IOSR Journal of Engineering (IOSR JEN) **www.iosrjen.org**
ISSN (e): 2250-3021, ISSN (p): 2278-8719
PP 01-04

A Survey on Sentimental Analysis on Social Media using Deep Learning

Suhashini Chaurasia[1], Dr. Swati Sherekar[2]
[1]*S.S. Maniar College of Computer and Management, Nagpur,*
[2]*P.G. Department of Computer Science, & Engineering, SantGadge Baba Amravati University, Amravati*

Abstract: *Sentiment analysis is currently treated as a famous topic as well as in the area of research with significant applications in both industry and academia. This paper represents the importance and applications of opinion mining and sentiment analysis in social networks. The paper described the basic concepts, background, methodology, analysis and discussion as different sections of the paper. This paper reflects a survey of deep learning and sentiment analysis in social networks.*
Keywords: *Sentiment analysis, Deep Learning, Social Media*

I. Introduction

Sentiment Analysis is also popularly known as Opinion Mining. It can be defined as a sub discipline of Natural Language Processing (NLP) and Computational Linguistics mainly concerned with the emotion, thought, or a mood expressed by a reader in any document. Here, the former term signifies evaluation of information when extracted and the latter denotes drawing or outing of subjective information from a text corpus or reviews. [1]

Sentiment analysis is an automatic process to determine whether a text segment contains objective or opinionated content, and it can furthermore determine the text's sentiment polarity. The goal of sentiment classification is to automatically determine whether a sentiment polarity is negative, positive or neutral. [2]

Opinion mining techniques can be used for the creation and automated upkeep of review and opinion aggregation websites, in which opinions are continuously gathered from the web and not restricted to just product reviews, but also to broader topics such as political issues and brand perception. [3]

Sentiment is a view, feeling, opinion or assessment of a person for some product, event or service. Sentiment analysis or opinion mining is a challenging text mining and NLP problem for automatic extraction, classification and summarization of sentiments and emotions expressed in online text. Sentiment analysis is replacing traditional and web based surveys conducted by companies for finding public opinion about entities like products and services. Sentiment Analysis also assists individuals and organizations interested in knowing what other people comment about a particular product, service topic, issue and event to find an optimal choice for which they are looking for. [4]

Knowing the overall sentiment inclination towards a topic may prove extremely useful in certain cases. For instance, a technological company would like to know what their customers think about the latest product, in order to receive helpful feedback that will be utilized in the production of the next device. Therefore, it is obvious that an inclusive sentiment analysis for a time period after the release of a new product is needed. Moreover, user-generated content that captures sentiment information has proved to be valuable among many internet applications and information systems, such as search engines or recommendation systems. [5]

II. Backgroud

Due to the huge growth of social media on the web, opinions extracted in these media are used by individuals and organizations for decision making. Each site contains a large amount of opinioned text which makes it challenging for the user to read and extract information. [6]

Besides the open problem of determining which documents are topically relevant to an opinion oriented query, an additional challenge is face in new setting is simultaneously or subsequently determining which documents or portions of documents contain review-like or opinionated material. [7]

NLP is used in the industry surrounding sentiment analysis has been flourished due to the proliferation of commercial applications. It offers many challenging research problems, which had never been studied before. Feature extraction identifies those product aspects which are being commented by customers, sentiment prediction identifies the text containing sentiment or opinion by deciding sentiment polarity as positive, negative or neutral and finally summarization module aggregates the results obtained from previous two steps. [8]

There is a problem when negative comments on social media are posted. Without forensic evidence, many offenders in complex criminal cases are still free and may repeat offending against the laws and harm the others.

A problem of children/ person being exposed to pornographic web sites on the internet has led to their safety issues. To prevent the children/person from these inappropriate materials, an effective web filtering system is essential. [9]

The unstructured form of data from the social media is needed to be analyzed and well-structured. The challenge for sentiment analysis is lack of sufficient labeled data in the field of NLP[10]. And to solve the problem the research is proposed. Sentiment classification, cross lingual problems, textual and visual analysis and product review analysis are also proposed in this research.

III. Methodology

The web contains both a huge amount of information in structured and unstructured texts. Analyzingunstructured texts is of great importance and perhaps even more important than extracting structured data because of the sheer volume of valuable information of almost any imaginable types contained in them.

Businesses always want to find public or consumer opinions on their products and services. Potential customers also want to know the opinions of existing users before they use a service or purchase a product. Moreover, opinion mining also known as sentiment analysis, can also provide valuable information for placing advertisements in web pages. If in a page people express positive opinions or sentiments on a product, it may be a good idea to place an ad of the product. However, if people express negative opinions about the product, it is probably not wise to place an ad of the product. A better idea may be to place an ad of a competitor's product.

Mining opinions on the web is not only technically challenging because of the need for NLP, but also very useful in practice.

The web has dramatically changed the way that people express their opinions. They can now post reviews of products at merchant sites and express their views on almost anything in Internet forums, discussion groups, blogs, Twitter, Facebook, Foursquare, Instagram, etc. This online word-of-mouth behavior represents new and measurable sources of information with many practical applications. Because of these features new techniques are needed and social media mining has become popular.

Social media Mining is the process of representing, analyzing, and extracting actionable patterns from social media data. Social Media Mining, introduces basic concepts and principal algorithms suitable for investigating massive social media data. It discusses theories and methodologies from different disciplines such as computer science, data mining, machine learning, social network analysis, network science, sociology, ethnography, statistics, optimization, and mathematics. It encompasses the tools to formally represent, measure, model, and mine meaningful patterns from large-scale social media data. Twitter and Facebook are two of the todays the most known applications:

Twitter is a rich source of social data that is a great starting point for social web mining because of its inherent openness for public consumption, clean and well-documented API, rich developer tooling, and broad appeal to users from every walk of life. Twitter data is particularly interesting because tweets happen at the "speed of thought" and are available for consumption as they happen in near real time, represent the broadest cross-section of society Global Journal of Emerging Trends in e-Business, Marketing and Consumer Psychology (GJETeMCP) at an international level, and are so inherently multifaceted. Tweets and Twitter's "following" mechanism link people in a variety of ways, ranging from short but often meaningful conversational dialogues to interest graphs that connect people and the things they care about Facebook is arguably the heart of the social web and is somewhat of an all-in-one wonder, given that more than half of its 1 billion users are active each day updating statuses, posting photos, exchanging messages, chatting in real time, checking in to physical locales, playing games, shopping, and just about anything else you can imagine. From a social web mining standpoint, the wealth of data that Facebook stores about individuals, groups, and products is quite exciting, because Facebook's clean API presents incredible opportunities to synthesize it into information (the world's most precious commodity), and glean valuable insights.

Companies have millions of tweets about their brands, thousands of Facebook "likes", hundreds of thousands of check-ins on Foursquare. Pinterest and Instagram are adding even more to social media data deluge. Companies can manage their brand awareness and brand loyalty through social media. Positive, negative and neutral comments of people are core of social media analysis. [11]

Opinion features such as reviews on a particular product are typically domain-specific. The feature appears frequently in the given review domain, and which are outside the domain is domain-independent corpus about product. Domain-specific opinion features are mentioned more frequently in the domain corpus of reviews, as compared to a domain-independent corpus. A domain-dependent review corpus and a domain-independent corpus is observed. Fig. 1 shows that first extract a list of candidate features from the review corpus

by defining manually syntactic rules. Each extracted candidate feature, will estimate its IDR, which represents the statistical association of the candidate to the given domain corpus, and extrinsic-domain relevance, will reflects the statistical relevance of the candidate to the domain-independent corpus. Only candidates with IDR scores more exceeding a predefined intrinsic relevance threshold and EDR scores less than another extrinsic relevance threshold are extracted as valid opinion features. It identifies opinion features that are domain-specific and at the same time domain-independent corpus are removed and ignored. [12]

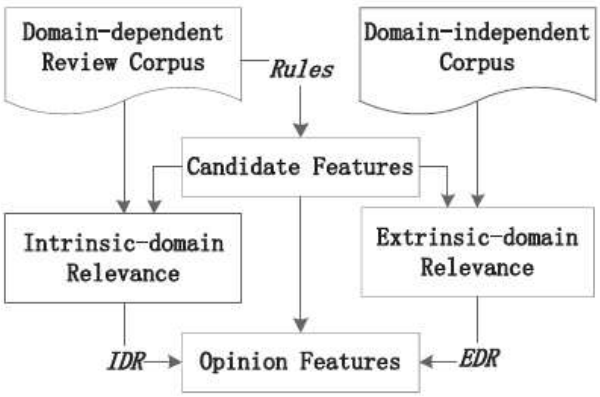

Fig. 1: *IEDR Workflow*

Automatic Summarization

Automatic Summarization is the process of reducing a text document with the help of a computer program in order to create a summary that retains the most significant points of the original document. Technologies that can make a coherent summary take into account variables such as length, writing style and syntax. The main notion of summarization is to find a representative subset of the data, which contains the information of the entire set.

Generally, there are two approaches to automatic summarization: Extraction and Abstraction. Extraction refers to selecting a subset of existing words, phrases, or sentences in the original text to form the summary. In contrast, abstraction builds an internal semantic representation and then use natural language generation techniques to create a summary that is closer to what a human might generate. Automatic Summarization system takes three basic steps namely, Analysis, Transformation and Realization.

In analysis, a concise and fluent summary of the most significant information is produced in the input. It requires the capability to reorganize, modify and merge information expressed in different sentences in the input. Transformation is an ordered text is generated by manipulating the internal representation post analysis in Auto Summarization. An analyzed summary text is generated using scores of transformation in the Realization phase. The process of Auto Summarization is depicted in Fig. 2. [13]

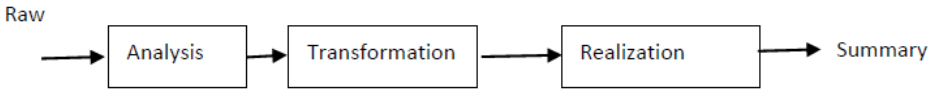

Fig. 2. Process of Auto Summarization

IV. Anaysis and description

The methodology is being applied on neurons which are trained. To train on set of aspects extracted for other set of training cases of customer reviews. First train the neural network then fine tune with backdrop. After this step it has to be tested on another set of separate customer reviews. One review will be picked query. Other reviews will be ranked using consine of angles between codes. This step has to be repeated for other test reviews as query as well. For performance measurement, the number of reviews retrieved has to be plotted against the proportion that is in the same hand labeled class as the query document. [14]

V. Conclusion

Sentiment classification of features of products is useful for shopping websites where it is possible to give more detailed information about the product from consumer's point of view. Showing user opinions about special features of products is great beneficiary to both online retailer and buyer of products. In this paper it is observed that sentiment analysis or opinion mining plays important role while making a decision towards a particular product or a service. But it is very important to consider certain quality measures like helpfulness, usefulness and utility while analyzing each review. In the literature survey there are many sophisticated methods explained which defines the sentiment analysis with respect to different aspects. In future, more research work is needed to improving the performance measures further. Sentiment analysis or opinion mining can be applied for any new applications which follow data mining rules. Although the techniques and algorithms used for sentiment analysis are advancing fast and giving high quality results, lot of problems in this field of study remainunresolved andalso it is hard to find the fake review by reading. Sometimes fake reviews also seen as good quality review and it was modified like no one can identify their actual intension. For further improvement, we can increase the database of our reviews search engine; bigger the search database will increase the reliability of the system. Phrase extraction patterns are crucial to implement as there is possibility of useless phrases. [15]

References

[1]. Archana Sonagi1, Deipali Gore, "Sentiment analysis and challenges involved A survey", International Journal of Science and Research (IJSR),vol. 5, pp 1928-1932, 2015

[2]. Zhao Jianquang , Gui Xiaolin, Zhang Xuejun, " Deep Learning for Sentiment Analysis A Survey", vol. 6,IEEE Access, pp 23253-23260, 2017

[3]. Soujanya Poria , Erik Cambria, Alexander Gelbukh, "Aspect extraction for opinion mining with a deep convolutional neural network", Elsevier, 2016

[4]. Muhammad Zubair Asghar, Aurangzeb Khan, Shakeel Ahmad, Fazal Masud Kundi, "A review of feature extraction in sentiment analysis", Journal of Basic and Applied Scientific Research, 2014

[5]. Andreas Kanavos, Nikolaos Nodarakis, Spyros Sioutas, Athanasios Tsakalidis,Dimitrios Tsolis, Giannis Tzimas , "Large sca, le implementations for twitter sentiment classification", MDPI algorithms, pp- 1-22,2017

[6]. Savitha Mathapati, S H Manjula, Venugopal K R, "Sentiment Analysis and Opinion Mining from Social Media: A Review", Global Journal of Computer Science and Technology: C Software and Data Engineering , ISSN: 0975-4350, vol 16, issue 5,pp - 77-90,2016

[7]. Bo Pang, Lillian Lee , "Opinion mining and sentiment analysis", Foundations and Trends in Information Retrieval, vol-2, pp 1-90, 2008

[8]. Bing Liu , "Sentiment Analysis and Opinion Mining", Ph.d. thesis published by Morgan & Claypool Publishers, 2012.

[9]. Megha S Anchatgeri, Thejaswini, "The Facebook Analysis using Natural Language Processing (NLP)", International Research Journal of Engineering and Technology (IRJET), p-ISSN: 2395-0072, Vol. 05, Issue: 05,pp 2518-2519, 2018

[10]. Qurat Tul Ain , Mubashir Ali , Amna Riaz , Amna Noureen , Muhammad Kamran , Babar Hayat, A. Rehman. "Sentiment Analysis Using Deep Learning Techniques: A Review", International Journal of Advanced Computer Science and Applications, Vol.8, No.6, pp 424- 433, 2017

[11]. Umman Tugba Gürsoy, "Social Media Mining and Sentiment Analysis for Brand Management", Global Journal of Emerging Trends in e-Business, Marketing and Consumer Psychology (GJETeMCP), ISSN: 2311-3170, Vol 3, Issue 1, pp 497-511, 2017

[12]. ProcessingSwati N. Manke, Nitin Shivale, "A Review on: Opinion Mining and Sentiment Analysis based on Natural Language Processing", International Journal of Computer Applications, Vol. 109, No. 4, pp 29-32, January 2015

[13]. Charu Virmani, Dr. Anuradha Pillai, Dr. Dimple Juneja, "Extracting Information from Social Network using NLP", International Journal of Computational Intelligence Research, ISSN 0973-1873, Vol. 13, No. 4, pp. 621-630, 2017

[14]. M. Lal, A. Jain, M. Avatade, "Sentiment Analysis on Customer Reviews using Deep Learning", International Journal of Computer Sciences and Engineering Open Access, Vol.-6, Issue-7, E-ISSN 2347-2693, pp 1023-1024, July 2018

[15]. Supriya Soni,Shivendra Dubey, Rakesh Tiwari,Mukesh Dixit, "Feature Based Sentiment Analysisof PrpductReviewsusing Deep Learning Methods", International Journal of Advanced Technology & Engineering Research (IJATER), ISSN No 2250-3536 Vol. 8, Issue 3, pp 16-24, May 2018