Sentiment Analysis: From Opinion Mining to Human-Agent Interaction

Adlene Ebenezer. P, Srijan Saumya, Anand Kumar, Sumit Saikia, Abhishek Raj

Abstract - This framework proposes the strategies for sentiment analysis considered for the eye to eye and human-specialist associations. We have thought about the techniques for our proposed framework and its convention of assessment regarding the current sentiment analysis strategies and have demonstrated how our proposed framework has coordinated the human-operator communication angles. Finally, we proposed a top to bottom execution analysis of the outcomes gotten by the assessment, with the talk on the diverse troubles and the significant difficulties of sentiment analysis in our proposed human-operator connections with the usage and working of the map reduce.

Keywords – ECA, Human agent interaction, Map Reduce, Opinion-mining, Sentiment-Analysis.

I. INTRODUCTION

The field of opinion mining and sentiment analysis has been an unexpected increment of enthusiasm with the quick development of accessible content data containing faultfinders, opinions, proposals on the web (motion picture surveys, gathering discusses, tweets). The test is the advancement of opinion identification techniques relying on these new assets. The differing data and the modern applications utilizing sentiment analysis offers ascend to different issues that are not all that completely tended to by the current existing framework. Another maturing research field is the improvement of ECAs (encapsulates conversational operators), virtual characters associating with individuals. The ECAs are occupied with different applications: - they can play out the job of collaborator, for example, characters present on the deal sites, or of a coach in some genuine diversions. From one perspective ECA needs to incorporate the human enthusiastic conduct just as social frames of mind, then again ECA needs to indicate them with significance.

Sentiment analysis regarding the matter's verbal substance is then basic for the ECA to decide the client's feeling and mentality and to adjust its conduct as indicated by the outcome . The setting of the paper shows improvement of a sentiment analysis module with its combination in an ECA stage managing multimodal social and enthusiastic collaborations. The verbal analysis of the substance will be associated to the acoustic analysis and video analysis of client's social and enthusiastic conduct. The last target is in this manner to figure out which response an ECA should appear as indicated by the client's financial conduct. There are two inquiries are imperative to examine in this unique circumstance: What are the sentiment – related wonders that are pertinent so as to improve the specialist multimodal observation space? What are the current sentiment calculation ways and how they are to be implanted in an ECA stage? In this paper, discourse about the setting identified with the current framework , its disservices proposed framework and its points of interest design are identified with opinion mining and module portrayal and future works and ends. The objective is to furnish the sentiment analysis with human collaboration to discover the financial conduct of the operators.

II. LITERATURE SURVEY

We had done literature survey using the following papers for the description of various resources of sentiment analysis with human interactions. This is the basic paper of how the different methods regarding the sentimental analysis works, different types of datasets e.g. Facebook, Twitter, MySpace, YouTube applications of opinion mining and ECA in different areas their comparision , algorithms of machine learning and their general process has been learned.[1] In this paper we got to know about the perspective of sentiment analysis with human interaction: opinion mining and Embodied Conversational Agent and the different phenomenons.[2] We came across the different tools and techniques of opinion mining like red opal, open NLP etc. with various datasets (blogs, reviews sites etc.)and how to use them for sentiment analysis.[3] In this we analyzed the different approaches of multimodal sentiment analysis like text speech videos vlogs facial expression etc. with opinion structures and approaches.[4] Demonstration of sentimental analysis from ECA points of view ,computational method for sentimental analysis and the architecture of GRETA platform with human agent is discussed and learned.
[5] The Liking dimension in ECA and detection system and evaluational protocols comparison in likes and dislikes also the issues and error related to SA is analyzed and studied.[6]

Retrieval Number D6234048419/19©BEIESP
Journal Website: www.ijeat.org

1450

Published By:
Blue Eyes Intelligence Engineering
& Sciences Publication (BEIESP)
© Copyright: All rights reserved.

III. EXISTING SYSTEM

The human - agent connection and opinion mining networks are at present demonstrating sentiment analysis from different perspectives that comprises of on one side sentiment-related unique techniques and instruments and portrayals identified with computational , and on the opposite side, disparate location and systems of discourse the board. The framework distinguishes just as looks at the growing degree for between disciplinary errand that may improve singular advancement. Sentiment or opinion recognition techniques that are utilized in cooperation among human and operator are very rare and, at whatever point they are locked in, they are like the ones utilized in opinion mining and in this manner not created for socio-full of feeling communications (time - imperative of collaboration and sentiment analysis as an information and association procedures as yield). To fortify what we have expressed, we have appeared relative cutting edge strategy that filters the sentiment identification techniques utilized in the two networks and the sentiment-related marvels and present a diagram of the destinations of socio-emotional human-specialist methodologies.

A. Existing system Disadvantages

- Cannot calculate the triumph of campaigns of marketing.
- Cannot determine the user-experience with specific products or the brands values or images.
- Cannot determine stock's price stability or fluctuation.
- Cannot deduce the user's attitude or emotions.
- Highly difficult to use.

IV. PROPOSED SYSTEM

The proposed framework presents various probabilities for symbiotic increase, featuring some examination tracks and considering the open finishes and expectations. To show the practicality of the typical principles considered we also handle them from a specific viewpoint by their application to the instance of the incorporated conversational agents (ECA) for human-agent interactions in two unique use cases: first job related meetings and after that conversations with visitors to the historical center. In this proposed application, identification and discussion of the growing possibilities for between disciplinary errand that may upgrade singular progress. Sentiment or Opinion detection which is utilized in interaction of human-agent are positively rare and, when employed they are not excessively various from those which are occupied with opinion mining and consequently are not developed for socio-full of feeling interactions. The goal of this paper is to produce a craftsmanship related state on sentiment-analysis from a viewpoint of the opinion mining and communities related conversational agent, recognizing the most allowable upgrades of both communities and discussing the open research inquiries and opportunities. Exploring territories of specific examinations, as or expect to develop applications including social relationships among the ECA and the human client. For modeling relationships, these examinations use huge dimensions. One of the most generally utilized is the enjoying dimension.

A. Proposed System Advantages

- To decide the accomplishment of showcasing campaigns.
- To decide the client experience with specific products or the brands picture.
- To decide stock's price strength or fluctuations.
- Can find the frame of mind or emotions of clients.

The Heider's Balance Theory that relates with "the way in which associations among persons involving couple of impersonal substances are judiciously felt by the person. It surveys the correlations connecting a person P, who is the prime-focus of the analysis, one more person O and X - an element, which can be anything like methods, occasions or even a physical things or objects. Every one of these relations can be modeled with the help of a triangle with vertices – O , X and P. On the off chance that the client and the agent experience disparate inclination towards the impersonal substance (X), at that point client's enjoying with the agent is either pessimistic or positive. After the description of Heider's Balance Theory we have shown the engineering of framework with implementation.
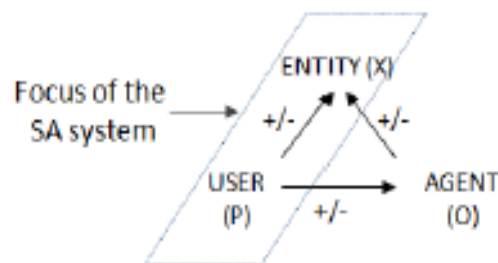

Fig.4.1. Heider's Balance Theory

V. SYSTEM ARCHITECTURE DESIGN

The interaction module will first collect all the data from the user. The data file is then stored in an SQL database. The processing mechanism on this file takes place in Hadoop system (single cluster) and the output generated will be analyses on the basis of the model trained (opinion mining).

Data will have various datagrams like: appreciations, emotions, opinions, social stances, sentiments, affects, moods and so on. Then the model will use various methods for detecting the sentiment-related phenomena in opinion mining. After the system analyses the data dialog management and socio-emotional interaction for ECAs.
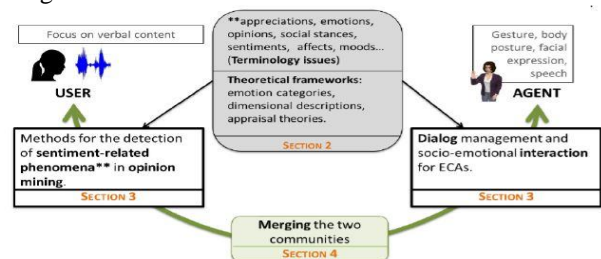

Fig.5.1. Architectural Overview

In order to process data, opinion data should be present. Web scraping which is defined as the process for gleaning information for the Internet is performed on various e-commerce platform for extracting reviews to act as training data.

Once training data is acquired, next algorithmic steps could be performed on the reviews text file. The review text file is in the following format, which eases processing in Hadoop Architecture.
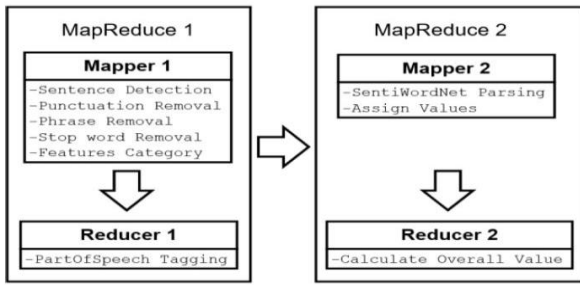

Fig.5.2. Map Reduction Algorithm

The proposed mechanism for feature based opinion summarization takes place in the Hadoop system and its architectural overview.
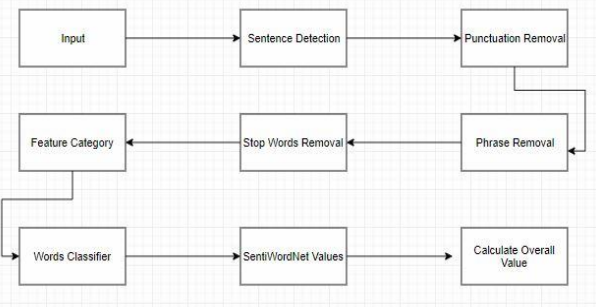
VI. IMPLEMENTATION

Fig.6.1. Workflow Diagram for Map Reduction

A. Sentence Detection

A review is not necessarily always written in a single line. Most of the time, it is in a form of paragraphs. Sentence Detection allows detection and segregation of sentences from the paragraphs which can then be processed.

B. Punctuation Removal

Punctuations and special characters are to be removed from the sentences such that only alphabets and number are left in the sentences. The sentences are also entirely converted into lower cases.

C. Phrase Removal

Phrases such as „could have been", "hope it will be" are removed and replaced by a negation word.

D. Stop Words Removal

Stop words are considered as meaningless words which are filtered out to reduce the processing time. This list consists of the preposition, conjunctions, articles, etc.

E. Feature Category

It represents searching for features related words in the sentence and then classifying in the same feature cluster. For example, the review data set is parsed for keywords/ feature (such as display: display, screen, gorilla glass, resolution, color, pixels) which are generated by finding frequent item set through Apriori Algorithm. After finding such word in the sentence, the sentence is classified in the cluster of only those features.

F. Parts-Of-Speech(POS) Tagging

The POS tagging model is applied to the sentences thereby providing part of speech of each word in the sentences.
Apache's Open NLP has been used to perform Sentence Detection and POS tagging. POS tagging model uses Maximum Entropy Model for analyzing information gain on training data and provides parts of speech tags to new sentences. The reason for removing opinion changing phrases before stop words can be understood by the given example. In this case , if the stop words are removed directly then the opinion of sentence changes. In the given example, the opinion is negative as they expect memory to be better, but if stop words are removed then the remaining words are: memory phone better. This, when processed for sentiment gives an influence on positive sentiment which is contradictory. Hence, using phrase removal before stop word removal acts as a solution, so that a negative word can be substituted in such cases and the remaining words left after phrase removal and stop word removal are memory phone not better which gives a sense of negative influence in the sentence. So the first Map-Reduce provides an output of POS tagged sentences placed in the feature cluster to the second Map-Reduce.

VII. PERFORMANCE-ANALYSIS

The system developed aims to achieve an efficient mechanism for summarizing the opinion posted by users to help other potential customer. It enable many ecommerce websites in the need of time to substitute there up vote system for surfacing helpful reviews with the proposed system which doesn't involve manual intervention in rating process. The system provides information regarding that. We assert that usage of techniques and mechanism provided by
Hadoop System such as Key – Value pair and Map Reduce significantly reduces the time complexity of system with such intensive processing.

VIII. RESULT AND DISCUSSION

We have obtained the results by comparing the two constituents of sentiment analysis method of opinion mining and ECA based applications with the consideration of the different module description approaches. We have shown the detection of emotions ,polarity, negative and positive categories and some other sentimental categories in each applications with different level module approaches. In this we have also depicted how the two different applications work in three different levels of module approaches that is dimensional approach module, discrete approach module and appraisal approach.The dimensional approaches are frequently used in module from psychological point of view while the appraisal approach mainly focus on process modeling for evaluation of events.

Sentiment Analysis: From Opinion Mining to Human-Agent Interaction

We have also done an in-depth analysis of the challenges in our human-agent proposed interaction framework. We have also investigated the contradiction between the human annotations and the framework output so as to distinguish tracks for the improvement of our framework. The principal trouble is identified with the processing of the spontaneous speech .The disruption of the linguistic structure of the speech turn and subsequently blocks both the annotation process and the detection framework. One of the solution is to coordinate the framework ready to automatically name disfluencies.

Second challenge is the shortage of context or words provided by some APIs. One of the solution to this is to involve the design of difficult disciplines that will take into consideration of a huge sets interaction context.

Table I: Opinion Mining vs. ECA applications

Opinion Mining Applications	ECA Applications
• Detection of the polarity (eg.positive vs negative:reviews, attitude towards brands,comparitives and preferences.	• Negative emotions detection are used in optimising ASR and for the interaction management
• Evaluation of the value of the emotions viewed in comment and contexts:used in evaluation of data like web contents, blogs,movies reviews etc.	• Detection of negative emotions categories
• Used in evaluation of categories such as positive and negative expression	• Used in ignoring user's frustation in driving situations like driving or for tutoring system or for an infant controversial machine game.
• Detection of emotion/affect in texts (dimensional approach mixed with discrete approach OCC categories)	• Various emotions detection are used according to the working of dialog systems.
• Other sentimental categories(such as agreement vs disagreement)is being analysed	• Avatar visualizations in textual affects.
• Used in detection of the user's target sentiment such as reviews,call-centers,for example Whites' appraisal model.	• Modelling the models of behavior and user decision making system along with the context of the interaction using various types of models for example dynamic appraisal EMA model or BICA.
• Presentation of effective study and learning of alphabetical resources and textual affect sensing.	

IX. FUTURE WORKS

In future work, these techniques and opinion mining process can be improved by taking into consideration the usage of words used by people. Features can also be clubbed together according to the score as good, neutral, and bad. Spam reviews can be detected and removed from the list to increase the overall efficiency (Algorithm can be developed to check whether features are present in the reviews posted or not).

X. CONCLUSION

The paper comprises the leads of multiple research for the combination of sentiment analysis with the eye to eye human-agent interactions. The difficulties obviously shows

that to achieve this goal are cross-disciplinary and numerous. To start with, our premise is on psycho-phonetic methods, so as to decide the phenomena of sentiment-related that are fascinating to be accepted by an ECA. We've also proposed the existing system of opinion mining and their disadvantages in contrast with our proposed system advantages. Then we have showed the architectural design of the ECA proposed model. Finally the approaches and the map reduce with their implementation of working is showed and finally we have demonstrated the performance analysis of our proposed system with map reduce with some key points.

REFERENCES

1. Mohammed Ali Al-Garadi , Shahid Shayaa, Shamshul Bahri, Ainin Sulaiman,Phoong Seuk, Yeong Wai Chung, Noor Ismawati Jaafar , Arsalan Zahid Piprani,and "Sentiment Analysis of Big Data: Methods,Applications Open Challenges" July 30, 2018.
2. Chlo_e Clavel and Zoraida Callejas "Sentiment Analysis : From Opinion Mining to Human-Agent Interaction" Ieee Transactions On Affective Computing, Vol. 6, No. X, Xxxxx 2015.
3. M. K. Prakash, G.Dinesh " A Perspective Survey on Opinion Mining Tools and Techniques" ISSN XXXX XXXX © 2018 IJESC
4. Mohammad Soleymani, David Garcia, Brendan Jou, Bj¨orn Schuller, Shih-Fu Chang, Maja Pantic, "A Survey of Multimodal Sentiment Analysis,Image and Vision Computing (2017)", doi: 10.1016/j.imavis.2017.08.003
5. Chlo´e Clavel1, Catherine Pelachaud2, Magalie Ochs2 "User's sentiment analysis in face-to-face human-agent interactions – prospects".
6. Caroline Langlet, Chlo´e Clavel Institut Mines-T´el´ecom ParisTech "Adapting sentiment analysis to face -to-face human-agent interactions: from the detection to the evaluation issues" 2015 International Conference on Affective Computing and Intelligent Interaction (ACII).
7. Nidhi R. Sharma , Prof. Vidya D. Chitre "Opinion Mining, Analysis and its Challenges", International Journal of Innovations & Advancement in Computer Science.
8. Y. Choi, C. Cardie , E. Breck , "Identifying expressions of opinion in context," in International Joint Conference On Artifical Intelligence, Sangal S., Mehta H., and Bagga R. K., Eds., San Francisco, CA, 2007, pp. 2683–2688, Morgan KoffMann Publishers.
9. C. Clavel , C. Langlet "Improving social relationships in face-to-face human-agent interactions: when the agent wants to know user's likes and dislikes," in The Association for Computer Linguistics, Beijin, China, July 2015, to appear.
10. Jayasree, Raisa Varghese. A Survey on Sentiment Analysis and Opinion Mining, International Journal of Research in Engineering and Technology (IJRET), Vol 2 Issue 11 Nov 2013.

AUTHORS PROFILE

Adlene Ebenezer.P Computer Science & Engineering, SRM Institute of Science and Technology, Chennai adlenepackiadoss@gmail.com

Srijan Saumya Computer Science & Engineering, SRM Institute of Science and Technology, Chennai. srijansaumya@gmail.com

Anand Kumar Computer Science & Engineering, SRM Institute of Science and Technology, Chennai kr.anand.srm@gmail.com

Sumit Saikia Computer Science & Engineering, SRM Institute of Science and Technology, Chennai. sumit25saikiaactive@gmail.com

Abhishek Raj Computer Science & Engineering, SRM Institute of Science and Technology, Chennai. abhishek.raj0001@gmail.com