

Sentiment Analysis by using Machine Learning

Krishna Koli¹, Ganesh Rupnar¹, Mahesh Deshmukh¹ and S. R. Waghchaware²

UG Students, Department of Electronics & Telecommunication Engineering¹

Faculty, Department of Electronics & Telecommunication Engineering²

Punyashlok Ahilyadevi Holkar Solapur University Solapur, Maharashtra, India

SVERI'S College of Engineering, Pandharpur, India

Abstract: *The Many businesses are using social media networks to deliver different services and connect with clients and collect information about the thoughts and views of individuals. Sentiment analysis is a technique of machine learning that senses polarities such as positive or negative thoughts within the text, full documents, paragraphs, lines, or subsections. Machine Learning (ML) is a multidisciplinary field, a mixture of statistics and computer science algorithms that are commonly used in predictive and classification analyses. The goal and primary objectives of this project is to analytically categorize and analyze the prevalent research techniques and implementations of Machine Learning techniques to Sentiment Analysis on various applications.*

Keywords: ML, NLP, CNN

I. INTRODUCTION

Due Social networking channels like Twitter, Facebook, Instagram, and Whats App have stormy contact environments, it is imperative to relay sensitive knowledge about people's opinions, moods and feelings on any product, concept or policy through these social network channels. To both customers and suppliers, this data is valuable. During any online shopping, consumers

usually check other people's opinions about the product. Based on the customer's sentiment, the manufacturer can learn about its product benefits and drawbacks. Although both business organizations and individuals can get profit from these opinions, the sheer number of these opinions on text data is daunting for the users. For researchers, it is a very interesting area to examine and sum up the opinions conveyed in this broad opinion text content. This modern area of study is also known as Sentiment Analysis or Opinion mining.

II. LITERATURE REVIEW

In general opinion research at the starting of the 20th Century, the science of sentiment analysis and opinion mining has a strong basis. When online product reviews were required and accessible in the middle of 2000, they finally became a major research subject. Just 101 articles

on this subject were published in 2005, while almost 5,699 were published in 2015. This means that over a decade sentiment analysis has increased almost 50 times, making it one of the most quickly expanding fields of study in previous year. Throughout the early days of the internet, a

person was able to seek feedback from his friends, neighbors and relatives before taking any decision. Opinion sampling, surveys, and general public opinion on its products or services were conducted by organizations. As the World Wide Web has come and particularly with the

production and adoration of Web 2.0, where the focus on content generated by users has changed significantly the way the individual expresses his opinion or views. Now people can offer their thoughts, opinions, feelings, blogs, social platforms, forums, and reviews on their own personal web pages. Thanks to rich and diverse data generated in Web 2.0 applications, the field of opinion mining has advanced quickly (A. Kumar & Sebastian, 2012). Research into the shifts in the subjects found that social networking such as Twitter and Facebook are more focused on the most recent articles from the year 2014 to 2016. In recent year's mobile devices, stock markets, and human emotions were other topics that have become popular.

III. DATA COLLECTION

One of the most vital and significant components of the sentiment analysis application is data collection. Due to the widespread use of machine learning models, higher performance cannot be guaranteed just by using huge datasets for a task that is specialized to a given area. The quality of the dataset and labeling/annotation have an impact on the model's performance. Automatic predictions are likely to reflect the human disagreement discovered during annotation as ML models gain knowledge from the data they are trained on. As a result, it is crucial to have a correct set of instructions for annotating data. Customer sentiment analysis involves collecting, analyzing, and leveraging data to understand customers' feelings. This article focuses on how to collect data for customer sentiment analysis. Customer sentiment analysis is the practise of utilizing machine learning (ML) to analyse customer feedback from reviews, forums, surveys, and other sources to determine the intent and opinion of customers regarding a brand. Customer experience data sentiment analysis provides businesses with comprehensive insight into the drivers of purchasing decisions, trends in the ebb and flow of brand sentiment depending on timelines or events, and market-gap analysis that can aid in the enhancement of products and services.

Ways to collect data for sentiment Analysis:

1. Observations and videos on social media : One way to gather recent client input about your business, including your product and service, is through social media listening. The best bet for utilizing this data source is a sentiment analysis model that can examine and interpret both video and social media comments.
2. Beyond NPS, CES, and CSAT: Quantitative Surveys :You must look beyond quantitative data in order to obtain accurate client sentiment information. And to do that, you must examine comments and open-ended survey answers that don't have a predetermined answer. This enables clients to leave open-ended feedback, which might provide you with knowledge about elements of your organization.
3. Examine testimonials from online forums and retail websites: By scouring forums like Reedit and product review websites like GoogleMyBusiness, you can also find a variety of client feedback information. Importantly, the audience that various platforms attract means that using insights from
4. Data on the voice of the consumer from unconventional sources:Customer emails, chat bot histories, customer care transcripts, and other non-traditional sources of customer feedback data are excellent places to look for insights on the customer experience. One benefit of these sources is that your customer relationship management (CRM) software already contain all of this data.
5. Review the news and podcasts: You may gain detailed insights into brand performance and perception with news data that includes both articles and news videos and podcasts. A company can use market input from news sources to help with efficient public relations (PR) initiatives for managing brand reputation.

IV. MACHINE LEARNING ALGORITHM

Sentiment analysis with naive Bayes:

Naive Bayes is a series of reasonably straightforward probabilistic algorithms that determines the likelihood that a word or phrase should be categorized as positive or negative for the purposes of sentiment analysis. In essence, this is how the Bayes theorem operates. The likelihood of A if B is true is equal to the likelihood of B if A is true multiplied by the likelihood of A being true divided by the likelihood of B being true.

$$P(A|B) = [P(B|A) * P(A)] / P(B)$$

Deep learning: Deep learning is a branch of machine learning that uses "artificial neural networks" to simulate how the human brain calculates data. Machine learning with a hierarchy is deep learning. In other words, it has multiple levels and enables a machine to 'chain' several human-made operations together automatically. Deep learning may tackle complex problems in the same way that humans do by allowing numerous algorithms to be utilized gradually while proceeding from step to step.

Linear Regression:

Given X features, a statistical procedure called linear regression is used to forecast a Y value. The data sets are evaluated using machine learning to reveal a relationship. In order to forecast additional relationships, the associations are then

plotted along the X/Y axis with a straight line passing across them. The relationship between the X input (words and phrases) and the Y output (polarity) is determined by linear regression. On a scale of polarity that ranges from "really positive" to "really negative," as well as everywhere in between, this will determine where words and phrases fall.

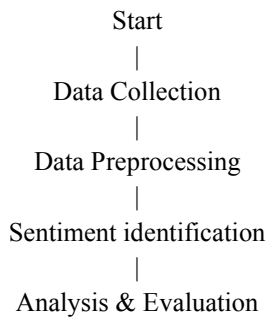
Support Vector Machines (SVM): Another supervised machine learning model that is more complex than linear regression is the support vector machine. The hyperplane that best separates the tags is then assigned by the SVM. This is just a line (like in linear regression) in two dimensions. Everything is red on one side of the line and blue on the other. Positive and negative would be used for sentiment analysis.

IV. STEPS REQUIRED TO TRAIN MODEL

1. Collect data set
 - a. Preprocessing
 - b. Stopword removal
 - c. Non Standard to standard word mapping
 - d. Tagging of Po's
 - e. Positive/Negative words tagging
 - f. Stemming
2. Utilize NLP to transform textual data into vector form.
3. Create training and testing groups for the dataset.
4. ML classifier training with training data. SVM, logistic regression, multi nominal Naive Bayes, random forest, and other techniques are recommended.
5. Determine the polarity of test results.
6. Use accuracy, precision, recall, and F1-score as evaluation measures for models.
7. To choose the best model, carry out the processes for algorithm selection and model selection.

The text data would have to be manually labeled before being used to train the machine learning models. For instance, machine learning models are trained using labeled sentiment data for sentiment analysis related to tweets and reviews.

V. ALGORITHM STEPS



VI. PROPOSED METHODOLOGY

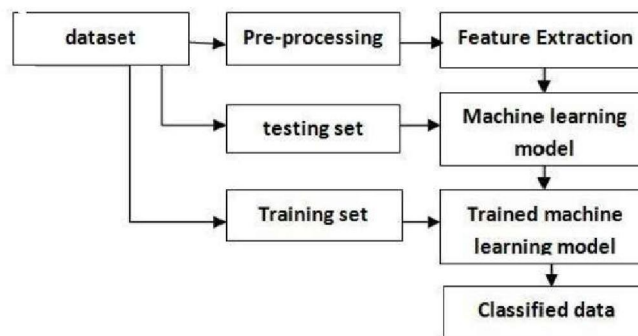


Fig.1. Proposed Methodology

VII. RESULT ANALYSIS

Sentiment analysis, often known as opinion mining, is a natural language processing (NLP) method for identifying the positive, negativity, or neutrality of data. Businesses frequently do sentiment analysis on textual data to track the perception of their brands and products in customer reviews and to better understand their target market. Classifying the polarity of textual data whether it is positive, negative, or neutral is the main goal of sentiment analysis. Numerous various business concerns can be resolved by determining whether the end-user mood is favourable, unfavourable, or neutral. Although the material that needs to be analysed for sentiment might come from a variety of sources, in the current context, sentiment analysis mostly concerns tweets and reviews.



Fig.2. Hidden Sentiment & Emotions

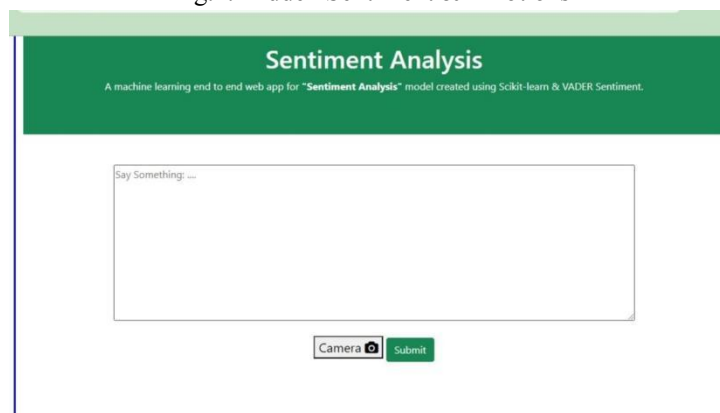


Fig.3. Output Window for entering text input



Fig.4 . Positive Result of Entered Text



Fig. 5. Facial Emotion -Neutral



Fig.6.Facial Emotion - Happy



Fig.7. Facial Emotion-Sad

VIII. CONCLUSION

Sentiment analysis is a method for determining the text's emotional undertone. It can be used to determine whether a piece of writing has good, negative, or neutral feelings. Business owners that want to comprehend how their clients feel about their brand may find this knowledge valuable. Sentiment analysis is a method for determining the text's emotional undertone. It can be used to determine whether a piece of writing has good, negative, or neutral feelings.

Business owners that want to comprehend how their clients feel about their brand may find this knowledge valuable. Understanding the tone of your customers' reviews and feedback will help you improve the areas that are causing them to be dissatisfied and foster more client loyalty. Sentiment analysis is a tool that marketers may use to better analyse client feedback and modify their plans as necessary. It can also be used to analyse whether a certain campaign or product has a

favourable or bad impact on customers. Companies can gain a deeper understanding of their target demographic through intelligence from the appropriate sentiment analysis data sources, which can result in more successful commercial strategies.

Particularly in an era of rapid innovation and evolving market dynamics, AI and data science are of utmost relevance to marketing activities. Customer feedback data that has been directly collected from them, along with customer sentiment analysis, can provide you all the power you need to ensure that your marketing plan is sustainable for long-term success. In this Machine learning sentiment analysis project & paper we tried to identify the hidden sentiment from the users entered text and the second part is that we have successfully classified the hidden sentiment or emotions from facial expressions whether the facial expression is angry, sad, happy, surprise & so many.

REFERENCES

- [1]. International Journal of Applied Engineering and Management Letters (IJAEML), ISSN: 2581-7000, Vol. 4, No. 2, August 2020.
- [2]. Neha S. Joshi et al, / (IJCSIT) International Journal of Computer Science and Information Technologies, Vol. 5 (4) , 2014, 5422-5425
- [3]. M.Govindarajan, Sentiment Analysis of Movie Reviews using Hybrid Method of Naive Bayes and Genetic Algorithm , International Journal of Advanced Computer Research (ISSN (print): 2249-7277 ISSN (online): 2277-7970), Volume-3 Number-4 Issue-13 December-2013
- [4]. David Zimbra, M. Ghiassi and Sean Lee, “Brand-Related Twitter Sentiment Analysis using Feature Engineering and the Dynamic Architecture for Artificial Neural Networks”, IEEE 1530-1605, 2016.
- [5]. Varsha Sahayak, Vijaya Shete and Apashabi Pathan, “Sentiment Analysis on Twitter Data”, (IJIRAE) ISSN: 2349-2163, January 2015.