I J C T A, 9(20), 2016, pp. 185-191

Dictionary lookup approach for emotion mining from rhymes in English language using Apriori algorithm

Jeenu Gupta* and Williamjeet Singh**

ABSTRACT

These days emotion mining is gaining popularity. Emotions are the main components of individual's day-to-day life. People can choose rhymes according to the emotions in which they have interest. In this paper, a new functionality has been proposed for characterizing emotions in rhymes provided as corpus and then distinguishing two rhymes for different emotions. The primary goal is to extract the emotions from the rhymes in text form in the English language. The main interest is to detect the emotion category to which a particular rhyme belongs. The functionality includes the database collection, root words dictionary and the text mining techniques. It considers the online rhymes and text available online only in the English language. The technique used is Apriori algorithm. Experiments show high accuracy for the system proposed.

Keywords: Dictionary Lookup, Emotion Mining, Apriori Algorithm

1. INTRODUCTION

Mining emotions have fascinated the world, as is clear from the tremendous assemblage of examination performed in business related to feeling in fields of science, brain, phonetics and correspondence. Feelings can be judged in many ways like using expressions, body language, compositions, activities etc [22]. This paper addresses the task of mining the emotions by automatically detecting the feelings from content**.**

Emotions are the result of our inside conscience and it performs a significant role in basic leadership and subjective connection forms. Emotion mining is a recent research area and focuses on feelings of a human [16].

Artificial intelligence should choose automatically without or with scanty human interference how to respond to alter their relations and surroundings [10].

1.1. Emotion Mining

Recognizing feeling in the content is the only way to make machines react to feelings or emotions. The human feeling can be detected from appearance, signals, discourse and works [4]. Only restricted work has been done toward determination of feeling in content.

Mining social feelings from the content or the corpus provided are doled out by social users with feeling like happiness, warmness and amusement [6]. The corpus in the context of feelings is classified which in turn help the users for related choice on the web. Mining regular trends are obvious amongst the most important ideas in information mining. It ought to be the start of any information mining as it gives the information what emotion mining is all about. The corpus content based on mining permits us to induce various contingent probabilities for concealed records [1] [7]. There are diverse strategies used to manage

* Mtech scholar Department of Computer Engineering Punjabi University, Patiala, Punjab, India, *Email: guptajeenu615@gmail.com*

** Assistant Professor Department of Computer Engineering Punjabi University, Patiala, Punjab, India, *Email: williamjeet@gmail.com*

the content. Earlier the concentration used to be on titles to mine the emotions. The results displayed treats terms independently and were unable to find the logical data inside the corpus. Previous models were unable to differentiate between the general and the emotion terms. But the recent model can find the actual emotion under the titles and can connect the associations between the social feelings and emotions writings [9] [11].

Dialect is a capable instrument to impart and pass on data [3]. In addition, it is a way to express feeling. For quite some time Natural Language Processing (NLP) strategies have been connected to consequently distinguish the emotional data content in the corpus. This work is a try to apply NLP procedures to recognize feelings communicated in content [14].

Today, Artificial Intelligence (AI) has centered expanding endeavors on creating frameworks that consolidate feeling. Emotion mining is an important procedure that is demonstrated in AI frameworks [21]. These incorporate recognition, thinking, learning and normal dialect handling [13]. Emotion mining is the most important aspect of creating sentimental interfaces – ones that can understand inputs, give fitting reactions and encourage online correspondence [8]. Such interfaces can enormously enhance client involvement in Computer-Mediated Communication (CMC) and Human-Computer Interaction (HCI) [2].

Identifying the feeling and investigating strategies are valuable in numerous applications within mental premises. For instance, they can be effectively connected to take in user interests and their compositions.

The rest of the paper is organized as follows. Proposed Emotion Mining algorithms are explained in section II. Experimental results are presented in section III. Concluding remarks are given in section IV.

2. PROPOSED ALGORITHM

2.1. Dictionary Lookup Approach

In the proposed system, an improved Apriori Algorithm along with dictionary lookup technique is used to determine the emotions in the given rhymes. In the proposed system, rhymes are used as a corpus. Rhymes are scanned word by word to detect the emotional words and the detected emotional words are compared with the dictionary which includes emotional category and accordingly emotional category tags are provided. Emotion category included are: “sad”, “happy”, “anger”, “disgust”, “fear”, “boredom”, “surprise”, “love”. The presence of one or more emotion words of a particular category in the given corpus helps to determine the overall emotion of the rhyme. In this approach to find the emotions in a rhyme prior knowledge of emotional words are required to classify the words into a particular emotion category, which will be used as a dictionary.

The proposed system works in two phases. In the first phase, the emotion words are detected from the given corpus and then enhanced Apriori algorithm is used to mine the emotions by making a comparison with emotion words, finally emotion category tags are provided to rhymes. In the second phase, new emotion words can be identified with the help of root words and these words are again used to mine the emotions from the rhymes.

To determine the particular emotion from rhyme, the proposed system counts the number of emotion words of each category in the given corpus and then assigns the category according to the largest number of words found in the rhyme. A rhyme in which emotion word of any type is not found is categorized as “sentence having no emotion”. An example of the dictionary or the database used at the back end is shown in Table 1.

With the help of rhymes available online and other sources, the dictionary contains at least 100 words from each category mentioned above. Table 1 shows few words from each category.

Table 1
Synonyms related to different categories of emotions.

S. No.	*Emotion*	*Related Words*
1	Happiness	blessed, joy, enjoy, blissful, cheerful, chirpy
2	Sad	unhappiness, sorrow, depression, anguish, dejection, regret
3	Anger	irritation, annoyed, crossness, rage, fury, wrath
4	Disgust	nauseated, fed up, repelled, abhorrence, aversion, loathing, repulsion
5	Surprise	amazing, abruptness, amazement, astonishment, shock
6	Fear	Terror, fright, fearfulness, horror, Alarm, apprehension
7	Love	adoration, liking, adulation, affection, allegiance, amity
8	Boredom	ennui, apathy, weariness, unconcernacidize, monotony

2.2. Dictionary Lookup Framework

In this section, we will first present the architecture of the proposed system shown in figure 1 and then elaborate each step followed by procedure.

The framework shown in Figure 1 is structured into six steps:

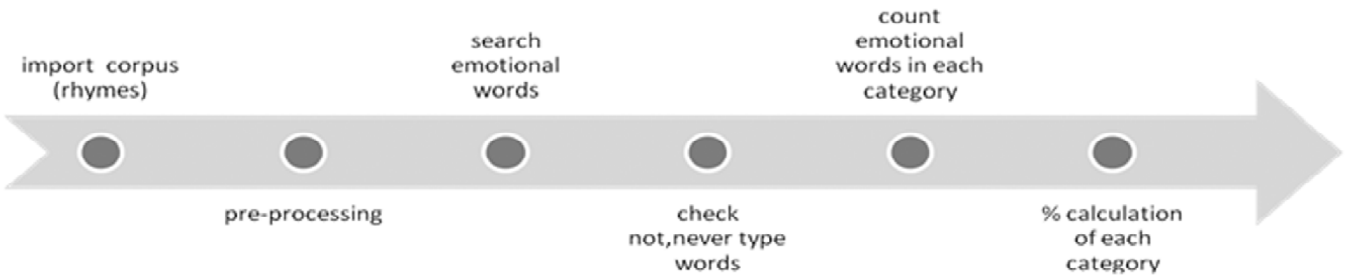

Figure 1: Framework architecture

1. *Import Corpus*: This step is applied to import the rhymes for which we want to find emotions. The imported data will be displayed in the text box.
2. *Pre-processing*: This step is applied to the imported data. It involves removing redundant attributes by discarding them.
3. *Emotional words search*: In this search imported text will be compared word by word with the database and if any word in rhyme matches database then it will check the corresponding emotion category. The system will display the matched emotional words along with corresponding emotional category in the imported rhyme.
4. *Check not, never words*: This step focus on the checking “not, never” type words before emotional words. Such words may change the meaning of emotion if added before it. For example not happy becomes sad.
5. *Counting of emotional words in each category*: This step is applied on the rhyme provided by the user after undergoing above 4 steps, the emotional words from each category are counted and displayed on the screen.
6. *% calculation of each category*: Output of above step is taken as input for this step to calculate the percentage of each category, so that will have an idea of emotion to which a particular rhyme belongs to.

2.2.1. Procedure (Mine_Emotion)

1. BEGIN;
2. Input the text paragraph from which emotions is to be determined.
3. Tokenize the input obtained from the previous step;
4. Initialize all the emotion counters to zero;
5. Initialize the Emotion_List to empty;
6. for each tokenized word do
7. If (word is NOT in Emotion_List) then
8. Check the word into emotion database for every emotion category;
9. If (word found) then
10. Increase Corresponding Emotion counter
11. Add the word into Emotion_List
12. Else
13. Check the word into root database for every emotion category;
14. If (word found) then
15. Increase Corresponding Emotion counter
16. Add the word into Emotion_List
17. End
18. End
19. End
20. Calculate the percentage of emotion for each category;
21. Display the results to the user;
22. Generate the graph for emotions for each category;
23. END;

The output of the model was the calculation done in step 5 for calculating the total emotion with respect to per category. Step 6 is the application of calculating the percentage of emotion classification. In this way, we will get the percentage of each emotion category present in particular rhyme. A particular rhyme can be a mixture of two or three feelings.

3. EXPERIMENT AND RESULT

The test set for this evaluation experiment of rhymes is randomly selected from the internet. ASP .net software platform is use to perform the experiment. The PC for experiment is equipped with an Intel i3 1.70GHz Personal laptop and 4GB memory. The proposed scheme is tested using 50 ordinarily rhymes extracted online.

The evaluation of the model was done at two different levels. First, we tested how accurate the Apriori algorithm was in recognizing different classes of rhymes with respect to emotions. Second, we tested the accuracy of the algorithm based solely on the emotions of the rhyme, the relationship's strength between two classes of emotions in a particular rhyme.

Table 2 gives the comparison between the existing and proposed system. Several poems were taken and accuracy was tested. On manual check we found 8 emotional words in a particular rhyme. The existing system was able to detect only 6 words whereas; proposed system was capable of finding 7 emotional words. Consequently, the accuracy has been increased.

Table 2
Comparison of the existing and proposed systems

Rhyme	*Total Emotion Words*	*Emotion Words Extracted by Existing system*	*Accuracy of Existing system*	*Emotion words extracted by proposed system*	*Accuracy of the proposed system*
Rhyme1	8	6	75%	7	88%
Rhyme2	10	7	70%	9	90%
Rhyme3	15	11	73%	13	86%
Rhyme4	20	14	70%	18	90%

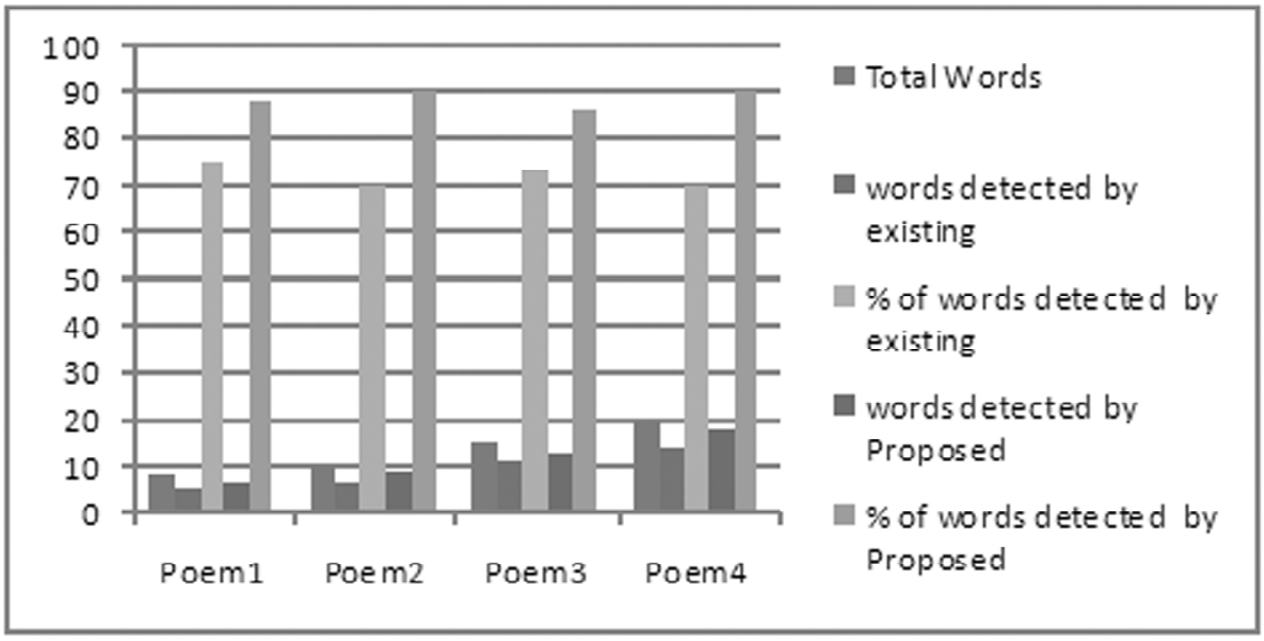

Figure 2: Accuracy comparison between existing and proposed systems

4. CONCLUSION

This paper discusses a novel emotion mining technique for rhymes provided by the user. It presents a new outlook for studying rhymes and emotions expression where it deals with the specific language used. The purpose of this paper was to identify emotions and feelings of a poet in his writings. The processed data was then used to identify percentage strength between two emotions. The main challenge in the current algorithm is the usage of new words that are not contained in the proposed dictionary; in this perspective, we can develop new root table that will cover common pre and post words for each emotion. Emotions were grouped into eight categories.

We propose as future work to first increase the number of emotion category as we have limited to eight categories, secondly database i.e. dictionary can be increased. There is no end to the words, so words under each category can be increased. Finally, one of the challenges was to deal with the variety in the language used in online rhymes consequently, it is important to develop new and creative ways to learn and cope with the changes of the language used in rhymes.

REFERENCES

[1] Yanghui Rao, Qing Li, Xudong Mao, Liu Wenyin, (2014), " Sentiment topic models for social emotion mining", an international conference on education and social sciences (intcess14),"ed:Elsevier,2014, pp. 90-100.

[2] R. Holden and J. Rubery, (2013)"Emotion," in Oxford Dictionaries, J. Pearsall, Ed., ed: Oxford University Press, 2013.

[3] Shenghua Bao, Shengliang Xu, Li Zhang, Rong Yan, Zhong Su, Dingyi Han, and Yong Yu (2012) "Mining Social Emotions from Affective Text "published In : IEEE transactions on knowledge and data engineering, vol. 24, no. 9.

[4] S. Bao, S. Xu, L. Zhang, R. Yan, Z. Su, D. Han, Y. Yu, (2012) Mining social emotions from affective text, IEEE Trans. Knowledge. Data Eng. 24 1658–1670.

[5] David Garcia, Frank Schweitzer(2011), Chair of Systems Design, ETH Zurich, Kreuzplatz, "Emotions in Product Reviews – Empirics and Models", 2011 IEEE International Conference on Privacy, Security, Risk, and Trust, and IEEE International Conference on Social Computing, Boston, MA, IEEE Publications, Oct 2011, pp. 483-488.

[6] Mohamed Yassine, Hazem Hajj (2010), "A Framework for Emotion Mining from Text in Online Social Networks", IEEE International Conference on Data Mining Workshops, Sydney, NSW, IEEE publications, Dec 2010, pp. 1136-1143.

[7] M. Thelwall, D. Wilkinson and S. Uppal.(2010) "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.

[8] A. Neviarouskaya, H. Prendinger, and M. Ishizuka(2010), "EmoHeart: Conveying Emotions in Second Life Based on Affect Sensing from Text," Advances in Human-Computer Interaction, vol. 2010, Article ID 209801, 13 pages.

[9] S. Bao, S. Xu, L. Zhang, R. Yan, Z. Su, D.Han, Y. Yu,(2009) Joint emotion-topic modeling for social affective text mining, In: Proc. 9th IEEE International Conference on Data Mining (ICDM), 2009, pp. 699–704.

[10] B. H. C. Cheng, R. De Lemos, H. Giese, P. Inverardi, and J. Magee(2009), "Software Engineering for Self-Adaptive Systems," Dagstuhl seminar 10431, ed: Springer, 2009, pp. 48-70.

[11] A. J. Gill, D. Gergle, R.M. French, and J. Oberlander(2008). Emotion rating from short blog texts. In Proc. of CHI'08 on Human factors in computing systems, pages 1121–1124.

[12] I. Titov and R. McDonald(2008). A joint model of text and aspect ratings for sentiment summarization. In Proc. of ACL'08.

[13] R. Tokuhisa, K. Inui, and Y. Matsumoto(2008). Emotion classification using massive examples extracted from the web. In Proc. of Coling'08, pages 881–888.

[14] C. Strapparava and R. Mihalcea(2008). Learning to identify emotions in text. In Proc. of SAC'08, pages 1556–1560.

[15] R. Cai, C. Zhang, C. Wang, L. Zhang, and W.-Y. Ma(2007)., "Musicsense: Contextual Music Recommendation Using Emotional Allocation," Proc. 15th Int'l Conf. Multimedia, pp. 553-556.

[16] C. Yang, K. H.-Y. Lin, and H.-H. Chen(2007). Emotion classification using web blog corpora. In Proc. of WI'07, pages 275–278.

[17] F. Chaumartin. Upar7(2007): A knowledge-based system for headline sentiment tagging. In Proceedings of SemEval-2007, pp. 422-425.

[18] X. Wang and A. McCallum.(2006) Topic over time: A nonmarkov continuous-time model of topical trends. In Proc.of KDD'06, pages 424–433

[19] A. Esuli and F. Sebastiani, "Sentiwordnet(2006) A Pub-Licly Available Lexical Resource for Opinion Mining," Proc. Fifth Int'l Conf. Language Resources and Evaluation (LREC '06).

[20] A. Esuli and F. Sebastiani. Sentiwordnet (2006) : A publicly available lexical resource for opinion mining. In Proc. Of LREC'06.

[21] C.O. Alm, D. Roth, and R. Sproat(2005), "Emotions from Text: Machine Learning for Text-Based Emotion Prediction," Proc. Joint Conf. Human Language Technology and Empirical Methods in Natural Language Processing (HLT/EMNLP '05), pp. 579-586.

[22] 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.

[23] T. Griffiths and M. Steyvers, (2004). "Finding Scientific Topics," Proc. Nat'l Academy of Sciences USA, vol. 101, pp. 5228-5235, 2004.

[24] T. Wilson, J. Wiebe, and R. Hwa. (2004). Just how mad are you? Finding strong and weak opinion clauses. In Proceedings of AAAI-04, 21st Conference of the American Association for Artificial Intelligence, pp. 761–769.

[25] Hugo Liu, Henry Lieberman, Ted Selker. (2003). A Model of Textual Affect Sensing using Real-World Knowledge.

Proceedings of the 2003 International Conference on Intelligent User Interfaces, IUI 2003, January 12-15, 2003 ACM 2003, pp. 125-132.

[26] D. M. Blei, A. Y. Ng, and M. I. Jordan(2003). Latent Dirichlet allocation. Journal of Machine Learning Research, (3):993–1022, 2003.

[27] H. Liu, H. Lieberman, and T. Selker(2003)., "A Model of Textual Affect Sensing Using Real-World Knowledge," Proc. Eighth Int'l Conf. Intelligent User Interfaces, pp. 369-374.

[28] M. Corney, O. de Vel, A. Anderson, and G. Mohay(2002)., "Genderpreferential text mining of e-mail discourse," in 18th Annual Computer Security Applications Conference, Vol. 13, pp. 21-27.

[29] P. Ekman, (1992), "An Argument for Basic Emotions", International Journal of Cognition and Emotion, Vol. 6(3), published by Lawrence Associates Ltd, US, Jan 1992, pp. 169-200.