

Analysing Opinions of People for different Products

¹Heena Girdher, ²Anshul Garg

^{1,2}Department of Computer Applications, Chandigarh Group of colleges, Landran, Mohali, India

¹ heena.1120@cgc.edu.in

Abstract: Ongoing years have seen the hazardous advancement of online internet based life. A huge number of messages are posted everyday on Social media like Twitter, Facebook, Linked In and Tumbler etc. Such large number of messages posted are the rich source of information for analysis purpose. Sentiment analysis is a technique of finding the judgments of people towards a particular topic that is positive, negative or neutral. With this technology, sales revenue, customer service and product quality etc. can be improved. This paper focuses on finding the sentiments of people for different categories of Products like Mobile Phones, Cars and Laptops. The analysis is done on Twitter data which is fetched online through Twitter API. Python Programming language is used for the task. In this research, thousand tweets are fetched online for different categories of products. The outcome shows the percentage of positive, negative and neutral tweets for all the products.

I. INTRODUCTION

In recent years, there have been the hazardous advancement of online web based life. A large number of texts are placed day by day on Social networking sites, for instance, Linked In, Google Plus, Facebook and Twitter. Weibo, a Twitter-like administration propelled in 2009, has pulled in more than 500million clients in under 5 years, with more than 100 million tweets being placed each day. An ever increasing number of users of social media tend to share their reviews and responses for a particular event online. Such information can be used to adjust marketing strategy, measure ROI of your marketing campaign and to improve customer service etc.

Sentiment Analysis (Opinion Mining) is the way toward identifying the relevant polarity of twitter posts, e-commerce websites and some other blogging sites. It decides if a writer's outlook towards a post is positive, negative or neutral. An elective term is opinion mining, as it determines the assessment, or emotional state of posts. This technique helps to predict how individuals' feel about a specific point. Sentiment Analysis is one field of NLP which is drawing in extraordinary consideration from specialists. News and web journals are typically great wellsprings of information for sentiment analysis, wherein individuals can express their views and sentiments on such discussions.

Sentiment analysis generally studies the polarity of posts such as positive, Negative and Neutral. For example, if someone writes "I love the summer in New York, but I hate the winter." The individual scores would show "love the summer" as positive and "hate the winter" as negative. However, the sentiment for the entire comment would be neutral, because the positive sentiment for the word love would cancel out the negative sentiment for the word hate [1].

II. APPLICATIONS OF SENTIMENT ANALYSIS

The applications of Sentiment Analysis are wide and great. Some of the applications are mentioned below:-

1. Numerous organizations utilize it to track their items, services or reputation by and large. For instance, on the off chance that somebody is assaulting your mark via web-based networking media, sentiment analysis will score the post as to a great degree negative, and you can make alarms with the post with hyper-negative assessments score.
2. Organizations can get comment on how glad or disappointed the client is, and utilize this understanding to pick up an aggressive edge.
3. Political parties might want to be informed about whether public is supporting their plans or not. For example- Prime Minister Narendra Modi has launched an android based application named "Narendra Modi Mobile app" where citizens can directly interact with the PM to promote good governance.
4. Sentiment analysis that is becoming hugely necessary because of opinion spam detection. While e-mail and Web spam are quite familiar, opinion spam is still new to the overall population. Because of the enormous growth of user-generated content on the Web, it is now a common practice for people to find and read other's opinions. Opinion spam refers to human activities that try to deliberately mislead readers or automated opinion mining systems by giving undeserving positive opinions to some target objects and/or by giving unjust, malicious or false negative opinions to other object.
5. Sentiment analysis can also be helpful for recommendation systems, since those systems should not recommend something that receives negative feedback, and for the development of new kinds of search engines.

III. REPORTED WORK

Yang et al. [1] presented their work on opinion mining and sentiment analysis. Traditional opinion mining methods are based on sentiment dictionary but they can't deal with buzzwords or domain sentiment words. In this paper, a domain sentiment dictionary is created using external textual data which can deal with different types of sentiments. Besides, an extremely powerful hybrid model is proposed which combines different single models to reach higher accuracy. The outcome depicts that hybrid model performs better than the single models.

Deshwal and Sharma [2] presented a new concept for twitter sentiment analysis using various classification algorithms. In this paper, many pre-processing and feature extraction techniques are combined to design improved sentiment classification system because the past studies of sentiment classification were not very conclusive. Besides, the result of six supervised classification algorithms is also compared.

Gupta et al. [3] proposed an approach for sentiment analysis using machine learning approach. In this paper, different machine learning approaches like K-Nearest Neighbours (KNN) and Support Vector Machines (SVM) are combined in a hybrid manner to improve accuracy of sentiment classification. The analysis of hybrid model is done in terms of both accuracy and f-measure. The results demonstrate that the hybrid approach performs better than the individual classifiers.

Tang et al. [4] developed a deep learning system (Coooolll) for twitter sentiment classification. Firstly, a neural network is built to learn SSWE (Sentiment-specific word embedding) features. Then SSWE features are combined with a state-of-the-art hand-crafted features and a sentiment classifier is built. The evaluation of the resulting deep learning system is done on twitter2014 test data and the system ranked 2nd on this data set.

Neri et al. [5] presented their work on sentiment analysis of social media. In this paper, sentiment analysis study is performed on Facebook posts. Around 1000 posts about newscasts were retrieved from Facebook and sentiments of people were analysed for both Rai and La7 Company. It also considers the data given by Auditel Company about newscast.

Bollen et al. [6] performed a sentiment analysis of all tweets published on twitter. The analysis may be performed on special events like response of the public to the death of famous American Singer Michael Jackson or response of the public to the win of particular political party. By performing analysis on such kind of events, they have extracted the mood of writers towards a particular event. From the results, it was observed that by performing online analysis on twitter data; the time, efforts and money required for performing surveys and questionnaires can be saved.

Wang et al. [7] presented a study that quantified self-data from internet-enabled smart scales with general social media data on Twitter. They used this combination of data sources to predict a user's weight using only their social media activity. They showed that it is possible to predict an individual's weight using their online social behaviors, such as their self-description and tweets. They capture weekly patterns, such as a peak of weigh-in activity on Saturday, and monthly patterns, such as a weight increase over Christmas. These results can be used to build models to monitor public health and to provide more customized personal training interventions.

Zhang et al. [8] proposed a novel entity-level sentiment method for twitter. This method first applies a lexicon based approach to perform sentiment analysis. To improve the recall, additional opinionated tweets are identified through chi square test. A classifier is then trained to assign polarities to the newly identified tweets whose training data is provided by the lexicon based method. The outcome shows that the proposed method is highly effective and promising.

Saif et al. [9] proposed a novel approach for extracting patterns of words of similar contextual semantics and sentiment on Twitter. This approach does not rely on syntactical templates and doesn't require deep analysis of the syntactic structure of tweets. The experiment is carried out on 9 twitter datasets and the performance of this approach is compared with the 6 state-of-art-baselines. The outcome depicts that proposed patterns consistently outperform all other baselines on all datasets by 2.19% at the tweet-level and 7.5% at the entity-level in average F-measure.

Mesnil et al. [10] proposed a new ensemble system for sentiment analysis. They combine three different baseline models. Each such model contributes to the success of the overall system. The experiment is carried out on the dataset of movie reviews. The outcome shows that the results are highly effective and promising. The link for the code is given so that the results can be reproduced.

IV. PROPOSED APPROACH

In this paper, we proposed a framework to perform sentiment analysis. We have utilized various products' posts as the area for the examination and acquired these tweets from Twitter. The posts are fetched online through twitter API. The tweets so gathered contain reviews with positive, negative and neutral emotions. In the first step, tweets are fetched online and pre-processing is undertaken on fetched tweets. Pre-processing step involves removing the hash tags, emoji's, special characters and links etc. from the tweets. Further, the sentiments of tweets are categorized as positive, negative or neutral. To perform sentiment analysis on tweets, we have developed a program in Python which is an interpreted high-level programming language for data

science. The steps used in the proposed system are following:

1. Register an app through twitter account.
2. Get keys and tokens from the Twitter Developer Console.
3. Authenticate twitter API client.
4. Fetch tweets through twitter API for a particular query.
5. Clean tweet text by removing links and special characters.
6. Parse the tweets and classify sentiments of reviews as positive, negative or neutral.
7. Find the percentage of positive, negative and neutral tweets for a particular query and display the result.

Input: - query {what kind of data you want to fetch online from twitter}; count {gives the number of tweets you want to fetch online}

Output: -Get data from twitter and calculate the percentage of sentiments i.e. positive, negative or neutral and display the result.

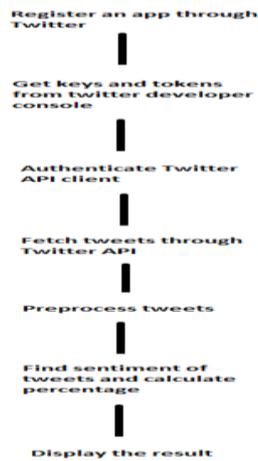


Fig1: Steps to perform Sentiment Analysis

V. ALGORITHM PROPOSED

Algorithm:

- Step 1:-** Authentication of twitter API client
- Step 2:-** Get tweets from the twitter API by passing the arguments query and count
- Step 3:-** for each tweet in fetched_tweets do
- Step 4:-** clean tweet text by removing links, special characters
- Step 5:-** If sentiment.polarity>0
 return positive
 else if sentiment.polarity==0
 return neutral
 else
 return negative
- Step 6:-** for each tweet in tweets do
 If tweet ['sentiment'] ==positive
 Add tweet to the list of positive tweets

- Find percentage of positive tweets
- If tweet ['sentiment'] ==negative
 Add tweet to the list of negative tweets
- Find percentage of negative tweets
- If tweet ['sentiment'] ==neutral
 Add tweet to the list of neutral tweets
- Find percentage of neutral tweets

Step 7:- Plot the bar chart containing percentage of positive, negative and neutral tweets.

VI. RESULT AND EVALUATION

In this research, we have conducted sentiment analysis on reviews of people for different categories of Products. We have used Python language to perform this task. The tweets for different products are fetched online through twitter API and about 1000 tweets are accumulated for each category of product. They were initially pre-processed and then analysed for sentiments. The percentage of tweets for each category of product is mentioned below in table 1.1.

Table 1: Percentage of tweets for different Products

	Categori es	Positiv e Comments	Negat ive Comments	Neutral Comments
Mobil e Phones	iPhone XR	26.3%	5.3%	68.4%
	Samsung Galaxy M20	18.3%	0%	81.7%
	Nokia 8.1	28.9%	4.3%	66.7%
	OnePlus 6T	23.8%	4.8%	71.4%
	Redmi 6 Pro Blue	33.3%	4.2%	62.5%
Cars	Mahindr a XUV300	33.3%	5.9%	60.7%
	Maruti Suzuki Baleno	58.3%	0%	41.7%
	Mercede s-Benz V-Class	56.9%	0%	43.1%
Lapto ps	Apple	35.2%	21.1 %	43.7%
	HP	25.6%	10%	64.4%
	DELL	45.5%	10.2 %	44.3%
	Lenovo	23.7%	12.4 %	63.9%

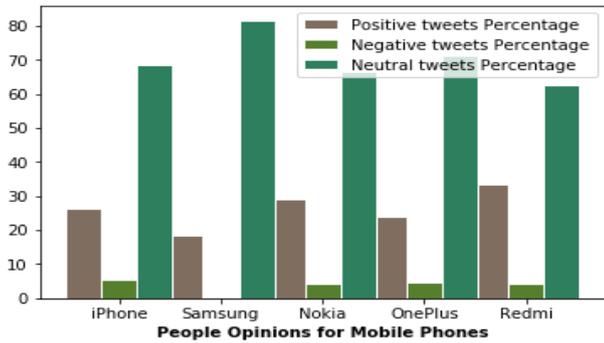


Fig 2: People opinions for Mobile Phones

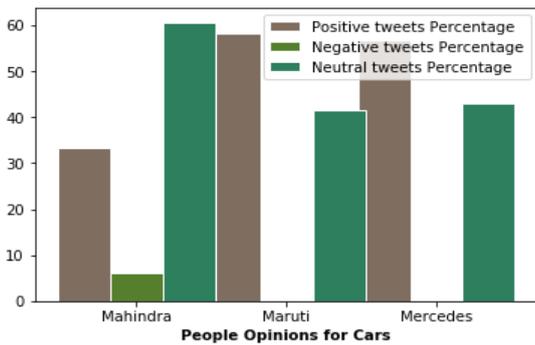


Fig 3: People opinions for Cars

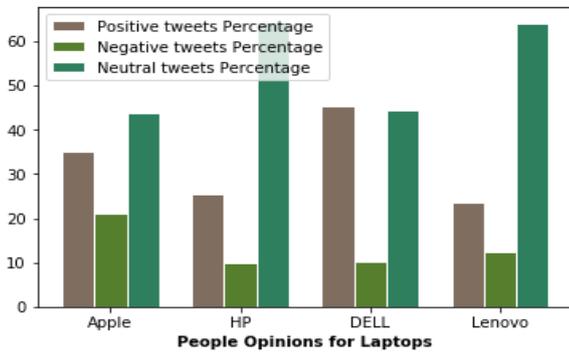


Fig 4: People Opinions for Laptops

VII.CONCLUSION

In this research, we have found the sentiments of people for various categories of Products. The analysis is done on Twitter data which is fetched online through Twitter API. Python programming language is used to carry out the research. Thousand tweets are fetched online for each Product category. The outcome shows the percentage of positive, negative and neutral tweets for all the products.

REFERENCES

- [1] K. Yang, Y. Cai, D. Huang, J. Li, Z. Zhou, X. Lei "An Effective Hybrid Model for Opinion Mining and Sentiment Analysis" 2017 IEEE International Conference on Big Data and Smart Computing (BigComp)
- [2] A. Deshwal, S. K. Sharma "Twitter Sentiment Analysis using Various Classification Algorithms" 2016 5th International Conference on Reliability, Infocom Technologies and Optimization (Trends and Future Directions) (ICRITO)
- [3] A. Gupta, J. Pruthi, N. Sahu "Sentiment Analysis of Tweets using Machine Learning Approach" IJCSMC, Vol. 6, Issue 4, 2017, pg.444 – 458
- [4] D. Tang, F. Wei, B. Qin, T. Liu, M. Zhou "Coooooll: A Deep Learning System for Twitter Sentiment Classification" Research Center for Social Computing and Information Retrieval Harbin Institute of Technology, China, Microsoft Research, Beijing, China
- [5] F. Neri, F. Capeci, M. Cuadros "Sentiment Analysis on Social Media" 2012 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining
- [6] J. Bollen, H. Mao, A. Pepe. "Modeling public mood and emotion: Twitter sentiment and socio-economic phenomena" ICWSM vol. 11, 2011, pp. 450-453
- [7] Y., Wang, I. Weber, P. Mitra. "Quantified Self Meets Social Media: Sharing of Weight Updates on Twitter." In Proceedings of the 6th International Conference on Digital Health Conference, pp. 93-97. ACM, 2016.
- [8] A. Khan, M. Atique, V.M. Thakare "Combining lexiconbased and learning-based methods for Twitter sentiment analysis." International Journal of Electronics, Communication and Soft Computing Science & Engineering (IJECSCE) 2015, 89.
- [9] H. Saif, Y. He, M. Fernandez, H.Alani. "Semantic patterns for sentiment analysis of Twitter." In The Semantic Web–ISWC 2014, pp. 324- 340.
- [10] G. Mesnil, T. Mikolov, M.Ranzato, Y. Bengio. "Ensemble of generative and discriminative techniques for sentiment analysis of movie reviews." arXiv preprint arXiv: 1412.5335 2014.