Contemporary Trends in Research on Sentiment Analysis

Silvia Das

Table Of Content

Abstract

Sentiment analysis research has been started in 2000s and now it is one of the most demanding field in research. Sentiment analysis is a rapid growing field in Natural Language Processing that aims to extract subjective information from text. Sentiment analysis has gained attention in recent years for understanding human emotion or opinion expressed in the form of textual information. Over the last few years researchers and practitioners they discuss sentiment analysis using machine learning and deep learning method. And Sentiment analysis is being studied into three levels aspect level, sentence level and document levels. The application of sentiment analysis is spread in every domain such as political field, marketing, business purpose and the most commonly used in social media. This book chapter is all about thorough review of sentiment analysis, level of sentiment analysis, applications, methods and techniques.

1. Introduction

Sentiment analysis is one of the most famous research area in computer science. Sentiment analysis can be defined as the ability to understand the emotions of users .Sentiment analysis also called as opinion mining, emotion analysis and opinion extraction. Opinion mining that analyses users' sentiment, opinion and emotions from text data. Sentiment analysis involves analysing the context, phrases, and words in the text to identify whether the sentiment or opinion expressed is positive, negative or neutral. It also considers the feelings that is happy, angry, sad, etc. It is an external of natural language processing (NLP) that has been widely studied in data mining and text mining in recent years. NLP is introduced in early 1940s, at this time the task was not as easy as people imagined, and some researchers have identified major issues in development field of NLP. Later on 1960s basic natural language processing was successfully developed. NLP is the field of Artificial Intelligence (AI), which is used to extract the information from data. The main goal of natural language processing is to understand various languages, which is used to process them and extract information from them. Sentiment analysis is a most powerful tool for an analysing social media data. Sentiment is defined as an attitude, thought by feeling, whereas opinion is defined as a judgement in the mind about a particular matter (According to Merriam-Webster). The decisions indicate

that an opinion is more of a person's particular view about something, where as a sentiment is more about feeling.

1.1Sentiment Analysis can be classified into various levels. Some common types of sentiment analysis:

- **Sentence level sentiment analysis** - The opinion of each sentences within a document is analysed, the task is to determine whether each sentences expressed a positive, negative or neutral opinion. Sentence level sentiment analysis is more about to subjectivity classification .

- **Document level sentiment** analysis - Document level sentiment analysis is the task of whole opinion expressed in a piece of text such as document, then determines positive or negative sentiment (Pang, Lee and vaithyannathan, 2002; Turney, 2000). It perform opinion on a single entity. This type of sentiment analysis takes part social media post, article, and review, into whether the document extract positive, negative or neutral opinion.

- **Aspect based Sentiment Analysis** - Aspect based sentiment analysis identify the sentiment towards the aspect within the text. Aspect level sentiment analysis provides more detailed analysis by identifying different aspect from the text information. In aspect level sentiment analysis we need to split our text into different aspect then we analyse the sentiment. For example, "The saree looks beautiful but the fabric is cheap" in this sentence two aspects are there one is the positive as the appearance and another is the material is negative.

- **Emotion based sentiment analysis** - Emotion based sentiment analysis detect emotion like happiness, sadness, anger, fear, rather than positive or negative sentiment.

1.2 Sentiment Analysis Applications:

Reviews are important to organizations as well as business purpose because they always want public opinions about products and services. Nowadays,

organizations are widely using the social media for decision making. If a customer wants to buy product online, they don't need to ask their family or friends opinions because there are many user reviews in public forums on the web about the particular product. In recent years, sentiment analysis applications are spread in every domain like customer support and feedback analysis, political analysis, market research, healthcare, hotel and restaurant reviews, online advertisement. Thousands of companies including both start-up and MNC now operate in this field. Now we will learn details about the customer support and feedback analysis, political analysis, market research, healthcare, hotel and restaurant reviews, online advertisement customer support and feedback analysis, political analysis, market research, healthcare, hotel and restaurant reviews, online advertisement.

Customer support and feedback analysis- Sentiment analysis identifying the need of customers requirement and recommend them the proper thing. Sentiment analysis is also help in e-commerce services for developing countries (Pare Dj, 2003). Mackey Tk, Miner A, Cuomo RE they explore the sentiment analysis framework can be trained to analyse the essential information (2015).

Political analysis- Sentiment analysis detect the political content. The content can be text, videos and images.

Market Research- In the market research sentiment analysis analyse the customer review feedback to understand market trends.

Healthcare- Patient reviews and feedback to understand the quality of healthcare services provided by healthcare centre. Jimenez-Zafra SM, Marin-Valdivia MT, Molina-Gonzalez MD , Urena-Lopez they found the challenges in applying sentiment analysis in medical science(2021).

Online Advertisement- Advertisers can use sentiment analysis to ensure their desired sentiment and are well-received by the audience.

Hotel and Restaurant Reviews- To improve customer experience by analysing reviews and feedback from platforms like Google My Business, Zomato, The infatuation and TripAdvisor.

1.3 Related work

Many researchers have contribute in sentiment analysis. A brief discussion on the work done previously on sentiment analysis is given in this section.

Moulaei ME, Abadeh MS, Keshavarz H suggest a technique for adaptive aspect-based lexicons for sentiment classification. The authors described two strategies for construction two dynamic lexicons to aid in the classification sentiments depending on their aspects : a strategy based on statistic and genetic algorithm.

Subhasini L, Li y, Z hang J, Atukorale AS, Wu y presents the results of a comprehensive review of contemporary opinion mining literature. It also covers how to extract text features from opinions with noise or uncertainly represent knowledge in opinions, and categorize them.

Sanchez-Rada JF, Iglesias CA stated that sentiment analysis has gained widespread acceptance in recent years, not just among researchers but also among business, governments and organizations.

Lighthart A, Catal C, Tekinerdogan B published an overview an opinion mining in the earlier stage.

Piryani R, Madhavi D, Singh VK discusses the study topic from 2000 to 2015 and provides a framework for computationally processing instructured data with the primary goal of extracting views and identifying their moods.

Hangya V, Farkas R both present a brief overview of machine learning algorithms used in social media analysis.

Balaji T, Annavarapu CSR, Bablani A conducted a thorough examination of the several applications of social media analysis utilizing sophisticated machine learning algorithms.

Yue L, Chen W, Li x, Zuo W, Yin M, conducted research on the effectiveness of internet reviews.

Jain PK, Pamula R, Srivastava G discuss machine learning applications that incorporate online reviews in sentiment categorization, predictive decision-making and the detection of false reviews.

Yousif A, Niu Z, Tarus JK, Ahmad A described the problem of sentiment analysis and suggested potential directions.

2. Methods

There are several methods and approaches to perform on sentiment analysis from traditional rule based methods to advanced machine learning approach.

2.1 Rule-Based Methods. Rule based approach in sentiment analysis involves the sentiment or emotion expressed in a piece of text. The sentiment of a document is determined based on positive and negative words.

2.2 Lexicon Based Methods. Lexicon-based approaches rely on the assumption that the text semantic orientation is related to the polarity of words, this is related to content words, phrases and sentences. The sentiment score of a document is calculated by aggregating the scores of its component words.

2.3 Supervised Methods. The most presiding methods are based on sequential labelling, supervised techniques, and manually labelled data for training.

3. Sentiment Analysis in Natural Language Processing

Sentiment analysis is seen as subfield of NLP, interact between computer and human .Present day NLP is based on machine learning algorithm. Sentiment analysis using NLP is a very promising area of research as well as applications. NLP is able to automatically extract the sentiment expressed in data, especially in order to determine whether the writer's attitude towards a particular topic is positive, negative or neutral. Sentiment analysis is most commonly used to analyse social media posts. Various issues are associated with natural language processing, such as sarcasm, irony, language specific challenges.

3.1 Feature Selection

Feature selection is important to identifying relevant features in dataset. Ritter A, Etzioni O, Clark S they present open domain event extraction from twitter (2011). Razon A, Barnden J proposed automated text readability classification based on concept indexing with Parts-Of-Speech n-gram features (2015). N-grams are contiguous sequences of n items from text. When two features are combined, the technique is called “Bi-gram”, and combination of three features technique is called “Tri-gram”.

3.2 Feature Extraction

Feature extraction is the process of transforming raw text data into a numerical representation. Feature extraction is to convert unstructured text data into structured data. Venugopalan M, Gupta D discuss sentiment analysis incorporated other features as it is challenging to extract features from text (2015).

Bag of Words is simplest method used in NLP, bag of words creates all unique words in corpus. Bag represents the vocabulary of words and each document is represented as vector.

Term Frequency – Inverse Document Frequency (TF – IDF) an extension of bag of words in a document is TF-IDF .Term frequency follows higher weight to words which appear more frequency in document. Term’s presence provides the value either 0 or 1.

Parts Of Speech Tagging is also known as grammatical tagging, the process of assigning the appropriate part of speech label to each word in a sentences. Straka M, Hajic J, Strakova J presents pos tagging, morphological analysis and parsing.

3.3 Algorithm Used In NLP

To perform nlp task the most two popular algorithm are Support Vector Machine (SVM) and Naïve Bayes, they have different approach to handle text data such as sentiment analysis.

Naïve Bayes – the definition of this algorithm is to classify text data into a predefined form and observing different features in each class. Basically the algorithm calculates the probabilities of a document that belongs to a particular class then assign the highest most probability.

Lopamudra Dey, Sanjay Chakraborty, Beepa Bose and Sweta Tiwari they worked on movie review and hotel review using Naïve Bayes Classifier (2016).

Support Vector Machine – Machine learning algorithm used for binary classification, SVM is design to find optimal hyperplane that separate data point belonging to different classes.

Zohreh Madhoushi, Abdul Razak Hamdan, Suhaila Zainudin they says sentiment analysis technique in recent works. In their research paper they used machine learning approach which contain support vector machine and naïve bayes classification (2021).

4. Challenges

Sarcasm detection is quite challenging, as the actual meaning of the words may communicate one sentiment while the intended meaning is completely opposite.

Unstructured Data suffer a predefined format, unstructured data doesn't follows pattern. So this is also challenging in sentiment analysis.

Subjectivity in Text. Comparison in positive, negative and neutral expression with varying tone is difficult to determine the sentiment.

5. Conclusion

In this book chapter, we discussed the overview of sentiment analysis, methodologies, applications, different levels of sentiment analysis and the different studies provided in research papers. Sentiment Analysis has proven useful resources in every industries especially in business field.

References

[1] C. C. T. B. Ligthart A, "Systematic reviews in sentiment analysis: a tertiary study," *Springer,* 2021.

[2] D. B. s. R.Piryani, "Analytical mapping of opinion mining and sentiment analysis research during 2000–2015," *Elsevier,* 2017.

[3] J. Fernando Sánchez-Rada,"Social context in sentiment analysis: Formal definition, overview of current trends and framework for comparison," Elsevier, 2019.

[4] Abdallah Yousif, Zhendong Niu, John K. Tarus ,Arshad Ahmad ," A survey on sentiment analysis of scientific citations," Springer, 2017.

[5] Marouane Birjali, Mohammed Kasri, Abderrahim Beni-Hssane, " A comprehensive survey on sentiment analysis: Approaches, challenges and trends," Elsevier, 2021.

[6] Mohammad Soleymani, David Garcia ,Brendan Jou, Björn Schuller,Shih-Fu Chang, Maja Pantic," A survey of multimodal sentiment analysis," Elsevier, 2017.

[7] Ashima Yadav, Dinesh Kumar Vishwakarma," Sentiment analysis using deep learning architectures: a review," Springer, 2019.

[8] Lin Yue, Weitong Chen, Xue Li, Wanli Zuo, Mingao Yin," A survey of sentiment analysis in social media," Springer, 2018.

[9] Praphula Kumar Jain, Rajendra Pamula, Gautam Srivastava," A systematic literature review on machine learning applications for consumer sentiment analysis using online reviews," Elsevier, 2021.

[10] Balaji T.K, Chandra Sekhara Rao Annavarapu, Annushree Bablani," Machine learning algorithms for social media analysis: A survey ," Elsevier, 2021.

[11] Viktor Hangya, Richárd Farkas ," A comparative empirical study on social media sentiment analysis over various genres and languages," Springer, 2017.

[12] L. D. C. S. Subhashini, Yuefeng Li, Jinglan Zhang, Ajantha S. Atukorale,Yutong Wu ," Mining and classifying customer reviews: a survey," Springer, 2021.

[13] Mohammad Erfan Mowlaei, Mohammad Saniee Abadeh, Hamidreza Keshavarz," Aspect-based sentiment analysis using adaptive aspect-based lexicons," Elsevier, 2020.

[14] Lopamudra Dey, Sanjay Chakraborty, Anuraag Biswas, Beepa Bose, Sweta Tiwari," Sentiment Analysis of Review Datasets Using Naive Bayes and K-NN Classifier," arXiv, 2016.

[15] Zohreh Madhoushi, Abdul Razak Hamdan, Suhaila Zainudin, " A Similarity Score Model for Aspect Category Detection," IJACSA, 2021.

[16] Milan Straka, Jan Hajič, and Jana Straková," UDPipe: Trainable Pipeline for Processing CoNLL-U Files Performing Tokenization, Morphological Analysis, POS Tagging and Parsing," ELRA, 2016.

[17] Manju Venugopalan, Deepa Gupta, “Exploring sentiment analysis on twitter data, “Eighth International Conference on Contemporary Computing (IC3), 2015.

[18] Alan Ritter, Sam Clark, Mausam and Oren Etzioni," Named Entity Recognition in Tweets: An Experimental Study," In Proceedings of the 2011 conference on empirical methods in natural language processing, 2011.

[19] Abigail R. Razon, John A. Barnden," A New Approach to Automated Text Readability Classification based on Concept Indexing with Integrated Part-of-Speech n-gram Features," Conference: Proceedings of Recent Advances in Natural Language Processing, 2015.

[20] Tim K. Mackey, Angela Miner, Raphael E. Cuomo, "Exploring the e-cigarette e-commerce marketplace: Identifying Internet e-cigarette marketing characteristics and regulatory gaps," Elsevier, 2015.

[21] Bo Pang, Lillian Lee, Shivakumar Vaithyanathan,"
Thumbs up? Sentiment Classification using Machine Learning Techniques, “arXiv, 2002.

[22] Peter D. Turney ," Learning Algorithms for Keyphrase Extraction," Springer, 2000.