

Emotion analysis using text mining on social networks

Rashmi Kumari¹, Mayura Sasane²

¹Student, M.E-CSE, Parul Institute of Technology, Limda, Vadodara, India

²Assistance Professor, M.E-CSE, Parul Institute of Technology, Limda, Vadodara, India

Abstract- Today, everyone has numerous possibilities to express his opinion, share content with others and access almost any information in a couple of seconds. Social networking Websites like facebook, MySpace, twitter welcome millions of visitors daily, the content produced by these millions of users is obviously enormous. They can create opportunities to establish interaction among people leading to mutual learning and sharing of knowledge, such as chat, comments. Sentiment Analysis is the technique to analyze the emotion from a text document, message, or similar content. Emotion Analysis algorithms are used to predict the public emotion on various issues getting discussed in the social network threads. In this research, we are discussing various techniques and methods to mining emotion from text. On the basis of text comments we categories which type of friendship exist.

Index Terms- Online social networks, Data mining, emotion mining, Text Mining, friendship

I. INTRODUCTION

The society, the world we live in, 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. Today for interacting with friends, relatives, acquaintances and among people of different communities, region and countries Social Networking becomes a very powerful and important tool. This technique affected and changed the way people interact with each other. Indeed, this becomes a very important part of day to day life. Social networking websites like Facebook, LinkedIn, Twitter provides a very powerful medium and a platform for communication among individuals leading to mutual learning and sharing of knowledge. Social Web based applications also provides opportunities to individuals or to a group of individuals for sharing of information in form of Wall posts, comments, chat and discussions. This provides an emotionally rich environment to the subjects to Interact. In this paper an attempt is taken to extract emotions shared in texts in online social networks

in form of messages, comments and wall posts etc. Purpose is not to mine a specific emotion but also to determine whether a text contains emotions or not i.e. to determine subjective message showing emotions of writer or objective message. This approach also predicts relationship strength between two individuals or a group of individuals on basis of content of comments they share. The focus is that the Individual shares his emotions or not, which is generally influenced by the strength of relationship between individuals or a group of individuals. Thus this technique also helps in finding the relationship status and to distinguish close friends, friends or acquaintances etc. An important aspect for this type of mining is use of informal language by the individuals which could not match with settings and formal language. These factors should also be taken into consideration in any type of text mining. This paper uses national facebook users as a case study and applies an unsupervised technique to categories text based on subjectivity.

II. DESCRIPTION OF TECHNIQUES FOR EMOTION ANALYSIS

3.1 Keyword spotting technique:

This technique is used for predict the emotion of the writer by identifying affective words from text. These techniques are so popular because of their simplicity and economical advantages.

3.2 Lexical affinity: These techniques assign for each word probabilistic affinity for a certain emotion. This technique is based on WordNet where two words can be connected by a string of stepwise synonyms.

3.3 Statistical Natural language processing:

This technique is not more feasible to use because training data are hard to find. Statistical techniques for the automatic analysis of natural (human) language data.

III. METHODS FOR EMOTION CLASSIFICATION

4.1 Keyword based detection method:

Emotions are detected based on the related set of keyword found in the input text. It uses synonyms and antonyms in WordNet to determine word sentiments.

Limitations:

Ambiguity in keywords definitions:

The meanings of keywords could be many and different. Besides those words conveying for emotional touch, most words could change their meanings according to different usages and contexts. Without all their synonyms could have different emotions in some extreme cases such as ironic or cynical sentences.

Incapability of recognizing sentences without keywords:

Keyword-based approach is totally based on the set of emotion keywords. For example, "I passed my qualify exam today" and "Hooray! I passed my qualify exam today" should imply the same emotion (joy), but the former without "hooray" could remain undetected if "hooray" is the only keyword to detect this emotion.

Absence of Linguistic Information:

Syntax structures and semantics also have influences on what we express. For example, "I laughed at him" and "He laughed at me" would suggest different emotions from the first person's perspective. As a result, to ignoring linguistic information also poses a problem to keyword based methods. Keyword based methods should also detect in linguistic information to detect emotions more accurately.

4.2 Learning based methods:

It try to detect emotions based on a trained classifier which apply various theories of machine learning such as a support vector machines. There is a disadvantage in determining emotions indicating factors from input text in different emotions.

Tough to Determining Emotion Indicators:

The first problem is, though learning-based methods can automatically determine the probabilities between features and sentiments, these methods still need keywords, but just in the form of features. The most something that is known or understood without proof or evidence features may be emoticons which can be seen as author's emotion annotations in the texts. Problems related to cascading are same in keyword-based methods.

Very-simplified Emotion Categories:

Nevertheless, lacking of efficient features other than emotion keywords, most learning-based methods can only classify sentences into two categories, which are positive and negative. Although the number of emotion labels depends on the emotion model applied, we would expect to refine more categories in practical systems.

4.3 Hybrid methods:

These methods utilize rule based approach to extract semantics related to specific emotions. Attributes and Semantics are associated with emotions in the form of emotion association rules. This method is combination of both keyword based method and learning based method.

IV. EXISTING FLOW CHART

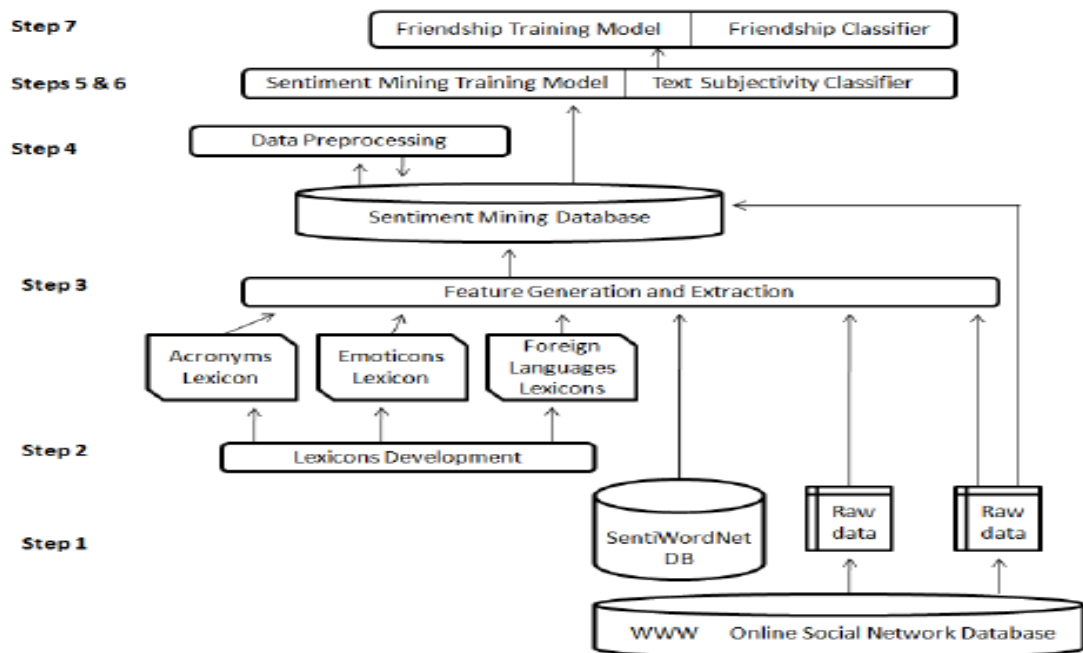


Figure 1: Existing flow chart

V. PROPOSED WORK

STEP 1: Raw Data

First collect data from social network i.e. facebook and this collection are done through creating Social network application interacting with social network API. This application fetches user information and stores them in a database.

STEP 2: It collect raw data i.e. comments between two users and separate raw data means which user having status and which user post comments on status.

STEP 3: Feature generation and extraction:

This step computes new features from available raw data. It separates emoticons and affective words from available raw data. Comments collected from facebook along with features computed in this step will be stored in the sentiment mining database which is the database used for analysis.

STEP 4: Linguistic variables:

Linguistic variable is a variable whose values are words in a natural language. It contains words with adverbs for example “happy”, “very happy”.

STEP 5: Construct rules:

5.5.1 If (affective word is simple positive and adverb is also positive) then comment is subjective.

5.5.2 If (affective word is simple negative and adverb is also negative) then comment is moderately subjective.

5.5.3 If (affective word is simple negative and adverb is positive) then comment is subjective.

5.5.4 If (affective word is simple positive and adverb is negative) then comment is moderately subjective.

5.5.5 If (emoticons is positive) then comment is subjective.

5.5.6 If (emoticons is negative) then comment is moderately subjective.

STEP 6: Evaluate comments with Rules:

Compare each comment with above rules. After comparisons we get result about which comment is subjective or moderately subjective and objective.

STEP 7: Combine result of each rule and all comments:

Combine each result according to output and on the basis of output we classify the relations between two users.

5.1 Attributes

- Number of affective words
- To measure how effective subjective words are on an average.
- Number of punctuation marks

- Number of repeated letters when letters are repeated
- Consecutively at least three times
- Number of interjections
- Number of social acronyms
- Number of emoticons
- Average rating of emoticons
- Number of reply

5.2 Tools

Nvivo10 and Ncapture tool is used as a tool. Ncapture tool is used for capture the facebook comments which is not private. Nvivo10 is the tool provides analysis of facebook comments. Nvivo10 is used for analyzing all forms of unstructured data and also Ncapture data show in Nvivo10.

VI. EXISTING WORK EXPERIMENTAL RESULTS

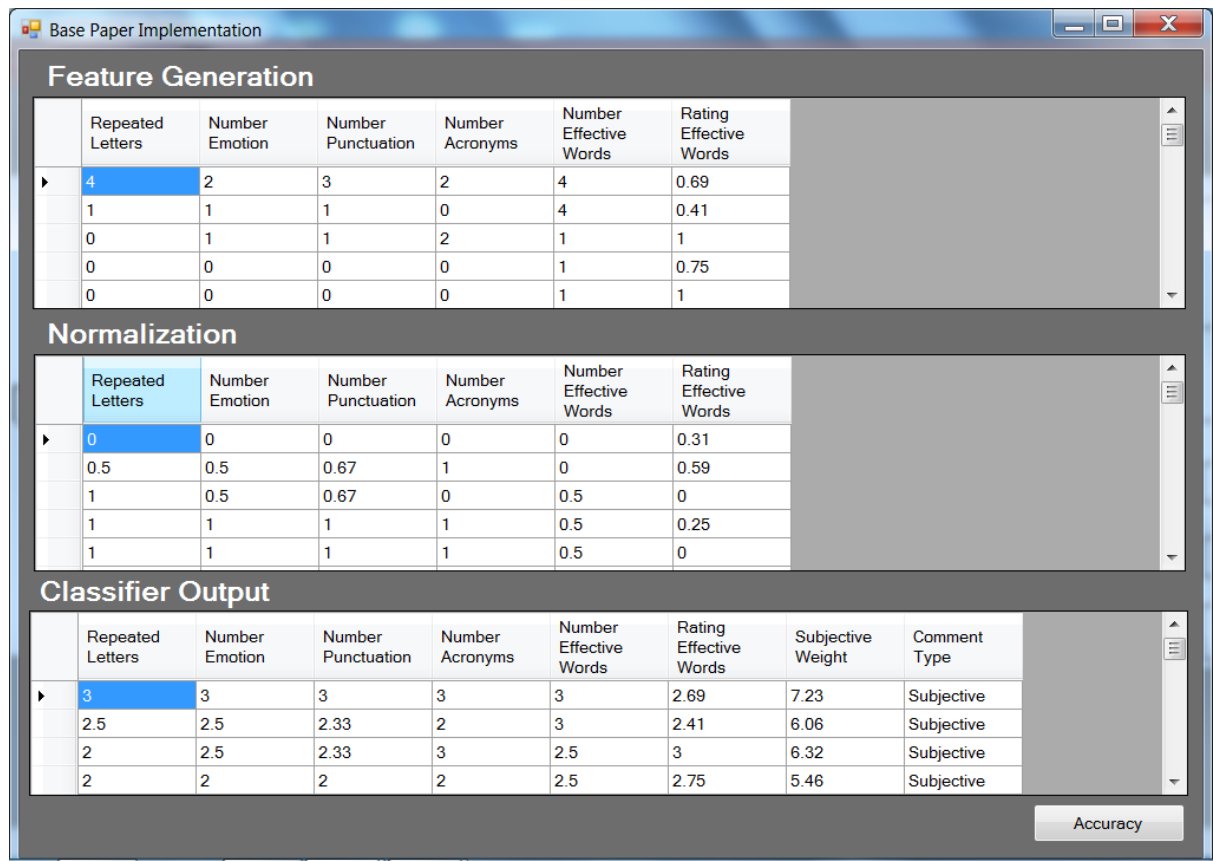


Figure2: Data Feature Generation and Normalization

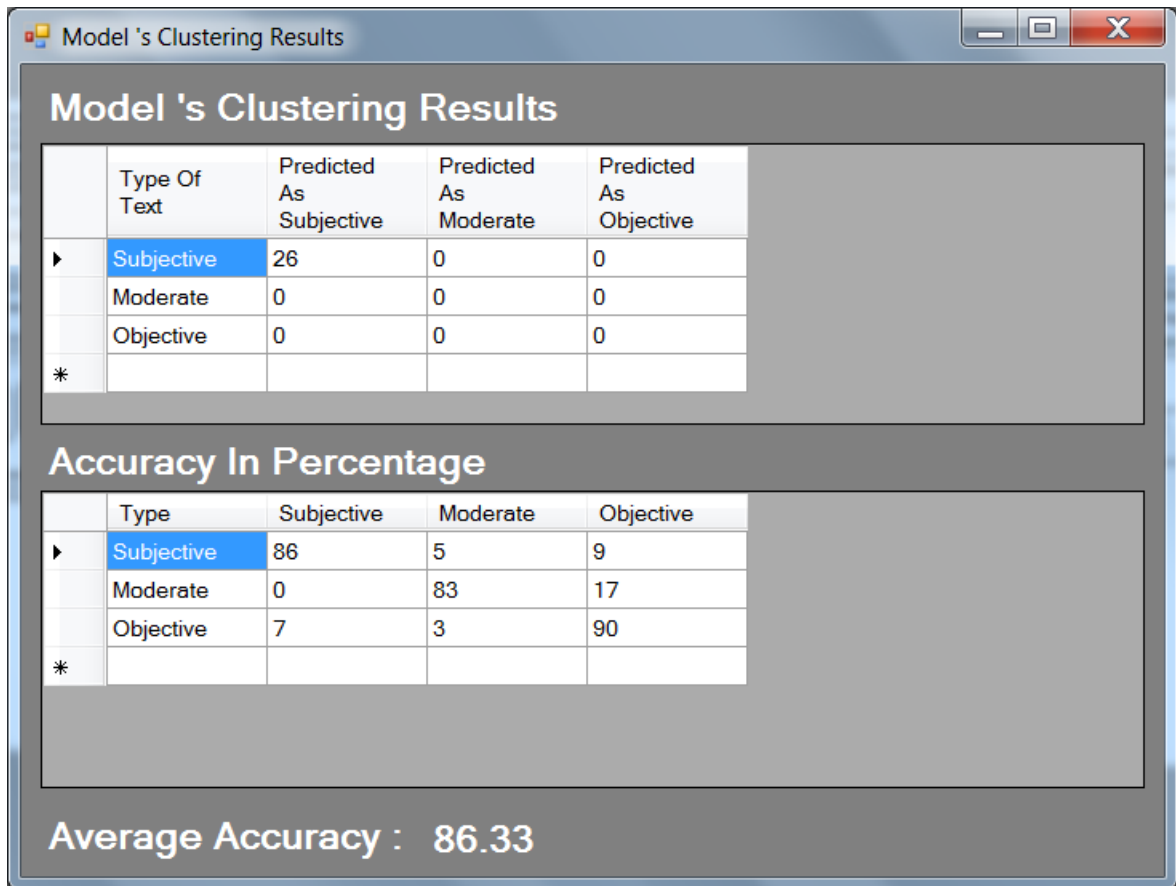


Figure 3: Existing result

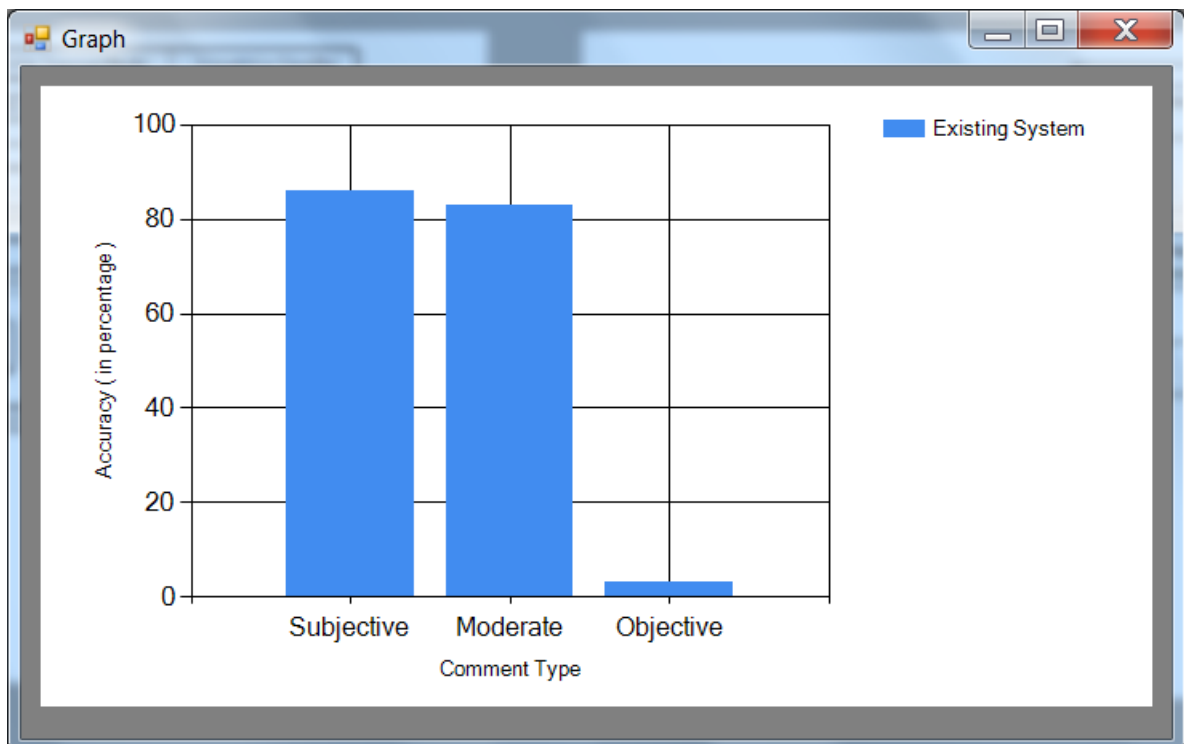


Figure 4: Existing graph

VII. PROPOSED WORK EXPERIMENTAL RESULTS

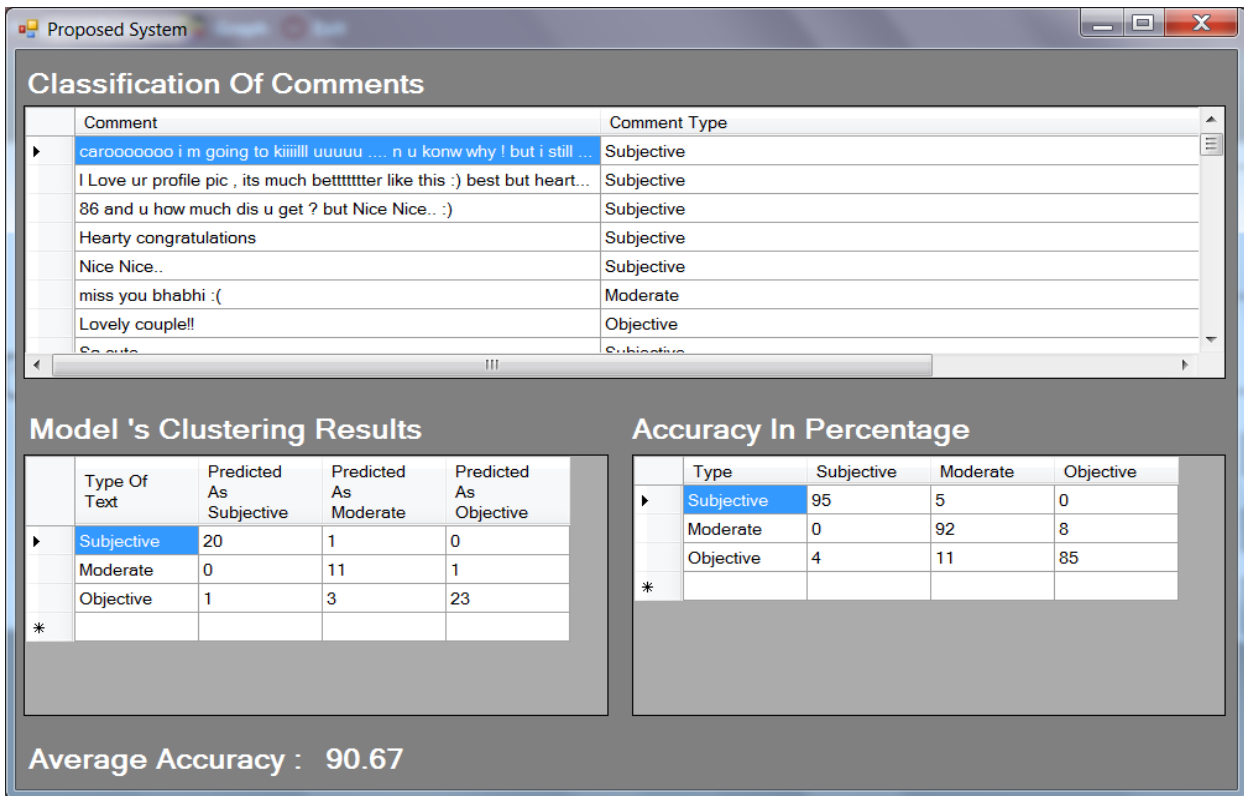


Figure 5: Proposed Result

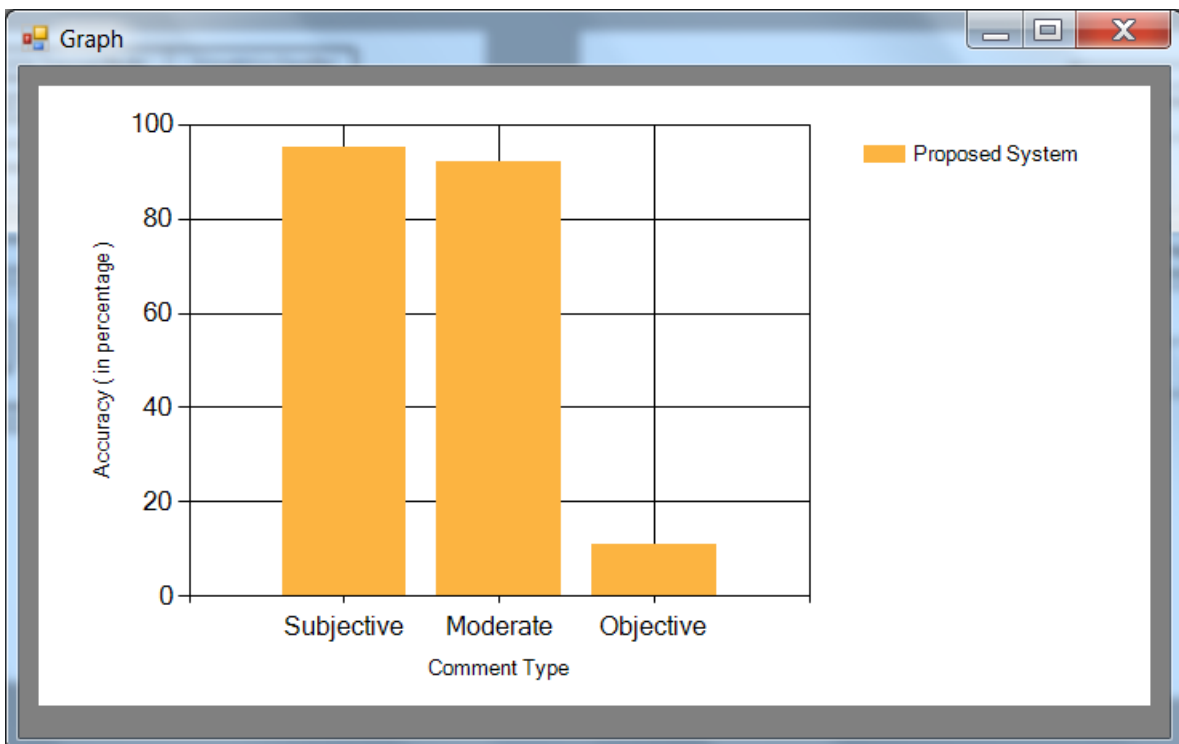


Figure 6: Proposed graph

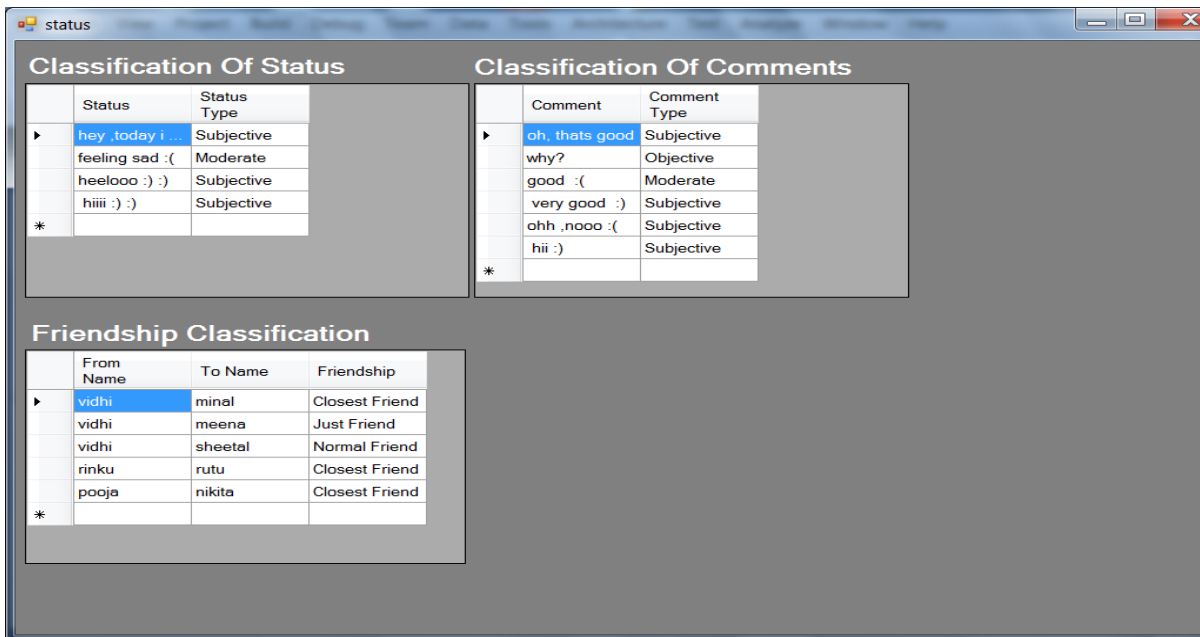


Figure 7: Classification of Status and Comments

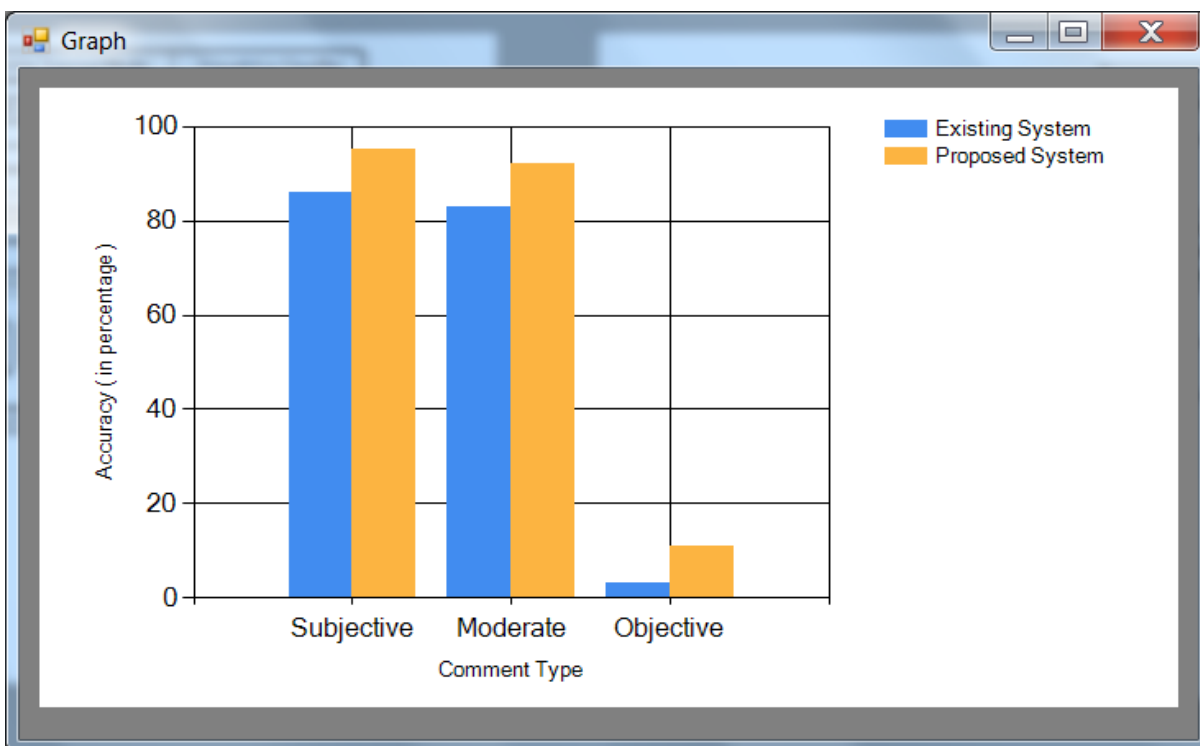


Figure 8: Comparison between Existing and Proposed Work

VIII. CONCLUSION

This paper discusses the various approaches to Emotion Analysis using emotion mining techniques. It presents different or new perspective for studying friendship relations and emotions' expression in online social networks where it deals with the specific nature of these sites and the

nature of the language used. The purpose was to identify whether the writer conveys his emotions and feelings in his writings. The processed data was then used to identify tie strength between two persons based on the subjectivity of the texts they share online. The first category consisted of objective or factual texts, the second contained moderately subjective texts suggesting some kind

of friendship between the users and the third consisted of subjective texts suggesting a close friendship between the two users.

REFERENCES

- [1] P. Ekman, "An argument for basic emotions," *Cognition & Emotion*, Vol. 6, No. 3, pp 169-200, 1992.
- [2] M. Thelwall, D. Wilkinson, S. Uppal, "Data mining emotion in social network communication: Gender differences in MySpace", In *Journal of the American Society for Information Science and Technology*, pp. 190-199, (2010).
- [3] A. Esuli and F. Sebastiani, "SentiWordNet: a publicly available lexical resource For opinion mining," in *Proceedings of the 5th International Conference on Language Resources and Evaluation*, pp. 417-422, Genoa, Italy, May 2006.
- [4] S. Aman, S. Szpakowicz, "Identifying Expressions of Emotion in Text", In *Proceedings of 10th International Conference on Text, Speech and Dialogue, Lecture Notes in Computer Science 4629*, pp. 196-205, 2007.
- [5] B. Pang, L. Lee, S. Vaithyanathan, "Thumbs up? Sentiment Classification Using Machine Learning Techniques", In *Proceedings of 2002 Conference on Empirical Methods in Natural Language Processing*, pp.79-86, 2002.
- [6] A. Neviarouskaya, H. Prendinger, and M. Ishizuka, "EmoHeart: Conveying Emotions in Second Life Based on Affect Sensing from Text", *Advances in Human-Computer Interaction*, vol. 2010, Article ID 209801, 13 pages.
- [7] H. Liu, H. Lieberman, and T. Selker, "A Model of Textual Affect Sensing Using Real World Knowledge," *Proc. Eighth Int'l Conf. Intelligent User Interfaces*, pp.369-374, 2003.
- [8] Yanghui Rao, Qing Li, Xudong Mao, Liu Wenyin, "Sentiment topic models for social emotion mining", *Information Sciences* 266 (2014) 90-100, 2014 Published by ELSEVIER Ltd.
- [9] Sanjeev Dhawan, Kulvinder Singh, Vandana Khanchi, "A Framework for Polarity Classification and Emotion Mining from Text", *International Journal Of Engineering And Computer Science* ISSN:2319-7242
- Volume - 3 Issue -8 August (2014), Page No. 7431-7436.
- [10] Sanjeev Dhawan, Kulvinder Singh, Deepika Sehrawat, (2014), "Emotion Mining Technique in Social Networking Sites", *International Journal of Information & Computation Technology*. ISSN 0974-2239 Volume 4, Number 12 pp. 1145-1153.
- [11] Nilesh M Shelke, Shriniwas deshpande PhD, vilas thakre PhD, "Approaches towards Emotion Extraction from Text", *National Conference on Innovative Paradigms in Engineering & Technology (NCIPET-2013) Proceedings published by International Journal of Computer Applications@ (IJCA)*.
- [12] C. Ma, H. Prendinger, and M. Ishizuka (2005), "Emotion Estimation and Reasoning Based on Affective Textual Interaction," In *Proc. First Int'l Conf. Affective Computing and Intelligent Interaction (ACII '05)*, pp. 622-628.
- [13] C. Strapparava and A. Valitutti, "WordNet-Affect: an affective extension of WordNet," in *Proceeding of 4th International Conference on Language Resources and Evaluation (LREC04)* pp. 1083-1086, Lisbon, Portugal, May 2004.