2020 IEEE International Symposium on Sustainable Energy, Signal Processing and Cyber Security (iSSSC) | 978-1-7281-8880-5/20/$31.00 ©2020 IEEE | DOI: 10.1109/iSSSC50941.2020.9358825

An Artificial Intelligence approach to Social Networks agent task scheduling analysis in map-reduce for Sentiment Opinion Analysis

Dr. YEGIREDDI RAMESH
Associate Professor, Department of Computer Science and Engineering, Aditya Institute of Technology and Management, Tekkali, Srikakulam, Jawaharlal Nehru Technological University Kakinada, Andhra Pradesh, India.
rameshyegireddi@gmail.com

BOSUBABU SAMBANA
*Assistant Professor and HOD, Department of Computer Science and Engineering, Viswanadha Institute of Technology and Management, Visakhapatnam, Jawaharlal Nehru Technological University Kakinada, Andhra Pradesh, India. **
bosukalam@gmail.com

Dr. MOLLI SRINIVASARAO
Professor, Department of Computer Science and Engineering, Dadi Institute of Engineering & Technology, Anakapalle, Visakhapatnam, Jawaharlal Nehru Technological University Kakinada, Andhra Pradesh, India.
sri_nivas_molli@rediffmail.com

Abstract — Now a day's everyone widely used social networks; it is one part of our daily life activity. In this way, the user can upload and view their profiles text messages, images and video etc. In the way, user can communicate either individual or group in their existing regions and post updated information to circulate all other followers or guest users like working the same as social networks sites. This research work, we propose a novel algorithm to analyze existing resources and predict sentiment analysis and prediction along with this user expectation. These results will be useful in Applied Economics, Expert Systems, Human Brain Analysis, Voting, Stock markets, Live updates on every field, unknown object identifications Etc.

Keywords — AI, Sentiment Analysis, Big Data, Social Networks, Map Reduce.

I. INTRODUCTION

Artificial Neural Network (ANN) is one of the significant parts of AI, comprising of hugely interrelated nonlinear memory less prepare mechanism known as neurons or hubs. Instead of the customary demonstrating procedures, ANN is an information-driven, self-versatile, discovery strategy, which gains from models. Intelligent machines frequently analyze existing resources and identified necessary changes have occurred, then predicted the system would resolve the problem and guide how to sequencing sufficient negative information, apart from of whether the initial connections are positive or response their existing resources. It is in the nonlinear idea of this present reality occasions.

Therefore, ANN has discovered use in numerous fields, including natural sciences and Technology [1, 2]. These days, high volumes of valuable information can be expertly gathered and created from various sources, for example, interpersonal organizations. Interpersonal organizations are commonly made of social substances that are connected by some particular sorts of interdependency. These days, high volumes of valuable information can be expertly gathered and created from various sources, for example, interpersonal organizations. Interpersonal organizations are commonly made of social substances that are connected by some particular sorts of interdependency (e.g., companionship, necessary intrigue).

In the present situation, long-range informal communication has left no segment of the general public immaculate. A large number of profiles are made each day on these sites. It has quickly turned into the best stage for associating with individuals as well as to frame and look after connections. With their enormous effort, these sites have turned into the most helpful approach to share media just as essential data. Nowadays, Facebook, Twitter, LinkedIn, Pinterest, and Instagram are the most prominent long range informal communication sites the whole way across the globe [3].These long-range informal communication sites create a gigantic measure of information like posts, remarks, and likes. As opposed to keeping this information in the information storage, this information can be dissected to enhance the person to person communication posts. Opinion analysis and examination is one such method that can be utilized for this reason [4]. It refers to the analysis of the content to recognize the feelings passed on through the content by the user. Sentiment analysis or feeling investigation should be possible utilizing multiple approaches, that is, the lexicon-based methodology and Machine Learning (ML) approach. This paper centers on the linear based way to deal with sentiment opinion analysis.

Individuals express their conclusion on a person to person communication blogs or sites through preferences expressions through emogies, comments or tweets etc. Clients may either like or comments on a post or do both. The critical region of the focal point of this investigate paper is applying users feeling inquiry on the comments, timelines and millions of users share their opinion in various social network platforms (Instagram, WhatsApp, Twitter, Hike, Facebook and many more) and making it an essential stage for following and breaking immediately store analyzed records making it an essential stage for breaking down open analysis, after examination data will categorization into various fields and that fields approach necessary decision making user's sentiment feelings. In this way, it has stood out in both the academic scholar's community and business [5].

In recent years, a few information mining calculations and systems have been proposed. A considerable lot of them are material to mine interpersonal organizations (e.g., location of networks, an expectation of networks [6, 7], just as the disclosure of 'following' designs and differing social elements). What is more, there are likewise ongoing works for companion mining and companion suggestion. Instances of the previous incorporate the mining of powerful companions and reliable companions. Instances of the last incorporate a proliferation structure that together targets client interests and predicts fellowships and a semantic-based companion suggestion framework called Facebook.

In this paper, we propose an elective information expository arrangement which incorporates the Map-Reduce model into successive example mining for the disclosure of mainstream gatherings of companions and hence suggestions of these companion gatherings. Enormous PFM finds prevalent companions and prescribes them to different clients. Such an answer is a non-inconsequential reconciliation of MapReduce into regular example mining [8]. The following segment exhibits our new MapReduce based information logical answer for continuous pattern mining.

II. BACKGROUND

A wide range of procedures has been utilized to group the given content as per extremity related. The most across the board methods are Machine Learning and Lexicon Based Approach. It is the investigation of calculations with information. Such calculations work by structure a model from information traits and along these lines utilizing that to make future forecasts. Models are: web indexes, grouping an email as spam or not [1]. AI systems utilize preparing and testing sets. Based on the preparation set, a model is assembled first and afterwards with the utilization of a testing set, and model testing is finished. Thus, a model is manufactured that can be utilized for any obscure example with some info ascribe to foresee the future class of that example.

Various Machine Learning (ML) procedures like Maximum Entropy (ME), Naïve Bayes are accustomed to arranging the given content. A little thought regarding the apparatuses used to follow the extremity or assessment of client created substance was likewise given. It incorporated the Review guide instrument, Web source, and Red Opal and Opinion observation [8].

Hemalatha et al. exhibited the work on dissecting the tweets from long range informal communication sites utilizing simple based calculation and most extreme entropy order. They acquainted an assumption investigation apparatus with distinguishing the positive, negative and nonpartisan tweets from the data assets, in this way given an approach to catch the innovation trend later on.

Neethu et al. proposed the work on sentence-level supposition examination by making a component vector for the twitter posts. The tweets were characterized by positive and negative classes, utilizing various classifiers, and the last opinion was inferred. In Lexicon Based strategy generally, the extremity of a total sentence or expression is determined by the whole of polarities of all the individual words or expressions.

Palanisamy et al. proposed a strategy for assumption investigation of twitter information utilizing Lexicon based methodology. They gave various systems to distinguish and remove assumptions from hash tags, emogies and gave a technique to change over non-syntactic words into linguistic words. The Lexicon was worked from the Sentiment of scientific categorization. In the preprocessing stage, the means performed were stemming, emogies location and standardization, misrepresented word shortening and hash tag recognition. The general rationale was to recognize the supposition related to the post, as emogies, comments expressions or tweets.

In this manner, realizing the careful insights related to each post from the proposed method will help in numerous essential leadership for more prominent firms. This strategy can be utilized as huge known firms or MNC's creating items could not want anything more than to know whether their clients prefer the item or not. during the proposed strategy, if they know the measurements [9], what is the definite number of preferences then that dreary work would be decreased as from the insights they will get a thought concerning the fame or need of their item. In this way, in necessary leadership, the proposed technique appears to be proper.

III. RELATED WORK

A. Big Data

As indicated by IBM reports [4] consistently "2.5 quintillion bytes of information" is made [10]. These figures are expanding every year. This is expected to the recently depicted omnipresent access to the Internet and a developing number of gadgets. Information is made and conveyed from different frameworks working continuously.

For instance, internet-based life always stages total data about client exercises and collaborations; for example, one of the most prominent social locales Facebook has more than 650 million day by day dynamic clients. The yield pace of the framework can be likewise significant when about continuous examinations are required. Such an on-the-fly investigation is required in proposals frameworks when the client's information influences substance gave by a site; genuine models are online retail stages, for example, Amazon.com.

This angle requires different methods for putting away the information to expand the speed and at times utilizing segments. Situated database or one of the diagrams less frameworks (NoSQL) can carry out the responsibility since enormous information is seldom all around organized but vast information is not just testing; however, essentially makes openings. They are, among the others: making straightforwardness, enhancement and improving execution, age of extra benefits and nothing else than finding new thoughts, administrations, and items.

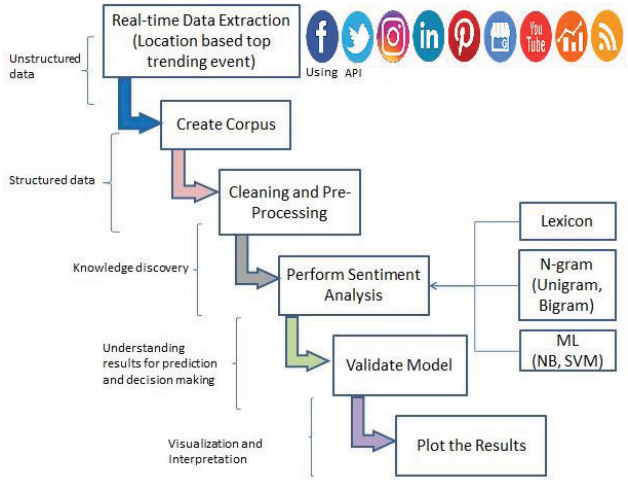

Fig.1: API Internal process of Sentiment Analysis

B. Social Networks and Communications

One of the patterns prompting the ascent of enormous information is Web 2.0. It is a significant move from static sites to intelligent ones with client created content (UGC). The advancement of Web 2.0 brought about numerous administrations, for example, blogging, podcasting, long-range interpersonal communication and bookmarking. Clients can make and share data inside open or shut networks, and that adds to volumes of enormous information.

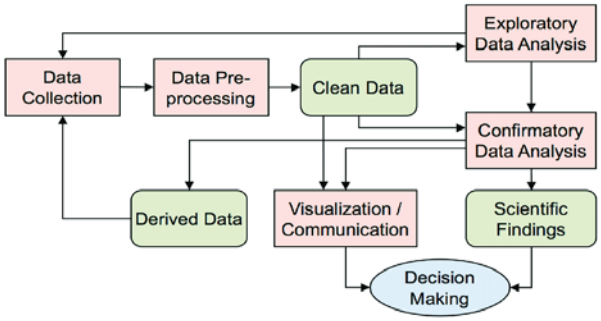

Fig. 2: Data Analytical Process in Cloud

Web 2.0 prompted the formation of online networking that presently is methods for making, contributing and trading data with others inside networks by electronic media. Internet-based life can be likewise condensed as "based on three key components: substance networks and Web 2.0". Every one of those components is a crucial factor and is essential for online networking. One of the most significant variables boosting online life is the expanding number of consistently Internet-associated cell phones, for example, cell phones and tablets. Twitter is a micro blogging stage, which consolidates highlights of online journals and informal community administrations. An individual who buys into different clients is alluded to as "adherent" and gets constant updates from that individual [10].

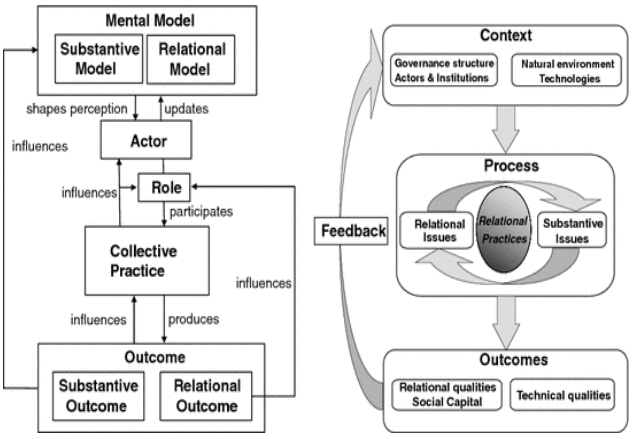

Fig.3 Procedural Operation of Social Influence

In any case, clients do not need to include individuals who are their adherents. Twitter can be gotten to from different administrations, for example, the official Twitter website page, versatile applications from outsiders and SMS administration. As Twitter is an amazingly tremendous help, particularly in the information structure is minimized.

Hence, it powers clients to post short remarks creators of this paper accept this is a decent wellspring of data in the feeling of previews of temperaments and emotions just as for forward-thinking occasions and current circumstance comments on texts. Besides, Twitter is a typical PDA specialized device for legislators and different VIPs forming or affecting the way of life and society of vast networks of individuals. Along these lines, Twitter was picked for a trial information hotspot for this work on foreseeing the securities exchange.

C. *Working Method*

It is a continuous analysis system; to start with, we need to sign in to Facebook. At that point, getting information is finished. At that point, it applies the tokenizing procedure. Also, evacuate stop words, spilling. It will do a post for registering grown-up remarks with it.

After characterization, the chart is produced. Simultaneously arrangement of a post is made in various classifications like games, legislative issues, and so on. At that point, it demonstrates the graphical exposé of sorted comments, texts, emogies or posts.

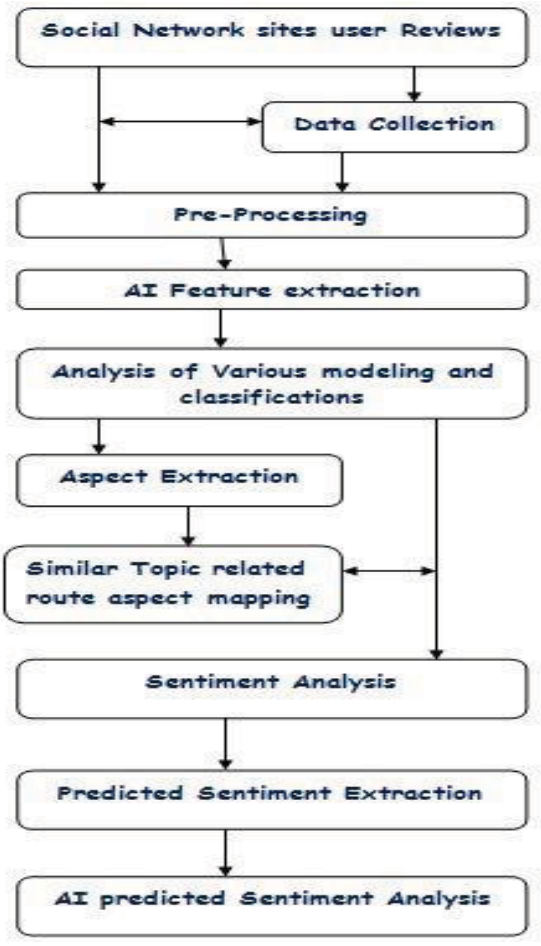

Fig.4: Overview of Entire working Mechanism

Algorithm

Step 1: Login through standard client or user, who contain the Twitter, Facebook or Instagram account beside by the account added into their Social Network website.

Step 2: Insert post resting scheduled the Twitter and Facebook, the Instagram report in favors of simple recovery of a post.

Step 2.1: Get User (Posts that is mostly related to Adult content) as Input to Web browser.

Step 2.2: for i=0 to MAX //Maximum = Maximum no of Posts

Step 2.3: Visit i (Social Network websites (Twitter/ Facebook /Instagram Portal) when logged in by a user.

Step 2.4: Go to step 2.2 until MAX.

Step 2.5: Output as CP can view all policy and schemes.

Step 3: Call RC Function.

Step 4: call to IE Function

Step 4.1 : Get Client Profile (Facebook posts) as Input.

Step 4.2: Call gathering NLP

Step 4.3: Process NLP as Removing Stop Word.

Step 4.4: Get Relevant Information of positive, Negative d neutral comments along with the adult satisfied verify.

Step 5: Display concerned the result in the cloud platform form how many posts Contain adult contents.

Step 6: Stop.

Proposed Technique

Twitter or Facebook and other social Network data gaining and preprocessing: Facebook and Twitter any social network sites messages are retrieved in real-time using Spilling API permits recovering tweets in semi- constant (server postponements must be mulled over). There are no exacting rate limit confinements; anyway, just a segment of mentioned tweets is conveyed. Spilling API requires a diligent HTTP association and authentication [8].

While the association is updated alive, messages are presented to the customer. Spilling API offers plausibility of separating tweets as indicated by a few classes, for example,

area, language, hash labels or words in tweets. One disservice of utilizing Streaming API is that it is challenging to recover tweets from the past along these lines. At first, we are gathered Posts or remarks or Tweets intermittently in an accessible way. Tweets were recovered for the most part for Apple Inc. (exchanged as 'AAPL') to guarantee that datasets would be adequately massive for characterizations. Recovered information contains many clam ours, and it is not legitimately reasonable for structure grouping model and afterwards for supposition location dependent on several times tweets or posts.

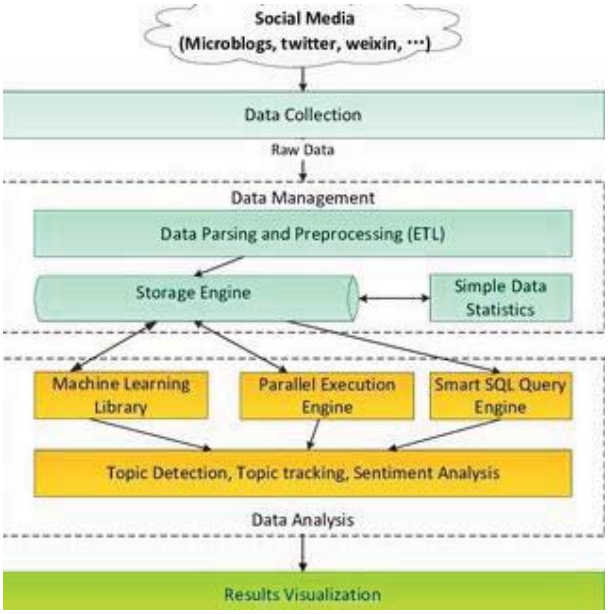

Fig.5 Data Collection and analysis procedure

During preparing information techniques following advances were taken [8, 9]. Language identification data about the language of the tweet is not always right. Just tweets in English or any nearby reasonable dialects are utilized in this exploration work. Copy evacuation – Facebook or Twitter permits reposting messages. After pre- handling, each message was spared as a sack of words model – a standard procedure of improved data portrayal utilized in data recovery [10].

In this area, we present our multiple arrange scientific information arrangement called Big-PFM, which leads to substantial interpersonal organization mining and examination utilizing prevalent fellowship mining with the Map-Reduce model.

The Following pseudo code speaks to the usefulness of the Comment or Post Analyzer in a proficient way:

1. Post or Comment _Analyzer
2. Input: Preprocessed comments
3. Output: Comment or Post arranged as positive, negative or neutral operation.

IV. RESULTS

We assess our scientific information arrangement BigPFM both logically and accurately. The expository outcomes demonstrate that BigPFM is adaptable with the measure of got to or potentially put away free community information. The exploratory outcomes with the Stanford Network Analysis Project (SNAP) sense of self Facebook datasets from the Stanford University huge Network Dataset collected works demonstrates the productivity and reasonableness of Big- PFM in speaking to kinships as charts, finding bunches now and again associated companions, and shaping principles for companion proposal.

A. Stage 1 - Representation of Social Users
B. Stage 2 - Discovery of Frequently Connected Friends
- Group the sets of <φ, v>returned by the guide work and
- Count the quantity of companions of every client u ? V. BigPFM

C. Stage 3 - Friend Recommendation
D. Stage 4 – Comments
E. Stage 5 – Posts
F. Stage 5 – Sentiment Calculation

Sentiment Calculation

The Following pseudo code represents the functionality of the Comment or Post Analyzer in an efficient manner:

Post or Comment _Analyzer
Input: Preprocessed comment

Output: Comment or Post categorized as positive, negative or neutral.

OverallPol: Polarity of the whole comment.

SentPol: Polarity of the sentence within the comment.

POS: Part of Speech of the word.

```
BEGIN
overallPol = 0
For each sentence in comment
{
     SentPol = 0
     For each word in sentence
     {
   If Polarity[word] == Positive
     {
     If POS[word] == Adjective
     sentPol = sentPol + 5
     Else if POS [word] == Adverb
      sentPol = sentPol + 4
     Else if POS[word] == Verb
        sentPol = sentPol + 3 Else:
```

```
        sentPol = sentPol + 2
        Else if POS[Symbol] == Emoji
        sentPol = sentPol + 1
    }
    Else if Polarity[word] == Negative
    {
    {
    If POS[word] == Adjective
          sentPol = sentPol – 5
    Else if POS[word] == Adverb
          sentPol = sentPol – 4
    Else if POS[word] == Verb
          sentPol = sentPol – 3
    Else
          sentPol = sentPol – 2
    Else if POS[Symbol] == Emoji
          sentPol = sentPol - 1
    }
     Else if word is a Negation
          word  sentPol = - sentPol
    }
    overallPol  =     overallPol +
     sentPol
    }
    If overallPol &gt; 0
    Return 'Positive Comment'
  Else if overallPol & lt; 0
 Return 'Negative Comment'
 Else
Return 'Neutral Comment'
End
```

Sentiment Analyzer:

Fig.6: Sentiment Opinion Views

V. CONCLUSION

A significant information logical arrangement of existing data by analyses based on Data collection (Raw data) directs interpersonal data mining with parsing and preprocessing mining by using Machine learning algorithms. In this research paper we proposed new working mechanism to identify and analyzed predictable sentiment analysis approaching methods based on existing datasets. Our work hugely help into identifies predictable critical decision making solutions for every applications.

REFERENCES

[1]. N. Collier, K. Takeuchi, A. Kawazoe, T. Mullen, and T. Wattarujeekrit. 2003. A framework for integrating deep and shallow semantic structures in text mining. In Proceedings of the Seventh International Conference on Knowledge-based Intelligent Information and Engineering Systems. Springer-Verlag.

[2]. Janyce Wiebe., Learning subjective adjectives from corpora. In Proc. 17th National Conference on Artificial Intelligence (AAAI-2000), Austin, Texas, July, 2000.

[3]. Ide, N.: Making senses: Bootstrapping sense-tagged lists of semantically-related words. In: CICLING (2006).

[4]. Andreevskaia, A., Bergler, S.: Mining wordnet for a fuzzy sentiment: Sentiment tag extraction from wordnet glosses. In: EACL (2006).

[5]. Wiebe, J.: Learning subjective adjectives from corpora. In: Proc. 17th Nat. Conf. on Artificial Intelligence and 12th Conf. on Innovative Applications of Artificial Intelligence, AAAI Press / The MIT Press,735–740, (2000).

[6]. Kil, J., Lloyd, L., Skiena, S.: Question answering with lydia. Proc. 14th Text Retrieval Conference (TREC 2005) (2005).

[7]. Jaap Kamps, Maarten Marx, Robert J. Mokken, and Marten de Rijke. 2002. Words with attitude. In Proceedings of the 1st International Conference on Global Wordnet, Mysore, India

[8]. J Wiebe., Instructions for annotating opinions in newspaper articles. Technical Report TR-02- 101, University of Pittsburgh, Pittsburgh, PA, 2002.

[9]. Kim, S.M., Hovy, E.: Determining the sentiment of opinions. In: Proceedings of the Coling Conference, (2004).

[10]. Wiebe, J.: Learning subjective adjectives from corpora. In: AAAI/IAAI, 735–740, (2000).