

Twitter Sentiment Analysis Using Machine Learning for Product Evaluation

Ariza Mirza

Submitted: 05-07-2022

Revised: 15-07-2022

Accepted: 18-07-2022

ABSTRACT

People from all over the world now use social media sites to share information. Twitter, for example, is a platform where users can send, read, and interact with different communities. Users post about their daily lives and their thoughts on various topics such as brands and locations. Companies can benefit from this massive platform by collecting data about their customers' opinions. The goal of this paper is to present a model that can perform sentiment analysis on real-world Twitter data. Twitter data is highly unstructured, making it difficult to evaluate. However, our proposed model differs from previous work in this field in that it employs both supervised and unsupervised machine learning algorithms. The following is the procedure for performing sentiment analysis: Tweet extracted directly from Twitter API, followed by data cleaning and discovery. Following that, the data was fed into several models for training. Each extracted tweet is classified based on its sentiment, whether positive, negative, or neutral. Data was gathered on two subjects, McDonalds and KFC, to determine which restaurant is more popular. Various machine learning algorithms were employed. These models' output was tested using various testing metrics such as cross validation and f-score. Furthermore, our model outperforms other models when mining text extracted directly from Twitter.

Keywords— Sentiment analysis, social media, Twitter, tweets.

I. INTRODUCTION

Users of online social media platforms such as Twitter, Facebook, and Instagram can communicate with people all over the world. Write their own opinions about products or share their experiences, and they can even influence politics and businesses. Twitter, for example, almost every large company has a Twitter account to learn about their customers' feedback on their services or products. Sentiment analysis, also known as opinion mining, is used to categorise specific words as positive or negative. [1-4]

The prediction of emotions in a word or a sentence is known as sentiment analysis. It's designed to be a tool for deciphering the attitudes, beliefs, and feelings conveyed in an online discussion. It's a way of dividing interactions into positive and negative categories. Unstructured information makes up a large portion of the information available in interpersonal organisations. Unstructured data makes up a portion of the information available everywhere. This makes it difficult to study information and make significant decisions based on it. The key system that aids in recognising conclusions is concept examination or evaluation mining.

Consumers, on the other hand, have complete control over what they want to see and how they react. As a result, the company's success and failure are made public and spread by word of mouth. However, social media can impact consumer behaviour and decision-making. For example, claims that 87 percent of internet users are influenced by customer reviews in their purchase and choice. As a result, if a company can catch up faster on what its customers are thinking, it will be more advantageous to organise to react quickly and come up with a smart strategy to compete against their competitors.

We are using sentiment analysis in this paper to classify specific English tweets about two restaurants, KFC, and McDonald's. Our research was aimed at determining which was superior to the other, and we specifically looking at whether specific tweets were positive, negative, or neutral.

Objectives

The study's first goal is to investigate sentiment analysis on Twitter data in order to analyse feedback from customers of an organization's product or service; and the study's second goal is to develop a programme for customer reviews on a product that allows an organisation or individual to sentiment and analyse a large number of tweets into a usable format.

II. LITERATURE REVIEW

In [5] focuses on analysing tweets in written English language belonging to various KSA telecommunication companies for opinion mining, and they used supervised machine learning algorithms for classification. Furthermore, they used TF-IDF (Term frequency – inverse document frequency) to determine the importance of a specific word in a tweet. [6] creates a sentiment analysis approach based on public Arabic tweets and Facebook posts comments.

They used supervised machine learning algorithms such as Support Vector Machine (SVM)

and Nave Bayes, as well as binary model (BM) and TF-IDF, to investigate the effect of various term weighting functions on sentiment analysis accuracy. In [7], they used natural language analysis for Arabic language text and applied sentiment analysis to a Twitter dataset of 4700 for Saudi dialect sentiment analysis with (k=0.807). Researchers in [8] present a sentiment analysis for Egyptian dialect based on a corpus of tweets and product reviews. They use natural language processing to comprehend Egyptian dialect. They also classified the data using a lexicon-based classification.

Proposed Methodology

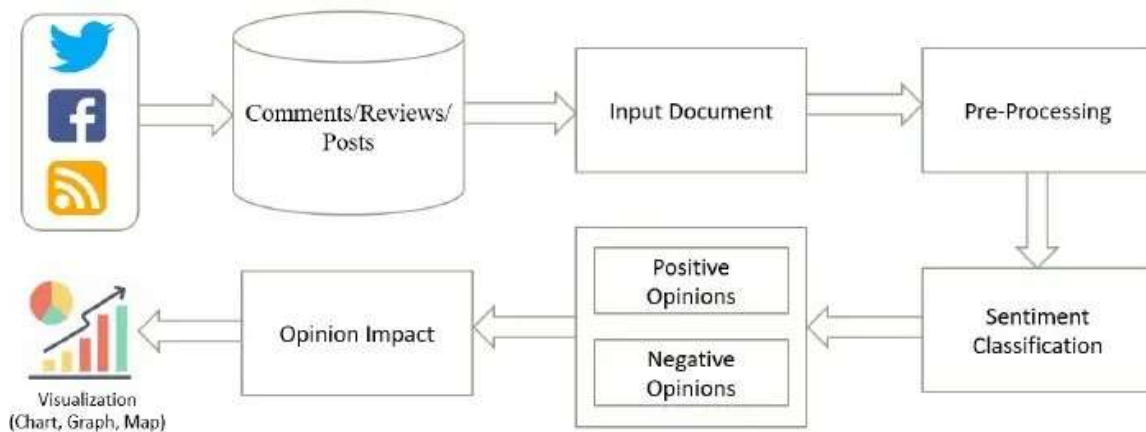


Figure 1 System Architecture

This paper focuses on mining English-language tweets. We're curious whether people think McDonald's or KFC is better in terms of good/bad reviews. Analysing people's opinions and what they think about a product based on their social media tweets could be beneficial to any business. In our project, we used Python programming to extract tweets from Twitter. Python is a popular programming language for statistical computing and machine learning algorithms. Twitter API was used to create a Twitter application and obtain authorization in order to extract tweets from Twitter. To complete this task, we use NumPy, Pandas, Matplotlib, Text blob, and request libraries.

sentiment analysis, is the process of determining a user's viewpoint toward a topic or a product. Opinion mining determines whether a user's opinion of a product, topic, or event is positive, negative, or neutral. The process of opinion mining and summarization consists of three major steps: opinion retrieval, opinion

classification, and opinion summarization. The review text is obtained from review websites. Opinion text in blog posts, reviews, and comments, for example, contains subjective information about a topic. Reviews are classified as either positive or negative. Opinion summaries are generated based on feature opinion sentences by considering frequently occurring features about a topic.

Business Benefits of Sentiment Analysis

Social media is now an essential part of almost everyone's life. They are gradually evolving into a platform for all forms of media communication. Businesses can capitalise on this by carefully listening to and monitoring their customers. It is critical to use sentiment analysis to accurately understand what the client wants. Sentiment analysis can be used to address a wide range of business issues, some of which are listed below.

- Sentiment analysis can provide useful information to businesses, allowing them to

develop effective business strategies. It can help businesses track the performance of their brands and products, resolve customer complaints, and collect detailed data for strategic research. Sentiment analysis can help track and develop successful marketing campaigns. The positive and negative thoughts and comments of customers can be used to calculate the true ROI of marketing activities. Today, sentiment analysis has become the backbone of most businesses' digital initiatives.

- With more modern tools and big data, sentiment analysis can capture, quantify, retrieve, and analyse customers more efficiently. Sentiment analysis can determine whether a phrase is positive, negative, or neutral, and thus reveal a customer's opinion of a brand or service. This aids in determining a company's strengths and weaknesses.
- As a result of sentiment analysis, product quality will improve. Market research teams would be better able to predict consumer preferences and tastes. Customers who have been specifically targeted can also make suggestions for product enhancements. It can help identify up-selling opportunities, reduce customer churn, increase client acquisition, improve customer retention, and resolve customer complaints.
- It can also aid in the identification of new business opportunities. Specific phrases and texts of target audiences can be tracked to efficiently generate new leads. Monitoring competitor brand mentions can also be used to evaluate competitor performance.

III. CONCLUSION

Sentiment analysis can be used in a wide range of business applications, from brand monitoring and product analytics to customer service and market research. Leading brands (not to mention entire cities) can work faster and more accurately toward more useful ends by incorporating it into their existing systems and analytics. Sentiment analysis has progressed beyond being merely an intriguing, high-tech whim, and will soon become an indispensable tool for all modern-day businesses.

Twitter is a massive platform and a rich supply of poorly structured and sentiment information that may be examined to generate trending emotions and many other things. We inspect or mine every piece of the tweet in Twitter sentiment analysis. This paper describes the many

procedures required in analysing Twitter sentiments, as well as the numerous tools that are used to do so. The tweet is used for sentiment analysis, and it goes through several stages. It is critical to understand the structure and parts of a tweet to analyse it. This review paper briefly describes each of these components and phases of sentiment analysis.

REFERENCES

- [1]. Levy, M. (2016). Playing with Twitter Data. [Blog] R-bloggers. Available at: <https://www.r-bloggers.com/playing-with-twitter-data/> [Accessed 7 Feb. 2018].
- [2]. Popularity Analysis for Saudi Telecom Companies Based on Twitter Data. (2013). Research Journal of Applied Sciences, Engineering and Technology. [online] Available at: <http://maxwellsci.com/print/rjaset/v6-4676-4680.pdf> [Accessed 1 Feb. 2018].
- [3]. Zhao, Y. (2016). Twitter Data Analysis with R – Text Mining and Social Network Analysis. [online] University of Canberra, p.40. Available at: <https://paulvanderlaken.files.wordpress.com/2017/08/rdataminingslides-twitter-analysis.pdf> [Accessed 7 Feb. 2018].
- [4]. Alrubaiee, H., Qiu, R., Alomar, K. and Li, D. (2016). Sentiment Analysis of Arabic Tweets in e-Learning. Journal of Computer Science. [online] Available at: <http://thescipub.com/PDF/jcssp.2016.553.563.pdf> [Accessed 7 Feb. 2018].
- [5]. Qamar, A., Alsuhibany, S. and Ahmed, S. (2017). Sentiment Classification of Twitter Data Belonging to Saudi Arabian Telecommunication Companies. (IJACSA) International Journal of Advanced Computer Science and Applications, [online] 8. Available https://thesai.org/Downloads/Volume8No1/Paper_50Sentiment_Classification_of_Twitter_Data_Belonging.pdf [Accessed 1 Feb. 2018].
- [6]. R. M. Duwairi and I.Qarqaz, "A framework for Arabic sentiment analysis using supervised classification", Int. J. Data Mining, Modelling and Management, Vol. 8, No. 4, pp.369-381, 2016.
- [7]. Hossam S. Ibrahim, Sherif M. Abdou, Mervat Gheith, "Sentiment Analysis For Modern Standard Arabic And Colloquial", International Journal on Natural Language Computing (IJNLC), Vol. 4, No.2, pp. 95-109, April 2015.

- [8]. Assiri, A., Emam, A. and Al-Dossari, H. (2016). Saudi Twitter Corpus for Sentiment Analysis. International Journal of Computer and Information Engineering, [online] 10. Available at: <http://waset.org/publications/10003483/saudi-twitter-corpus-for-sentiment-analysis> [Accessed 1 Mar. 2018].
- [9]. L. Wasser and C. Farmer, "Sentiment Analysis of Colorado Flood Tweets in R", Earth Lab, 2018. [Online]. Available: <https://earthdatascience.org/courses/earth-analytics/get-data-using-apis/sentiment-analysis-of-twitter-data-r/>. [Accessed: 01-Mar-2018].
- [10]. D. Robinson, "Text analysis of Trump's tweets confirms he writes only the (angrier) Android half", Variance explained, 2016.
- [11]. A Common Database Interface (DBI)", cran.r-R, 2003. [Online]. Available: <https://cran.rproject.org/web/packages/DBI/vignettes/DBI-1.html>. [Accessed: 25- Mar-2018].
- [12]. V. Kharde and S. Sonawane, "Sentiment Analysis of Twitter Data: A Survey of Techniques", International Journal of Computer Applications, vol. 139, p. 11, 2016.
- [13]. Product Quality Assessment using Opinion Mining in Persian Online Shopping Fatemeh HosseinzadehBendarkheili; Rezvan MohammadiBaghmolaei; Ali Ahmadi 2019 27th Iranian Conference on Electrical Engineering (ICEE) Year: 2019 | Conference Paper | Publisher: IEEE
- [14]. News Comments Modeling for Opinion Mining: The Case of Senegalese Online Press Lamine Faty; Marie Ndiaye; Edouard Ngor Sarr; Ousmane Sall; Sény Ndiaye Mbaye; Tony Tona Landu; Babiga Birregah; Mamadou Bousso 2020 International Conference on Advances in Computing and Communication Engineering (ICACCE) Year: 2020 | Conference Paper | Publisher: IEEE
- [15]. Chinese Explanatory Opinion Relationship Recognition Based on Improved Target Attention Mechanism Xue Cao; Chenghao Zhu; Chengguo Lv 2020 3rd International Conference on Advanced Electronic Materials, Computers, and Software Engineering (AEMCSE) Year: 2020 | Conference Paper | Publisher: IEEE
- [16]. Opinion Mining Using Live Twitter Data Andleeb Aslam; Usman Qamar; Reda Ayesha Khan; Pakizah Saqib; Aleena Ahmad; Aiman Qadeer 2019 IEEE International Conference on Computational Science and Engineering (CSE) and IEEE International Conference on Embedded and Ubiquitous Computing (EUC) Year: 2019 | Conference Paper | Publisher: IEEE