Descriptive Study on EmoMining from SoNet Sites

Jaskaranjit Kaur[1*], Navneet Kaur[2]

[1]Dept. of Computer Science & IT, Lyallpur Khalsa College, Jalandhar, Punjab, India
[2]Dept. of Computer Science & IT, Lyallpur Khalsa College, Jalandhar, Punjab, India

[*]*Corresponding Author: jaskaranjitkaur11@gmailcom, Tel.: 7710289748*

Abstract-As social networking sites are popular, so they become major part of person's social interaction. These social networking sites are rich in emotions where people share their feelings, opinions, emotions. Extracting emotions from these social networking sites play an important role in various fields. Many techniques are proposed by various authors to extract emotions from these social networking sites. This paper presents various studies carried out in the field of EmoMining. The basic objective is to extract the emotional content of texts in online social networks. For this purpose, text mining techniques are performed on comments retrieved from a social network. This paper includes data collection, database schemas, data pre-processing and data mining steps. The informal language of online social networks is a main point to take into account before performing any emotion mining techniques. Here EmoMining related to tweets from social networking site is presented. Also emotion mining based on fuzzy rule base is also discussed along with brief description of fuzzy set theory.

Keywords-Emotion mining, social networking, tweets, twitter, fuzzy rule, fuzzy set theory.

I. INTRODUCTION

Emotions are a natural part of human nature which will be thought of as hereditary. Additionally it's been found that expression of a specific feeling by totally different individual is identical. Some persistent emotions that last for much longer time end in mood. Mood may be a result of a mix of certain emotions of an individual. On the entire emotions may be classified into two categories: basic and complex. Basic emotions are joy, sadness, anger, fear, disgust and surprise as discussed by Ekman [1]. The complex emotions are a combination of two or additional basic emotions that are experienced by an individual at an instance [2].

In the globe where we live, our ancestors had once seen an era when people communicate through letter posts, telegram etc. A very modern technique people used is telephone, wired and gradually became wireless. Nowadays social networking becomes a very powerful and vital tool for interacting with friends, relatives, acquaintances and among people of different communities, region and countries. This system affected and altered the way people communicate with each other. Indeed, this becomes a very important part of day to day life. Social networking websites like Facebook, LinkedIn, and Twitter provides an awfully powerful medium and a platform for communication among people resulting to mutual learning and sharing of knowledge. Social Web based applications also provides opportunities to people or to a group of individuals for sharing of data/information in form of Wall posts, comments, chat and discussions. This provides an emotionally rich environment to the subjects to Interact [5].

Data mining is a powerful tool that can facilitate to search out patterns and relationships within our information. Data mining discovers hidden information from large database. The overall goal of the data mining method is to extract information from a knowledge set and transform it into an apparent structure for further processing. Social networks can be used in many business activities like increasing word-of-mouth marketing, marketing research, General marketing, Idea generation & new product development, Co-innovation, Customer service, Public relations, Employee communications and in Reputation management [7].

Section I contains the introduction of Data Mining and Emotion Mining , Section II explains Text Mining followed by Section III which explains Emotion Mining, Section IV & V contain Emotion Classification and Emotion Analysis resp. ,section VI explain Fuzzy Logic, Section VII contains Background Work, Section VIII describes role of Fuzzy Logic in Emotion Mining and Section IX concludes research work with future directions.

II. TEXT MINING

Text mining is also named as Text Analytics. Text mining finds out the various facts, relationships and assertions which would otherwise remain suppress under big data. These extracted facts are then turned into well-defined structured form for further analysis, visualization and which are then integrated with structured data in databases or warehouses. And its refinement can be performed using machine learning (ML) systems. Text mining or knowledge discovery is that sub procedure of information mining, which is generally being utilized to find concealed examples and noteworthy data from the tremendous measure of unstructured composed material. The multiplication of clouds, research and advancements are in charge of the making of immense volumes of information. This sort of information can't be utilized until or unless particular data or example is found. For this text mining utilizes strategies of various fields like machine learning, perception, case-based thinking, content analysis, knowledge management, natural language processing and information retrieval. This paper contains the audit of text mining methods, instruments and different applications. Text mining is the technique for inspecting huge accumulations of composed data to produce relevant information and to convert the unstructured content into some organized information which can be used in further study.

III. EMOTION MINING

EmoMining from text comprises of following phases:
- Raw Data Collection
- Data pre-processing utilizing natural language processing methods
- Feature Selection
- Identification of Emotions
- Classification of Emotions
- Evaluation stage

An example of emotion analysis from twitter dataset is shown in table below:

Table 1: Emotion Mining from Twitter Dataset

Emotion	Example
Happiness	eryHappy #Enjoying #PartyyyTime
Sadness	#HeartBreak #FeelingSick
Love	#LoveMyFamily #Beautiful
Surprise	#incredibleBuilding #Wowww
Anger	#SoDisgusting #Bastard #HateIt

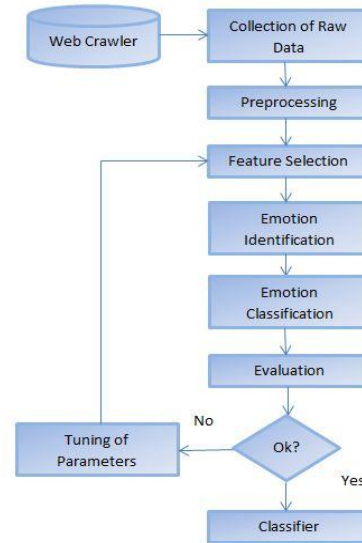Basic steps in emotion extraction framework are presented in the flowchart shown in below figure:

Figure 1: Basic Steps in Emotion Mining

The above diagram describes the basic steps which we follow for extracting emotions from any social networking site. In the first step raw data is collected from the web crawler. Then the collected data is pre-processed in order to remove noise and missing values to make it suitable for data mining. Data pre-processing is a data mining practice that includes changing crude information into an understandable format. Real-world data is frequently incomplete, inconsistent and lack in certain behaviours and contain numerous mistakes. So it is necessary to pre-process the collected data first. After data pre-processing features are selected. Author profiling in online social media differs from other platforms as the dialect that has been used by social media users is casual, unstructured and boisterous. A common approach of inducing user's traits in social media is to model the writing habits of various users by extracting out various features from texts they have posted. These writing habits can be inferred by extracting the various features of the user like in case of twitter the number of Female users, Male users, Avg. tweets per male user, Avg. tweets per female user, Avg. length of tweets per user, Majority age group, Avg. Openness score, Avg. Conscientiousness score, Avg. Agreeableness score, Avg. Emotional Stability score, Avg. Extroversion score. After selecting features the user's emotions are identified whether it comes under negative emotion or positive emotion categories which are further classified into following emotion categories: sadness, happiness, surprise, fear, disgust and anger.

IV. EMOTION CLASSIFICATION

- Keyword Classification: Detection of emotions is done on the basis of related set of keyword found in input. It utilizes similar words and opposite words in WordNet to determine word sentiments.

- Learning Classification: In this, emotions are detected on the basis of trained classifier that applies many machine learning theories like support vector machines. The drawback to determine emotions indicating factors from input.
- Hybrid Classification: Rule based technique is utilized in this method for extracting semantics relevant to particular emotions. Semantics and attributes are linked with emotions in form of emotion association rules. This technique is mixture of learning based technique and keyword based technique.

V.   EMOTION ANALYSIS

- Spotting Keywords: This technique is utilized to predict the emotion of writer by identifying affective words from text. These techniques are so accepted due to their economical benefits and simplicity.
- Lexical Affinity: These techniques assign for each word probabilistic affinity for a certain emotion. This method depends on WordNet where two words can be associated by a string of stepwise equivalent words.
- Natural Language Processing: This technique is not more feasible to use because training data are hard to find. Statistical methods for the programmed examination of human language data.

VI.   FUZZY SET THEORY

The essential to extract knowledge from domain specialist or from training data obtained within the globe, referred to as knowledge acquisition, is a task chief to applications of knowledge engineering. However, the imprecision intrinsic to human expertise and to the actual world makes this venture quite difficult to accomplish with the usage of traditional scientific model, as such a model tries to precisely characterize all the characteristics of the expert system we are constructing and absence of flexibility. Fuzzy set theory, introduced by Zadeh [8], should provide the needed tools needed to deal with the vagueness of our knowledge, and allows us to represent the parameters of a system using unclear linguistic terms rather than exact mathematical values. However, we are left with the venture of constructing the affiliation features of these parameters and the Fuzzy policies of the system. This is the necessary block in the development of FES, and if finished manually, would require high priced resources from domain specialist and expertise engineers. Fortunately, for many real-world applications, numerical data can be often simply obtained from tools or from the environment. Several researchers have come up with approaches to use this facts to permit the system to construct its very own fuzzy rules with the usage of a variety of sorts of learning algorithms.

The main features of fuzzy logic are as follows [9]:

- It contains matter of degree.
- Fuzzy logic is flexible
- Any system can be fuzzified
- Information is decomposed into cluster of variables.

Fuzzy logic has two unique meanings [10]. In a narrow experience, fuzzy logic is a logical system, which is an extension of multi valued logic. However, in a wider experience FL is nearly synonymous with the conecpt of fuzzy sets, a principle which relates to lesson of objects with smooth boundaries in which membership is a matter of degree. Fuzzy logic is the codification of common sense — use frequent feelings when you implement it and you will probably make the right decision.

A fuzzy rule is defined as a conditional declaration in the form:

$$\text{IF } x \text{ is } A \text{ THEN } y \text{ is } B$$

Where x and y are linguistic variables: A and B are linguistic values decided by fuzzy sets on the universe of discourse X and Y, respectively.

Let us consider an example of AGE parameter, linguistic variables for AGE parameter will be young, old, very old. Similarly another example height short, medium and tall is the linguistic variables.

A membership function (MF) is a curve that defines how every point in the input space is mapped to a MV (or degree of membership) between 0 and 1. Membership functions are μ_{young}, μ_{old}.

VII.   BACKGROUND WORK

As internet becomes universal mean to communicate, world has changed at a very fast pace. Social networking sites are a common means to interact with each other. These networking sites are most searched pages on the internet. Dhawan et al. [2] reviewed emotion mining techniques in social networking sites. He stated that latest research shows that it is advisable to deal with various kinds of sentences by various strategies. Also author describe the requirement of some tools to mine certain emotions from different sources of data that gives accurate outcomes.

Mohamed et al. [3] surveyed enhancement done in the field of emotion mining along with comparative study for various techniques. Also, this paper presented investigation on the technology utilized in this area and also on the implementation of these techniques. These techniques are categorized into two areas: lexical based techniques and ML based techniques. A descriptive process cycle to build such type of systems is also presented in this paper.

A novel technique to study friendship emotions and relations is presented by Yassine et al. [4]. The main aim of this work is to extract information about emotions from social networking sites. Author shows interest to depict if writer shows his/her emotions through their writings. This study

considers a case study called Lebanese Facebook users. It is based on k-means clustering technique and is unsupervised technique.

Kumari et al. [5] discussed different methods and techniques to mine emotion from text. Emotion Analysis algorithms are utilized to estimate the public emotion on different problems discussed in the social network. On the basis of the comments, the author can categorize the type of friendship.

Dhawan et al. [6] presented a novel perspective to study expressions of emotions' in online social networks. This paper adopts unsupervised technique; mainly it utilizes the nearest neighbour algorithm and k-means clustering algorithm. Results demonstrate enhanced accuracy for model in predicting emotions and determining subjectivity of texts.

Sharma et al. [7] discussed about data mining techniques. This paper presented a survey of the work performed in the area of social network analysis and also concentrates on the future scope in research on social network analysis. This paper presented study related to social networks utilizing Web mining methods.

Qamar et al. [8] proposed to utilize fuzzy logic to detect emotional content from text. Fuzzy logic was developed to handle ill-defined concepts. The transfer from one physiological state of emotion to another is gradual and it is easy to model by fuzzy logic technique.

Dutta et al. [9] presented the task based on fuzzy logic execution to perceive emotions from text. This technique is based on execution of fuzzy logic to recognize emotions from text in MATLAB environment.

Colneric et al. [11] explored the use of deep learning for emotion detection and created three large collections of tweets labelled with Ekman's, Plutchik's and POMS's classifications of emotions. Recurrent neural networks indeed outperform the baseline set by the frequent bag-of-words models. Experiments advise that it is better to train RNNs on sequences of characters than on sequences of words.

Stojanovski et al. [12] exploit an convolution neural network architecture for emotion analysis in Twitter messages related to sporting events on 2014 FIFA world Cup. In this paper, seven different kinds of emotions were evaluated using hash tag labelled tweets that were collected from Twitter Streaming API. The training of the network is performed on two samples containing 1000 and 10000 tweets on which this approach achieves 50.12% and 55.77% accuracy respectively. Moreover, they have presented the analysis of this approach on three different games that have great impact on Twitter users.

Mishne et al. [13] addressed the task of classifying blog posts on the basis of mood of the writers. They obtained a huge corpus of blog posts from one of the largest online blogging communities live journal. The author took the advantage of the live journal that allows writers to update their current mood from the 132 given categories. Yahoo API was used to get a list of 1000 web pages containing a live journal blog post with each kind of mood.

Roberts et al. [14] presented a publicly available corpus of tweets annotated with seven different kinds of emotions comprising Ekman's six emotions and love. The main goal of this research is to present emotion distributions in different emotion evoking topics. Tweets were collected from twitter using Twitter API on 14 different topics that evoke emotions. Moreover, they described a baseline method for automatically annotating of emotions for tweets in which they used previously annotated tweets by professional annotators for training the data. They used series of binary SVM classifiers to detect each of the seven emotions annotated in the corpus they created. Additionally, linguistic style features are also presented in this paper for topic "U.S Elections 2012".

Wang et al. [15] exploit a technique to automatically annotate a large amount of data. They extracted large amount of tweets (2.5 million) from twitter instead of using already annotated corpus which consists of just thousands of tweets. The main focus of this research is to think about the viability of different element mixes and in addition the impact of the measure of the preparation information on the feeling analysis work.

Table 2: Study of Various Techniques

S. No.	Author	Title	Description
1.	Dhawan et al.	Emotion Mining Techniques in Social Networking Sites	Reviewed emotion mining methods in social network sites.
2.	Mohamed et al.	The Road to Emotion Mining in Social Network	Surveyed enhancement done in the field of emotion mining along with comparative study for various techniques.
3.	Yassine et al.	A Framework for Emotion Mining from Text in Online Social Networks	Presented a novel technique to study friendship emotions and

			relations
4.	Kumari et al.	Emotion analysis using text mining on social networks	Discussed different methods and techniques to mine emotion from text
5.	Dhawan et al.	A Framework for Polarity Classification and Emotion Mining from Text	Presented a new perspective for studying emotions' expression in online social networks.
6.	Sharma et al.	Literature Review and Challenges of Data Mining Techniques for Social Network Analysis	Discussed about data mining techniques.
7.	Qamar et al.	Emotion Detection from Text using Fuzzy Logic	Proposed to use fuzzy logic in detecting emotional content from text.
8.	Dutta et al.	An Effectual Emotion Recognition System Based on Fuzzy Logic	Presented the work that depends on execution of fuzzy logic in perceiving emotions from text.
9.	Colneric et al.	Emotion Recognition on Twitter: Comparative Study and Training a Unison Model	Explored the use of deep learning for emotion detection.
10.	Stojanovski et al.	Emotion identification in FIFA world cup tweets using convolutional neural network	Exploited an convolutional neural network architecture for emotion analysis in Twitter messages.
11.	Mishne et al.	Experiments with mood classification in blog posts	Addressed the task of classifying blog posts on the basis of mood of the writers.
12.	Roberts et al.	EmpaTweet: Annotating and	Presented a publicly
13.		Detecting Emotions on Twitter	available corpus of tweets annotated with seven different kinds of emotions comprising Ekman's six emotions and love.
13.	Wang et al.	Harnessing twitter" big data" for automatic emotion identification	Exploited a technique to automatically annotate a large amount of data.

VIII. ROLE OF FUZZY LOGIC IN EMOTION MINING

When a sentence starts, there are no **cues** about the mood of the user. So, it is equitable to assume that the mood of a new user is neutral. As the emotion extraction engine acquires more data (e.g., when the user interacting with other user), then the below fuzzy rules can be applied and the mood of the user can be evaluated.

Fuzzy rules are created on the basis of users comments on social networking website.

Rule 1: In a sentence if feelings are positive AND adverb is also positive
        Then comment is subjective.
Rule 2: In a sentence if feelings are negative AND adverb is also positive
        Then comment is moderately subjective.
Rule 3: In a sentence if feelings are negative AND adverb is also negative
        Then comment is objective.
Rule 4: If emoji is positive
        Then comment is subjective.
Rule 5: If emoji is negative
        Then comment is moderately subjective.

Below Table 3 describes the classifications of comments based on above specified Fuzzy Rules.

Table 3: Result of Classification of comments based on Fuzzy Rules

Comment	Comment Type
Beautiful Profile Pic	Subjective
Good ☹	Moderate Subjective
Very Good ☺	Subjective
How?	Objective
Hii	Moderate Subjective
Hello ☺	Subjective

Table 4: List of Abbreviations

Abbreviations	Meaning
EmoMining	Emotional Mining
SoNet	Social Networking
Avg.	Average
FL	Fuzzy Logic
MF	Membership Function
MV	Membership Value
FES	Fuzzy Expert System

IX.   CONCLUSION AND FUTURE SCOPE

In this paper, various techniques related to EmoMining are discussed. Twitter emotion mining is presented as an example. Fuzzy set theory is explained here in brief as fuzzy rules are used to mine emotion. Now, emotion mining is the attracting area of research as it is vital to mine emotions of people for various purposes. EmoMining has gain much more attention after people increase the use of social networking sites (like facebook, twitter, so on) to share their views, opinions, their feelings and so on. Major methodology such as fuzzy logic towards Emotion Extraction from text has been discussed in this paper. We tried to present an overview of text mining approach with Fuzzy set theory and its applications. In future hybridization of both AI and classification algorithm can be applied in which fuzzy logic and neural network will be utilized. Various tools can be used for capturing the social networking websites comments like Nvivo and Ncapture.

REFERENCES

[1] P. Ekman, *"An argument for basic emotions",* Cognition and Emotion, **1992**, pp. **169-200**.

[2] SanjeevDhawan, Kulvinder Singh, DeepikaSehrawat, *"Emotion Mining Techniques in Social Networking Sites",* International Journal of Information & Computation Technology, Vol. **4**, No. **12**, **2014**, pp. **1145-1153.**

[3] Hany Mohamed, AymanEzzat, Mostafa Sami, *"The Road to Emotion Mining in Social Network",* International Journal of Computer Applications, Vol. **123**, No. **18, 2015**, pp. **41-47.**

[4] Mohamed Yassine, Hazem Hajj, *"A Framework for Emotion Mining from Text in Online Social Networks",* IEEE, International Conference on Data Mining Workshops, **2010**, pp. **1136-1142.**

[5] RashmiKumari, MayuraSasane, *"Emotion analysis using text mining on social networks",* International Journal of Innovative Research in Technology, **2015**, pp**. 20-27.**

[6] SanjeevDhawan, Kulvinder Singh, VandanaKhanchi, *"A Framework for Polarity Classification and Emotion Mining from Text",* International Journal Of Engineering And Computer Science, **2014**, pp. **7431-7436.**

[7] Anu Sharma, Dr. M.K Sharma, Dr. R.K Dwivedi, *"Literature Review and Challenges of Data Mining Techniques for Social Network Analysis",* Advances in Computational Sciences and Technology, Vol. **10**, No. **5**, **2017,** pp. **1337-1354.**

[8] SaqibQamar, Parvez Ahmad, *"Emotion Detection from Text using Fuzzy Logic",* International Journal of Computer Applications, Vol. **121**, No.**3**, July **2015**, pp. **29-32**.

[9] Er. VibhaDutta, Dr. R. C. Gangwar, Er. MohitMarwaha, *"An Effectual Emotion Recognition System Based on Fuzzy Logic",* International Journal of Latest Trends in Engineering and Technology, Vol. **8**, Issue. **1**, pp.**122-131**.

[10] Mohan Arava, TalasilaVamsidhar, *"Driver Emotional Status Recognition Using Fuzzy Logic in ADAS: An Indian Perspective",* International Journal of Advanced Research in Computer Science and Software Engineering, Vol. **4**, Issue. **10**, October **2014**, pp. **328-334**.

[11] NikoColneric, JanezDemsar, *"Emotion Recognition on Twitter: Comparative Study and Training a Unison Model",* IEEE Transactions on Affective Computing, **2018**, pp. **1-14**.

[12] Stojanovski, Dario, et al. *"Emotion identification in FIFA world cup tweets using convolutional neural network."*Innovations in Information Technology (IIT), **2015** 11th International Conference on.IEEE, **2015**.

[13] Mishne, Gilad. *"Experiments with mood classification in blog posts."*Proceedings of ACM SIGIR 2005 workshop on stylistic analysis of text for information access.Vol.**19**. **2005**.

[14] Roberts, Kirk, Michael A. Roach, Joseph Johnson, Josh Guthrie, and Sanda M. Harabagiu. *"EmpaTweet: Annotating and Detecting Emotions on Twitter."*In LREC, vol. **12**, pp. **3806-3813**.**2012.**

[15] Wang, Wenbo, Lu Chen, KrishnaprasadThirunarayan, and Amit P. Sheth. *"Harnessing twitter" big data" for automatic emotion identification."* In Privacy, Security, Risk and Trust (PASSAT), 2012 International Conference on and 2012 International Confernece on Social Computing (SocialCom), pp. **587-592**. IEEE, **2012**.

AUTHORS PROFILE

Ms.JaskaranjitKaur did Master of Technology in Computer Science & Engineering from DAV University Jalandhar in 2015. She has teaching experience of 3 years. She is currently working as an Assistant professor in Department of Computer Science & IT, Lyallpur Khalsa College, Jalandhar, Punjab, India. Her main research work focuses on Data Mining.

Ms.NavneetKaur. She did her Master of Engineering in Software Engineering from Thapar Institute of Engg. & Technology, Patiala in 2005. She has teaching experience of 13 years. She is currently working as Assistant Professor in Department of Computer Science & IT, Lyallpur Khalsa College, Jalandhar, Punjab, India. Her main research work focuses on Software Engineering, Software Reuse, Genetic Algorithms and ACO.