SURVEY ON METHODOLOGIES FOR STANCE DETECTION AND PRODUCT ASPECT MINING

[1]Dr. M. JAYAKANDAN
[1]Assistant Professor

[2]Mr. H. UTHUB OLI
[2]PG Student
II-M.Sc Computer Science

[3]Mr. B. MURALIDHARAN
[3] PG Student
II- M.Sc Computer Science

PG & RESEARCH DEPARTMENT OF COMPUTER SCIENCE,
E. G. S. PILLAY ARTS & SCIENCE COLLEGE (AUTONOMOUS),
NAGAPATTINAM-611 002, TAMIL NADU, INDIA.

ABSRACT:

The prevalence of social media and online forum, opinion mining, aiming at analyzing and discovering the latent opinion in user-generated reviews on the Internet, has become a hot research topic. This paper focuses on two important subtasks in this field, stance detection and product aspect mining, both of which can be formalized as the problem of the triple target, aspect, opinion extraction. In this paper, we first introduce the general framework of opinion mining and describe the evaluation metrics.Then, the methodologies for stance detection on different sources, such as online forum and social media are discussed. After that, approaches for product aspect mining are categorized into three main groups which are corpus level aspect extraction, corpus level aspect, and opinion mining, and document level aspect and opinion mining based on the processing units and tasks. And then we discuss the challenges and possible solutions. Finally, we summarize the evolving trend of the reviewed methodologies and conclude the survey.

KEYWORDS:

Opinion Mining, Stance Detection, Product Aspect Mining, Sentiment Analysis, Text Mining, Natural Language Processing, Social Media Analysis, Online Reviews, Triple Extraction.

INTRODUCTION:

The pervasiveness of online discussion forum and social media platform, user generated text containing opinions on some hot issues has increased significantly. Due to the large amount of such emotional reviews and posts on Internet, it is impossible for users to digest such information manually. Therefore, automatically mining opinion from online texts, aiming at discovering user concerned topics and the corresponding opinion, becomes essential. In general, opinion mining aims to extract a quintuple $<e, a, s, h, t>$ from texts, where e is the entity or the target, a is the aspect of the entity e, h is the opinion holder, t is the time when the opinion holder expresses her opinion on the entity e, and s is the opinion which h holds to the aspect a of the entity e at t.

For example, opinion mining processes the review text. Sentiment analysis cares more about the sentiment polarity s of the text, stance detection aims to identify the opinion s to the specific target e, and product aspect mining focus on extracting the aspect a and corresponding opinion s from text.

In the last decade, stance detection and product aspect mining have attracted many scholars. Following the general opinion mining framework, stance detection can be formalized as the task of extracting tuple $<e, s>$ (e means target and s represents opinion) without considering other elements. Stance detection focuses on detecting the user stance (favor, against) on a particular debate topic event. It is similar to sentiment analysis,

but with big difference. In specific, sentiment analysis aims to identify the sentiment polarity (positive, negative) of the text while stance detection cares about the stance on the target. For example, the tweet "Jeb bush is the only sane candidate in this republican lineup, I support him" will be assigned positive by sentiment analysis, but extracted with „against" stance to the topic "Donald Trump as President" by stance detection. Research on stance detection can be categorized into four groups based on debate settings, such as congressional floor debates, company-internal discussions, online forums ideological debates and hot-event oriented debates on social media. The latter two are open domain and flexible, therefore more challengeable. Furthermore, the participants prefer to use colorful and emotional language to express their viewpoints, such as the tweet "It looks like they like Hilary more and that plain stupid" related to the hash tag. Therefore, we concentrate on the prevalent work of stance detection for online debate forums and social media in this survey. Different from stance detection, product aspect mining aims at detecting relevant aspects and opinions. Following the general opinion mining framework, product aspect mining can be formalized as the task of extracting triple <e, a, s> (e means target, a and s represent aspect and opinion respectively). Based on text granularity, it could be categorized into corpus level and document/sentence level mining.
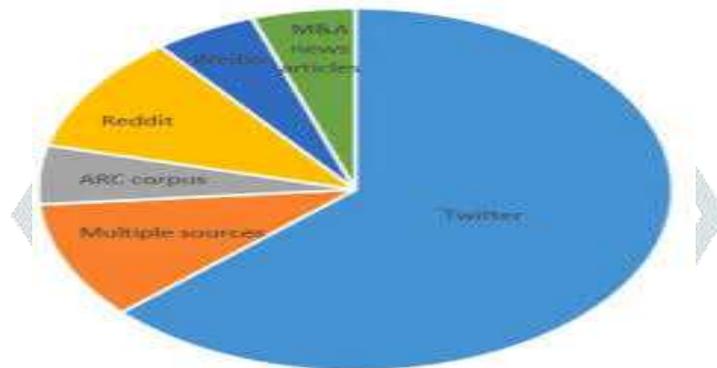

The corpus level mining could be further divided into two categories: aspect extraction, aspect and opinion mining. Corpus level aspect extraction aims to mine the aspect terms or aspect phrases in the corpus while ignoring where the aspects are discussed. Similarly, corpus level aspect and opinion mining extract both the aspects and the corresponding opinions without considering where they are expressed. Actually, corpus level mining pays attention to the aspects that most reviews are interested in and the corresponding opinions while document/sentence level mining focus on extracting aspect and opinion terms in a single review. After the first attempt of corpus level extraction, numerous approaches have been proposed which could be categorized as rule-based and unsupervised learning based models. Early aspect mining systems employ frequency pattern mining technique to extract aspect terms in the reviews.
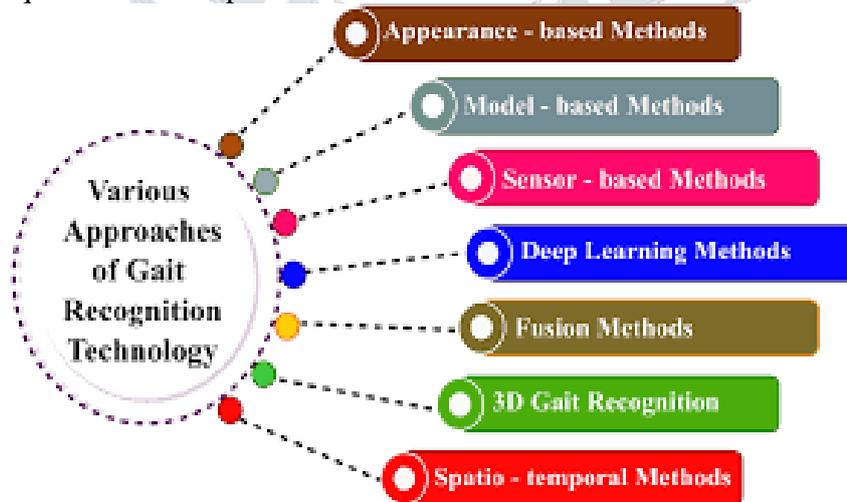
RELATED WORK:

Most existing approaches tackle the tasks using linguistic analysis or unsupervised learning topic models and clustering algorithms. Due to the arbitrariness of syntactic rules and the lack of necessary supervision in unsupervised model, those approaches have drawbacks such as aspect redundancy and poor performance.
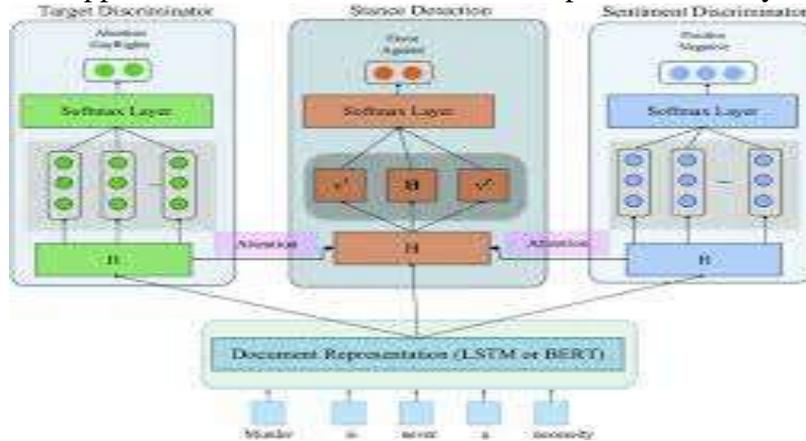

Thus, another group of researchers focus on the existing system methodologies in document/sentence level mining and view this task as a sequential tagging problem. This subsection will survey the corresponding supervised methodologies. Thus, the existing lexicons which just collect positive word sand negative words are not suitable for the task.
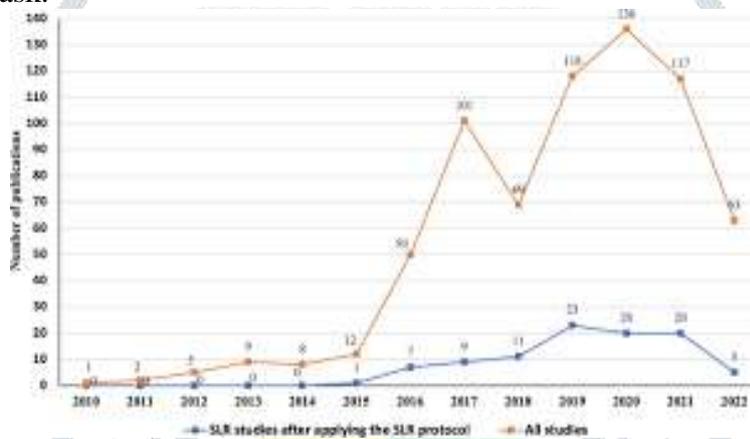

PROPOSED METHODOLOGY:

This paper proposed which could be categorized as rule-based and unsupervised learning based models. Early aspect mining systems employ frequency pattern mining technique to extract aspect terms in the reviews. To overcome the disadvantages of missing low frequency aspects and ignoring semantic similarity of aspects, unsupervised learning based approaches are proposed by casting the task into a clustering problem. Different from corpus level mining, document/sentence level aspect and opinion mining concentrates on detecting the aspects and opinions in each individual document or sentence. Thus, collective models which also employ the relationship between posts (e.g. disagreement, argument) or users are proposed.
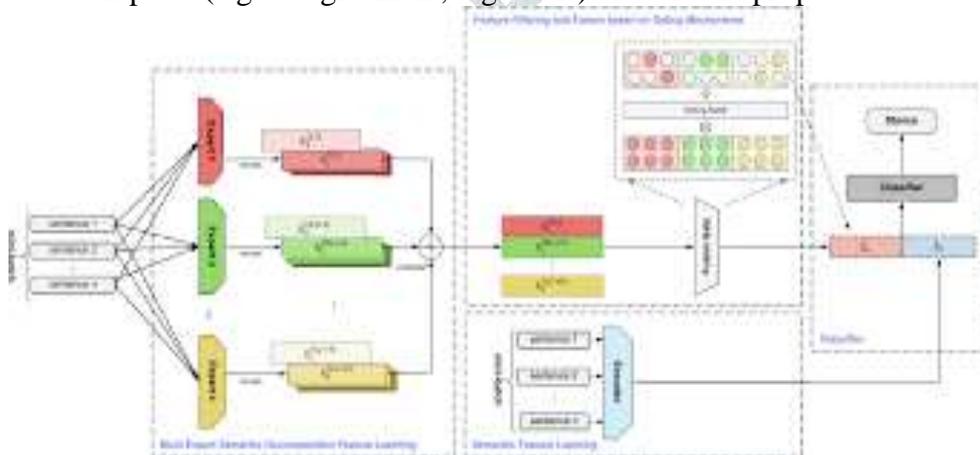

The proposed approach infers the probabilistic relationship between posts and the stance labels with good performance, and other relational information could be incorporate into the framework easily. the proposed method improves the performance in both user level and post level stance detection While knowledge based approaches have only recently begun to emerge and the knowledge use dare outdated, recently proposed representation techniques like and Bert might also help mining approaches to enhance the extraction ability.
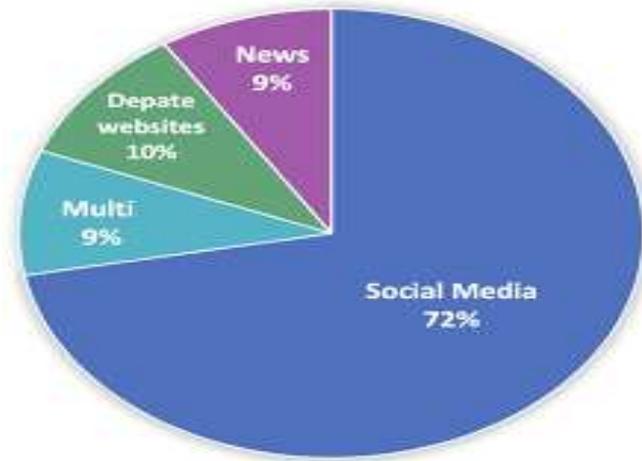
EXPERIMENTAL RESULTS AND DISCUSSION
## EXPERIMENTAL RESULTS:

Due to the large amount of such emotional reviews and posts on Internet, it is impossible for users to digest such information manually. Therefore, automatically mining opinion from online texts, aiming at discovering user concerned topics and the corresponding opinion, becomes essential.
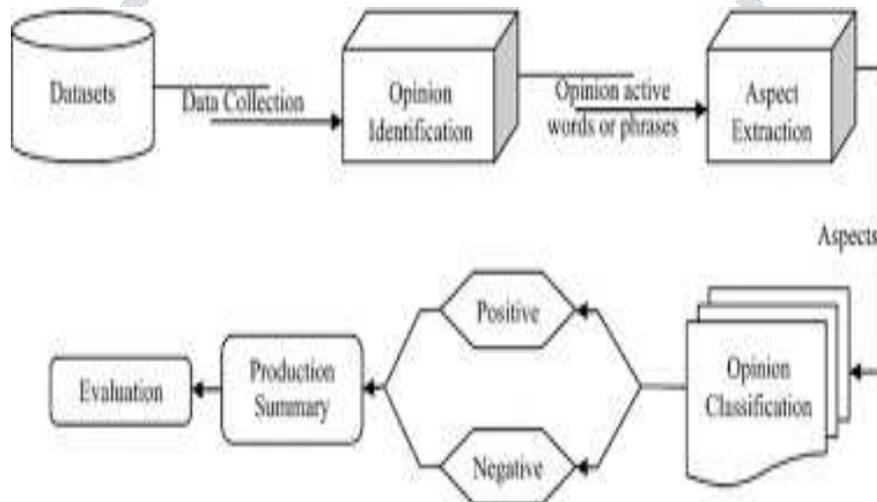
In this paper, we focus on the methodologies for two subtasks of opinion mining which are online stance detection and product aspect mining. Thus, we provide a brief survey and categorization on the previous methodologies. Earlier work mainly relies on frequency, relation rules and feature engineering. Later, computational linguistic model and machine learning techniques are explored for the two tasks.
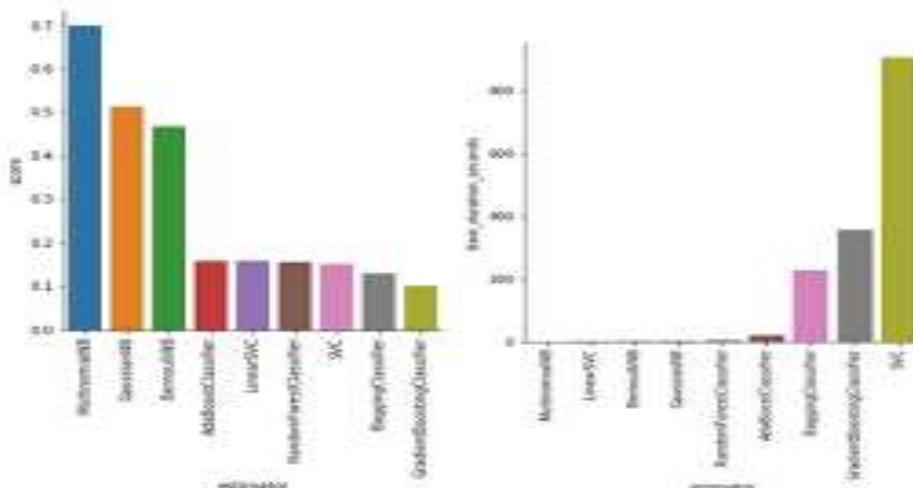
DISCUSSION:
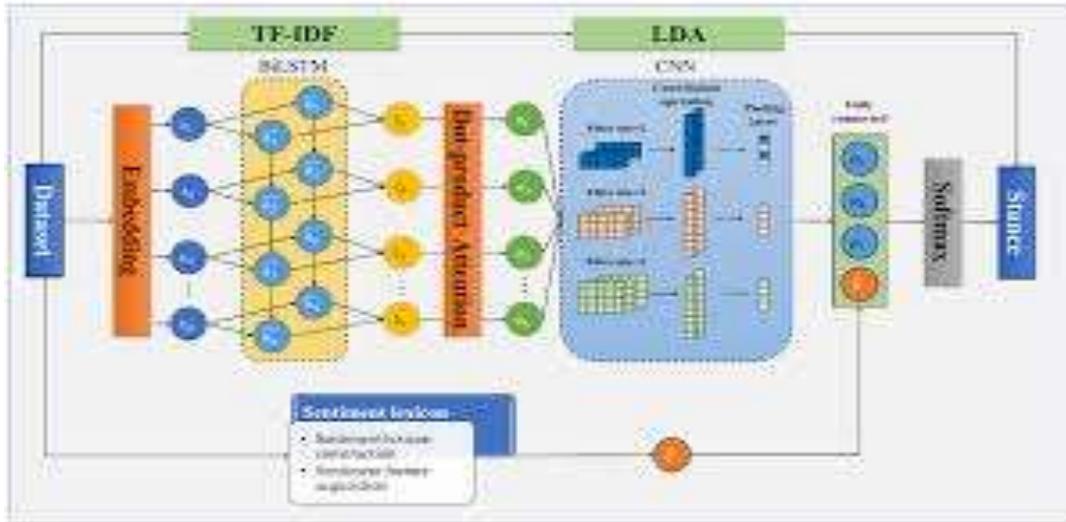
However, it conveys a favor stance to „atheism" target. In, a target based attention model is presented that explicitly models the relationship between target and text content will be helpful for stance detection. Likewise, the methodologies of product aspect mining also have a trend which evolves from the traditional frequency and relation based approaches to the knowledge based topic models and neural based models. Improves the extraction performance, uses two different word embedding to help with aspect and opinion term extraction. While knowledge based approaches have only recently begun to emerge and the knowledge use dare outdated, recently proposed representation techniques like also help mining approaches to enhance the extraction ability. Combining the external knowledge with the power of machine learning will give rise to models which are able to reason with expert instruction, and it could reduce the influence of insufficient data.
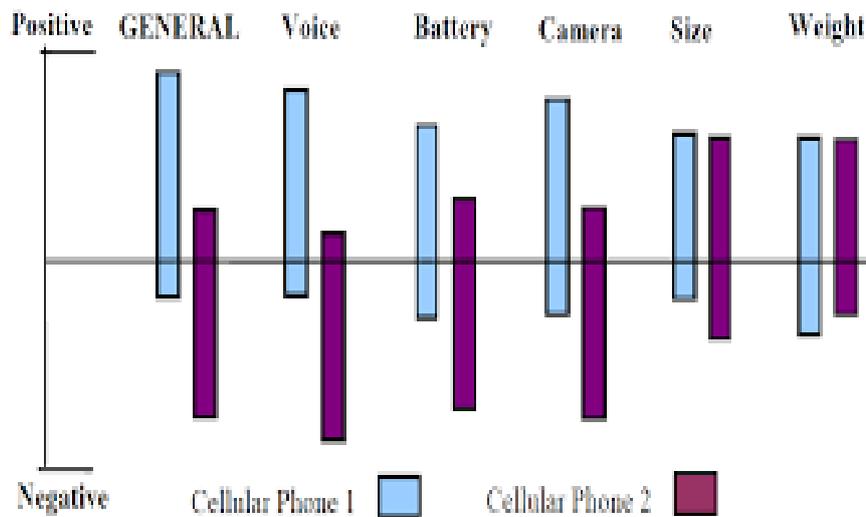
CONCLUSION:

From the overview of the state-of-the-art in stance detection and product aspect mining in this survey, it is clear that the field is transcending its early stage. We could observe another fact that the mining methodologies have evolved through time together with the increasing model complexity. Besides, we want to stress that transparency and standardization is needed in terms of evaluation methodology and datasets in order to draw firm conclusions about the current state-of- the art. For example, in "Religion has destroyed the ability for some to say know", (target: Atheism), it may be assigned an against stance by relation and pattern based approaches due to the negative sentiment of the content.

REFERENCES

[1]. Mr.M.Jayakandan and Dr.A.Chandrabose,"Land Weber Iterative Supervised Classification and Quantized Spiking Network for Crime Detection Emotion Analysis", International Journal of Intelligent Systems and Applications in Engineering, Volume 12, No 21S, (**2024**), Page No: 2219-2224, ISSN NO: 2147-6799 (**G-II-SCOPUS**).

[2]. Mrs.P.Ananthi and Dr.A.Chandrabose, "The Socio-technical Opportunities and Threats of Crowdsensing", The Scientific Temper, Volume: 15 (spl-1), 2024, Page No.:291-297, ISSN Print: 0976-8653 and E-ISSN: 2231-6396 (**G-II_WEB OF SCIENCE**).

[3]. Mr.M.Jayakandan and Dr.A.Chandrabose, "Emotion Analysis Using Iterative Supervised Classification Algorithm for Crime Detection", International Journal of Intelligent Systems and Applications in Engineering, Volume: 12, No 21S, (**2024**), Page No: 2225-2231, ISSN NO: 2147-6799 (**G-II-SCOPUS**).

[4]. Mrs.P.Ananthi and Dr.A.Chandrabose, "Exploring Learning-Assisted Optimization for Mobile Crowd Sensing", The Scientific Temper, Volume 15 (spl-1), 2024, Page No: 283-290. ISSN Print: 0976-8653 and E-ISSN: 2231-6396 (**G-II_WEB OF SCIENCE**).

[5]. Mr.M.Jayakandan and Dr.A.Chandrabose, "An ensemble-based approach for sentiment analysis of covid-19 Twitter data using machine learning and deep learning techniques", The Scientific Temper Volume: 15 (spl-1), **2024**, Page No:114-120, ISSN Print: 0976-8653 and E-ISSN: 2231-6396 (**G-II_WEB OF SCIENCE**).

[6]. Mrs.P.Ananthi and Dr.A.Chandrabose, "Volterra Integral Equation and Logistic Drop-Offloading for Collaborative Mobile Fog Crowd Sensing", International Journal of Intelligent Systems and Applications in Engineering, Volume: 12, No 21S, (2024), Page No: 2186-2192, ISSN NO: 2147-6799 (**G-II-SCOPUS**).

[7]. Mr.M.Jayakandan and Dr.A.Chandrabose, "Machine Learning Classifications for Automatic Sentiment Analysis on Twitter", Indian Journal of Natural Sciences Volume:15, Issue: 86, **October-2024**, Page No: 81858-81867, ISSN: 0976 – 0997 (**G-II_WEB OF SCIENCE**).

[8]. Mrs.P.Ananthi and Dr.A.Chandrabose, "Utilizing Mathematical Modelling and Offloading to Conduct Crowdsensing in A Collaborative Setting", International Journal of Intelligent Systems and Applications in Engineering, Volume: 12, No 21S, (2024), Page no: 2193-2197, ISSN NO: 2147-6799 (**G-II-SCOPUS**).