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 rapidly expanding domain in Natural Language Processing that seeks to derive subjective data 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 previous few years researchers and practitioners they talk sentiment analysis with the help of deep learning models and machine learning algorithms. And Sentiment analysis is being studied into three stages aspect level, sentence level and document levels. 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 overview of sentiment analysis, level of sentiment analysis, applications, methods and techniques.

1. Introduction

One of the most famous research area in computer science is sentiment analysis. The definition of sentiment analysis can be defined as the ability to understand the emotions of users. Sentiment analysis additionally known as opinion mining, emotion analysis and opinion extraction. Opinion mining is the method of analysing 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. Sentiment analysis is a subfield of natural language processing (NLP) and it is the most powerful application of NLP. Natural language processing has been extensively studied in data mining and text mining. 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) that 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 of data. Sentiment is described as an attitude and more about feeling.

1.1 Sentiment Analysis can be classified into various levels. Some common types are :

- **Sentiment Analysis Based On Sentence :** The opinion of each sentences within a document is analysed, and determine the polarity of each sentence. Sentence level sentiment analysis is more about to subjectivity classification.
- **Sentiment analysis Based On Document :** It implements opinion on a single entity. Document level sentiment analysis involves the overall sentiment from a whole document or it can determine any other piece of text then shows the result is positive, negative or neutral sentiment. This type of sentiment analysis focus on review and article.
- **Sentiment Analysis Based On Aspect :** It Aims to detect the sentiment associated with different aspect in a text. Aspect based sentiment analysis breaks down the text into different aspect and determine the text is positive, negative or neutral. 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. ABSA provides more detailed analysis.
- **Sentiment Analysis Based on Emotion :** However traditional sentiment analysis focus on classifying text as positive or negative or neutral, emotion based sentiment analysis focus on psychological conditions 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. Pare dj (2003) stated that sentiment analysis is very useful in e-commerce service area for emerging countries. Later Mackey Tk, Miner A, Cuomo RE (2015) they explore the sentiment analysis framework can be trained to analyse the essential information.

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 (2021) they " found the challenges in applying sentiment analysis in medical science".

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.

"A technique for adaptive aspect-based lexicons for sentiment classification" by Moulaei ME, Abadeh MS, Keshavarz H. The authors defined techniques for construction dynamic lexicons to aid in the classification sentiment relying on their aspects a method primarily based totally 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".

"Sentiment analysis has gained widespread acceptance in recent years, not just among researchers but also among business, governments and organizations" proposed by Sanchez-Rada JF, Iglesias CA.

Lighthart A, Catal C, Tekinerdogan B published "an overview on 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 in structured data with the primary goal of extracting views and identifying their moods".

"A brief overview of machine learning algorithms used in social media analysis" present by Hangya V, Farkas R.

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

"The effectiveness of internet reviews" proposed by Yue L, Chen W, Li x, Zuo W, Yin M.

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

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

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: Involves the predefined rules to describe the sentiment expressed in a piece of text. Rule based methods rely on linguistic pattern and can provide reasonable result for context. It is a straightforward approach.

2.2 Lexicon Based Methods: Processes depend on the idea that the textual semantic orientation is associated with the polarity of words, this is related to

content words, phrases and sentences. Lexicon based methods are simple to implement.

2.3 Supervised Methods: The most powerful approach in sentiment analysis is supervised methods. Supervised methods deals with the large number of data and complex language patterns and delivers high performance.

3. Sentiment Analysis is a subfield of Natural Language Processing

Sentiment analysis is a natural language processing method used to determine the emotional tone of a given text. Mostly it interact between computers and human language. 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 relies on machine learning and deep learning.

3.1 Feature Selection

The goal is to take a relevant features from a data set. Ritter A, Etzioni O, and Clark S (2011) they present “open domain event extraction from twitter”. Razon A, Barnden J (2015) in their work they said “automated text readability classification based on concept indexing with Parts-Of-Speech n-gram features”. N-grams are contiguous sequences of n items from text. The combination of two features is called Bi-gram and the combination of three feature is called Tri-gram. In sentiment analysis Bi-gram and Tri-gram can help in classification of text, language model etc.

3.2 Feature Extraction

Transforming raw text data into a numerical representation (from unstructured text data into structure data) is the method of feature extraction. Venugopalan M, and Gupta D (2015) discuss “The challenges to extract features from text”. The most popular techniques of feature extraction are Bag of Words (BOW) and Term Frequency-Inverse Document Frequency (TF-IDF) and Parts Of Speech Tagging (POS).

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

Term Frequency and 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 involves information about grammatical parts of speech in a text. The tagging assigns tags to words based on their syntactic roles in a sentence.

3.3 Algorithm Used In NLP

To perform on 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 (2016) they worked on "movie review and hotel review using Naïve Bayes Classifier".

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.

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] 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.

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

[17] 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.

[18] 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.

[19] 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.

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

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