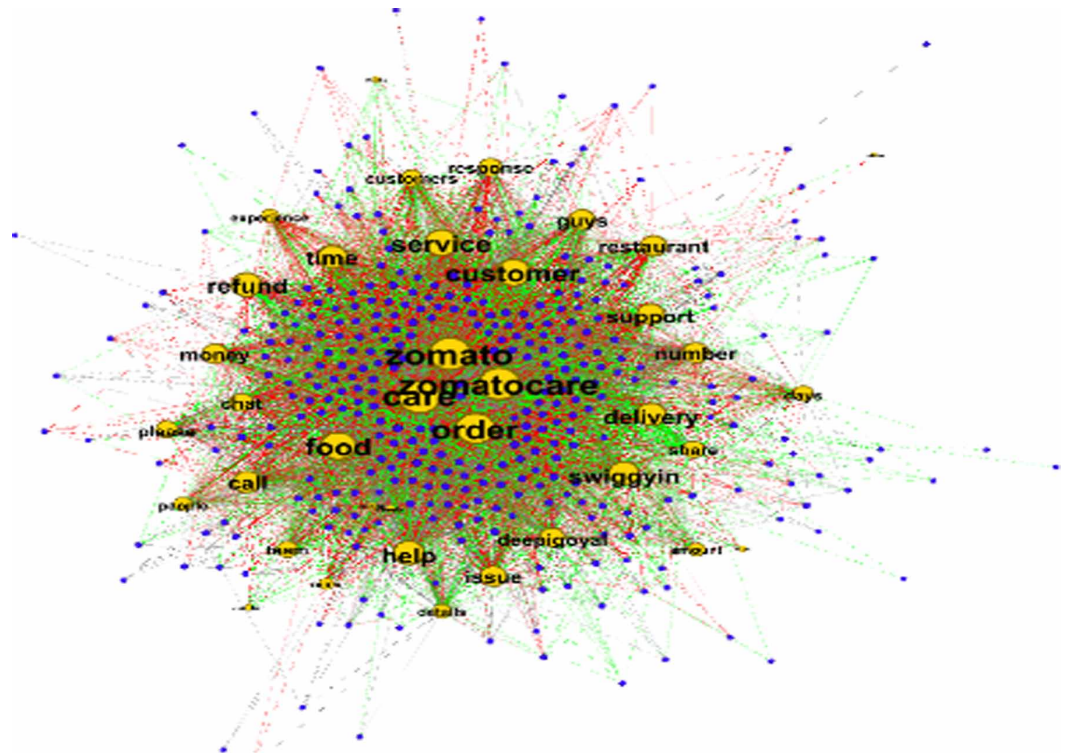**Figure 13. Simple Signed Network Graph for Swiggy**

Figure 14. Simple Signed Network Graph for Zomato

The color of an edge depends upon the sign on tweet sentiment. In the signed network, edges contain the sign according to the sentiment value of the chunk. The positive, negative, and neutral sentiments are highlighted with green, red, and grey edge colors. Fig. 15-16 represents a network graph of top context words using hashtag or screen name. Overall sentiment value for all the context words is computed individually. The yellow nodes are the context words and the red nodes denote the search screenname/hashtag. The positive, negative, and neutral sentiments of the individual context are highlighted with green, red, and grey edge colors respectively.

Tables 5 and 6 provide a prediction of the positive and negative context words which directly indicate the current mood of the customers about the products and services which can aid in decision making and improving strategy using our proposed context-based language analytics.

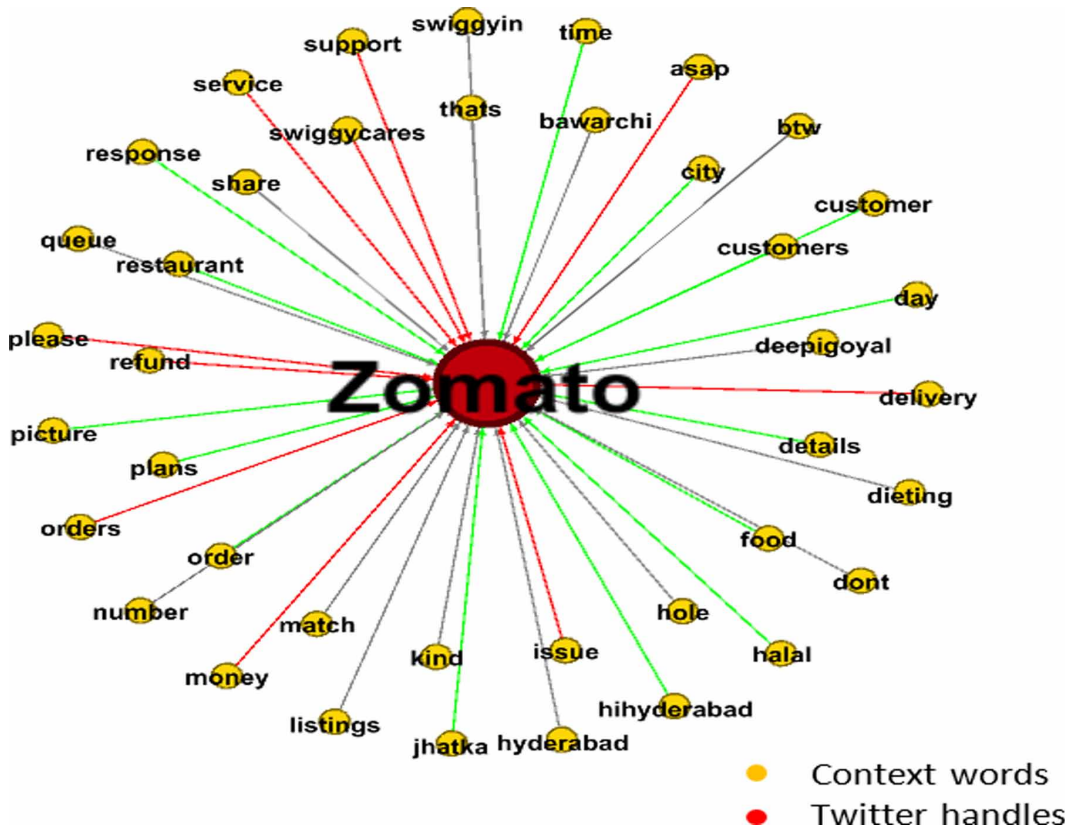**Figure 15. Context graph for Zomato**

Figure 16. Context graph for Swiggy

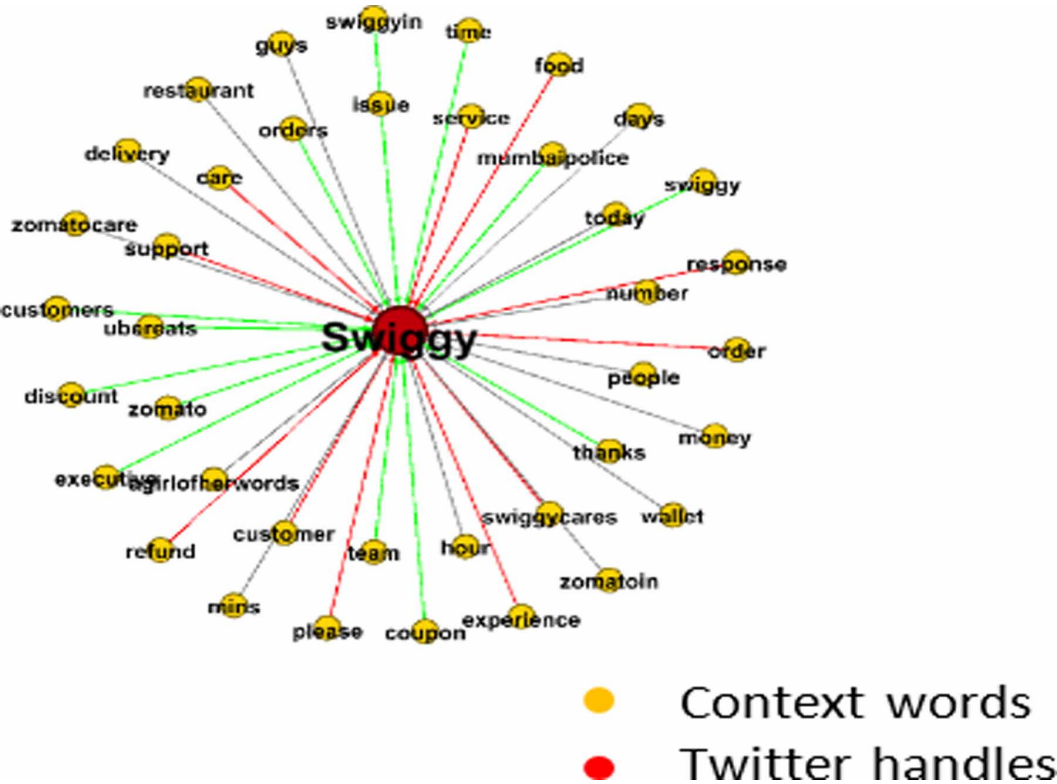**Table 5. Prediction of Positive & Negative Contexts for Zomato**

S.No.	Positive	Negative
1.	Food	Issue
2.	Order	Delivery
3.	Restaurants	Money
4.	Time	Refund
5.	Hyderabad	Service

Table 6. Prediction of Positive & Negative Contexts for Swiggy

S.No.	Positive	Negative
1.	Time	Food
2.	Discount	Response
3.	Coupon	Experience
4.	Thanks	Refund
5.	Executive	Support

5. CONCLUSION

Our research is motivated towards assisting large-scale organizations to perform analysis of user behavior in the form of artificial intelligence-based sentiment mining using natural language processing that could aid in decision making and service quality improvement. For this purpose, we have extracted real-time user comments from the Twitter social network for Zomato and Swiggy food delivery portals. The user experience posted on social media is considered highly genuine, and hence the feedbacks form a sound basis for research. In this paper, we have built a self-aware computational model for context-based sentimental language analysis, visualized through a signed social network. The overall sentiment analysis provides knowledge about an abstract sentiment, while context-based analysis helps us to find the major contexts from a list of topics and the mood associated with each such context. This is achieved using natural language processing, which involved chunking and lemmatization to compute the hidden contexts in user comments over social media. Further, our proposed algorithm is executed to generate the signed social network based on the real-time user experience, and the contents computed. The network data was processed and we found 26.2% of positive, 49.0% of neutral, and 24.8% of negative tweets for Zomato, and 30.3% of positive, 41.8% of neutral, and 27.9% of negative tweets for Swiggy. Moreover, our study revealed 50 contexts from both the food delivery portals, of which 10 contexts were found most relevant. Our proposed work aims at benefitting organizations in understanding their flaws in winning maximum customer satisfaction. Future work in the same field can be extended by adding emoticons to the analysis. We can also check the sentiment of the paragraph by evaluating the emoticons in it. Emoticons used will add up to the sentiment value. Furthermore, the main limitation is the changing jargon used in the day-to-day world of social media where words like lol, FOMO, OOTD, etc. come up every day and we have limitations in our system to study them in terms of their literal meaning.

REFERENCES

Adarsh, M. J., & Ravikumar, P. (2018, February). An Effective Method of Predicting the Polarity of Airline Tweets using sentimental Analysis. In *2018 4th International Conference on Electrical Energy Systems (ICEES)* (pp. 676-679). IEEE.

Al-Sharuee. (2018). Sentiment analysis: An automatic contextual analysis and ensemble clustering approach and comparison. *Data & Knowledge Engineering*. Advance online publication. doi:10.1016/j.datak.2018.04.001

Bahrainian, S. A., & Dengel, A. (2013, December). Sentiment analysis and summarization of Twitter data. In *2013 IEEE 16th International Conference on Computational Science and Engineering* (pp. 227-234). IEEE.

Gangrade, S., Shrivastava, N., & Gangrade, J. (2019). *Instagram Sentiment Analysis: Opinion Mining. In Proceedings of Recent Advances in Interdisciplinary Trends in Engineering & Applications*. RAITEA.

Gross, T., & Blasius, B. (2008). Adaptive coevolutionary networks: A review. *Journal of the Royal Society, Interface*, *5*(20), 259–271.

Hemmatian, F., & Sohrabi, M. K. (2019). A survey on classification techniques for opinion mining and sentiment analysis. *Artificial Intelligence Review*, 1–51.

Ji, Luo, Guo, Wang, Yu, & Li. (2020). Community detection in online social networks: a differentially private and parsimonious approach. *IEEE Transactions on Computational Social Systems*, *7*(1), 151-163.

Kumar, Sathis, Nabeem, Manoj, & Jeyachandran. (2020). Sentimental Analysis (Opinion Mining) in Social Network by Using Svm Algorithm. In *Fourth International Conference on Computing Methodologies and Communication (ICCMC),* (pp. 859-865). IEEE.

Kumar, R., Pannu, H. S., & Malhi, A. K. (2020). Aspect-based sentiment analysis using deep networks and stochastic optimization. *Neural Computing & Applications*, *32*(8), 3221–3235. doi:10.1007/s00521-019-04105-z

Lek, H. H., & Poo, D. C. (2013, November). Aspect-based Twitter sentiment classification. In *2013 IEEE 25th International Conference on Tools with Artificial Intelligence*(pp. 366-373). IEEE.

Leskovec, J., Huttenlocher, D., & Kleinberg, J. (2010, April). Predicting positive and negative links in online social networks. In *Proceedings of the 19th international conference on World wide web* (pp. 641-650). ACM.

Liu, B. (2012). *Sentiment Analysis and Opinion Mining* (Vol. 16). Synthesis Lectures on Human Language Technologies. Morgan & Claypool.

Lu, Z., Sun, X., Wen, Y., Cao, G., & La Porta, T. (2014). Algorithms and applications for community detection in weighted networks. *IEEE Transactions on Parallel and Distributed Systems*, *26*(11), 2916–2926.

Muhammad, A., Wiratunga, N., & Lothian, R. (2016). Contextual Sentiment Analysis for Social Media Genres. *Knowledge-Based Systems*. Advance online publication. doi:10.1016/j.knosys.2016.05.032

Pradhan, R., Khandelwal, V., Chaturvedi, A., & Sharma, D. K. (2020). Recommendation System using Lexicon Based Sentimental Analysis with collaborative filtering. In *International Conference on Power Electronics & IoT Applications in Renewable Energy and its Control (PARC),* (pp. 129-132). IEEE.

Rana, T. A., & Cheah, Y. N. (2016). Aspect extraction in sentiment analysis: Comparative analysis and survey. *Artificial Intelligence Review*, *46*(4), 459–483.

Sánchez-Núñez, P., Cobo, M. J., De Las Heras-Pedrosa, C., Peláez, J. I., & Herrera-Viedma, E. (2020). Opinion Mining, Sentiment Analysis and Emotion Understanding in Advertising: A Bibliometric Analysis. *IEEE Access : Practical Innovations, Open Solutions*, *8*, 134563–134576. doi:10.1109/ACCESS.2020.3009482

Sangeetha, T., Balaganesh, N., & Muneeswaran, K. (2017, June). Aspects based opinion mining from online reviews for product recommendation. In *2017 International Conference on Computational Intelligence in Data Science (ICCIDS)* (pp. 1-6). IEEE.

Schouten, K., & Frasincar, F. (2016). Survey on aspect-level sentiment analysis. *IEEE Transactions on Knowledge and Data Engineering*, *28*(3), 813–830.

Teixeira, A. S., Santos, F. C., & Francisco, A. P. (2017, March). Emergence of social balance in signed networks. In *International Workshop on Complex Networks* (pp. 185-192). Springer.

Tran, T. A., Duangsuwan, J., & Wettayaprasit, W. (2018, July). A Novel Automatic Sentiment Summarization from Aspect-based Customer Reviews. In *2018 15th International Joint Conference on Computer Science and Software Engineering (JCSSE)* (pp. 1-6). IEEE.

Weller, K., Bruns, A., Burgess, J., Mahrt, M., & Puschmann, C. (2014). *Twitter and society* (Vol. 89). Peter Lang.

Yang, B., Cheung, W., & Liu, J. (2007). Community mining from signed social networks. *IEEE Transactions on Knowledge and Data Engineering*, *19*(10), 1333–1348.

Zhao, G., Lei, X., Qian, X., & Mei, T. (2018). Exploring Users' Internal Influence from Reviews for Social Recommendation. *IEEE Transactions on Multimedia*, *21*(3), 771–781.

Zhu, J., Wang, H., Zhu, M., Tsou, B. K., & Ma, M. (2011). Aspect-based opinion polling from customer reviews. *IEEE Transactions on Affective Computing*, *2*(1), 37–49.

Adwitiya Sinha has received Ph.D. from Jawaharlal Nehru University (JNU), New Delhi, India in 2015. She has obtained Masters of Technology in Computer Science and Technology in 2010 from JNU. She was awarded with First Rank certificate in M.Tech. batch 2008. She received Senior Research Fellowship (SRF) from the Council of Scientific & Industrial Research (CSIR), New Delhi and also awarded research scholarship from the University Grants Commission (UGC) for her research in wireless sensor networks. She has delivered several invited talks in lecture series on Networks & Graphs, Wireless Sensor Networks, and Performance Analysis of Computing Systems, organized by the Consortium for Educational Communication (CEC), UGC, New Delhi in the form of EDUSAT live lectures. Presently, she is working as an Assistant Professor in Jaypee Institute of Information Technology, Uttar Pradesh, India. She is a senior member of IEEE. Her research areas mainly include performance analysis of wireless sensor networks, complex networks, data science, and large-scale graphs analytics.