Sentiments of Public Opinion

Aniket Sharadrao Gudankwar*, Pratik Milind Mendhe, Lukesh Namdeo Oghare, Atharva Ramu Yemde

Computer Science & Engineering, Guru Nanak Institute of Technology, Nagpur, Maharashtra, India

ARTICLEINFO

ABSTRACT

Sentiment analysis, also known as opinion mining, is a computational technique that aims to identify and extract subjective information from text data, allowing us to understand and categorize the underlying sentiment expressed by individuals. With the exponential growth of social media, online reviews, and user generated content, sentiment analysis has gained significant attention as a valuable tool for businesses, researchers, and decision-makers to gain insights into public opinion, customers satisfaction, and market trends. The project is about searching the web for public opinion about people, movies, places, companies or any entities. The user of this software can type in a query that is any entity and the system mines the web for public opinion about that entity and generates a sentiment score. A positive score represents a positive public sentiment about that entity and negative score represents a negative sentiment or opinion.

Keywords: Sentiment Analysis, Opinion Mining, Computational Linguistics, Machine Learning, Deep Learning, Social Media Analysis.

I. INTRODUCTION

Sentiment is an attitude, thought, or judgment prompted by feeling. Sentiment analysis, which is also known as opinion mining, studies people's sentiments towards certain entities. From a user's perspective, people are able to post their own content through various social media, such as forums, micro-blogs, or online social networking sites. From a researcher's perspective, many social media sites release their application programming interfaces (APIs), prompting data collection and analysis by researchers and developers. However, those types of online data have several flaws that potentially hinder the process of sentiment analysis. The first flaw is that since people can freely post their own content, the quality of their opinions cannot be guaranteed. The second flaw is that ground truth of such online data is not always available. A ground truth is more like a tag of a certain opinion, indicating whether the opinion is positive, negative, or neutral. SENTIMENT ANALYSIS is a kind of text classification based on Sentimental Orientation (SO) of opinion they contain. Sentiment analysis of product reviews has recently become very popular in text mining and computational linguistics research.

- Firstly, evaluative terms expressing opinions must be extracted from the review.
- Secondly, the SO, or the polarity, of the opinions must be determined.
- Thirdly, the opinion strength, or the intensity, of an opinion should also be determined.
- Finally, the review is classified with respect to sentiment classes, such as Positive and Negative, based on the SO of the opinions it contains.

II. METHODS AND MATERIAL

Sentiment analysis is the process of determining the sentiment or emotion expressed in a piece of text. Here are the general steps involved in sentiment analysis:

Data Collection: Gather the textual data that you want to analyze. This can include social media posts, customer reviews, surveys, or any other source of text data.

Text Preprocessing: Clean and preprocess the text data to remove noise and irrelevant information. This step typically involves removing punctuation, converting text to lowercase, removing stop words (commonly used words like "the," "and," etc.), and handling issues like spelling mistakes or abbreviations.

Tokenization: Break down the text into smaller units called tokens, such as words, phrases, or even characters. Tokenization is the process of splitting the text into meaningful units to analyze each unit separately.

Feature Extraction: Identify the relevant features or attributes that can help in determining sentiment. This can include word frequencies, n-grams (sequences of adjacent words), parts of speech, or any other linguistic or contextual information that might be useful in capturing sentiment.

Sentiment Classification: Use a machine learning or natural language processing (NLP) technique to classify the sentiment of each token or the overall sentiment of the text. Common approaches include using rule-based methods, lexicon-based methods (using sentiment dictionaries), or supervised learning techniques like

Naive Bayes, Support Vector Machines (SVM), or deep learning models like Recurrent Neural Networks (RNN) or Convolutional Neural Networks (CNN).

Model Training and Evaluation: If you're using a machine learning approach, you'll need to train your model using a labeled dataset where sentiments are already annotated. Split your dataset into training and testing sets, train your model on the training set, and evaluate its performance on the testing set. This step helps you measure the accuracy and effectiveness of your sentiment analysis model.

Sentiment Analysis Output: After training and evaluating the model, you can apply it to new, unseen text data to obtain sentiment analysis results. The output can be binary (positive/negative), multi-class (positive/negative/neutral), or even a continuous sentiment score.

Post-processing and Analysis: Analyze the sentiment analysis results to gain insights and draw conclusions from the data. This could involve aggregating sentiment scores, visualizing sentiment trends over time, or correlating sentiments with other factors of interest.

III. RESULTS AND DISCUSSION

Our results, compared to the performance of other terms:

SN	Methods Used	Data Size	Accuracy Obtained
1	Bag of words model	N.A.	39.80%
2	NaAfve Bayes model	10,000 tweets	42.51%
3	Support Vector Machine	1.6 Million tweets	69.02%
4	SVM + Bag of Words Model (Our Results)	9000 tweets	68.36%

IV. CONCLUSION

Sentiment analysis deals with the classification of texts based on the sentiments they contain. This focuses on a typical sentiment analysis model consisting of three core steps, namely data preparation, review analysis and sentiment classification, and describes representative techniques involved in those steps. Sentiment analysis is an emerging research area in text mining and computational linguistics, and has attracted considerable research attention in the past few years. Future research shall explore sophisticated methods for opinion and product feature extraction, as well as new classification models that can address the ordered labels property in rating inference. Applications that utilize results from sentiment analysis is also expected to emerge in the near future.

V. REFERENCES

[1]. S. ChandraKala1 and C. Sindhu2, "OPINION MINING AND SENTIMENT CLASSIFICATION: A SURVEY," Vol.3(1), Oct 2012, 420-427

[2]. G. Angulakshmi, Dr. R. Manicka Chezian ,"An Analysis on Opinion Mining: Techniques and Tools". Vol 3(7), 2014 www.iarcce.com.

[3]. Callen Rain, "Sentiment Analysis in Amazon Reviews Using Probabilistic Machine Learning" Swarthmore College, Department of Computer Science.

[4]. Padmani P .Tribhuvan, S.G. Bhirud, Amrapali P. Tribhuvan," A Peer Review of Feature Based Opinion Mining and Summarization"(IJCSIT) International Journal of Computer Science and Information Technologies, Vol. 5 (1), 2014, 247-250 ,www.ijcsit.com

[5]. Zhu, Jingbo, et al. "Aspect-based opinion polling from customer reviews." IEEE Transactions on Affective Computing, Volume 2.1,pp.37-49, 2011.

[6]. Li, Shoushan, Zhongqing Wang, Sophia Yat Mei Lee, and Chu-Ren Huang. "Sentiment Classification with Polarity Shifting Detection." In Asian Language Processing (IALP), 2013 International Conference on, pp. 129-132. IEEE, 2013