May-June 2020
ISSN: 0193-4120 Page No. 3653-3657

User Review Analysis in Social Forums Based on Sentiment Cateloging

[1]Jeyakumar Durairaj, [2]Balamanigandan R, [3]Selvanayaki S, [4]S. John Justin Thangaraj

[1,3]Assistant Professor, [2,4]Associate Professor
[1]Rajalakshmi Institute of Technology, Chennai, Tamil Nadu, India
[3]Vel Tech Multi Tech Dr.Rengarajan Dr.Sakunthala Engineering College, Chennai
[2,4]Saveetha School of Engineering, Saveetha Institute of Medical and Technical Sciences, Chennai
[2]bala16385@gmail.com

Article Info
Volume 83
Page Number: 3653-3657
Publication Issue:
May-June 2020

Abstract
Sentiment cataloging is a essential chore in sentiment identification and analysis, and its aim is to categories the sentiment or point of view found in the social webpages with the given feedback or command section text. The process of sentiment identification and classification we proposed not only follows the practical methods in subject-based text classification but also involves the sentiment analysis. Those bags of words are used in information retrieval. Machine learning approaches as well will not perform an advanced analysis on sentiment classification as on outdated theme based classification and categorization. We propose a simple yet efficient model, called Globalized User Review Analysis (GURA) by using the property of feedback sense with the sentiment classification of basic two opposite classes of labels, we proposed an algorithm with the data expansion technique first by creating sample sentiment toggled comments. The unique and transferred comments are then constructed in accordance with the one-to-one correlation. Thereafter, we enhance the dual training (DT) algorithm and a dual forecasting (DF) algorithm separately, in order to make use of the existing original samples and the stored switched samples in pairs. This model helps in systematic training and an automated statistical classifier that can be achieved with the estimated predictions. The overall polarity of the reviews can be viewed in the sentiment graph.

Keywords: *Sentiment Analysis, User Review, Dual Training, Dual Forecasting, Sentiment Cateloging*

Article History
Article Received:*19 August2019*
Revised: *27 November 2019*
Accepted: *29 January 2019*
Publication: *12 May 2020*

1. Introduction

This invention is related to User Review Analysis in Social Forums based on Sentiments. We propose a Globalized User Review Analysis (GURA) model in order to address the issues for Sentiment Classification using two methods namely Dual training method and Dual Prediction method [1]. In the initial phase, the original stored training samples are estimated reversed in terms of their synonyms to their opposite counterparts by referring them as the original and the reversed sample training sets respectively. In prediction algorithm it measures on how much positive the original test sample is, and also considers how much the negative reversed test review is and vice versa [2]. By applying this algorithm neutral reviews can be found accurately.

A customer will not be able to post any kind of unusual/vulgar words in his review which is detected with the help of Bag of Words (BOW). Those kind of detected words will be shown to other viewers in the form of asterisk (*) symbols. The overall polarity of the user reviews can be seen by the sentiment graph which will be generated automatically based on the reviews.

2. Background Study

Sentiment Analyzing is a very challenging task which involves the concepts of NLP, text and web mining

May-June 2020
ISSN: 0193-4120 Page No. 3653-3657

and machine learning. For the text based sentiment analysis, it requires large quantities of linguistic features [5]. Many of the existing method are relying on the supervised machine learning approaches that are well trained from the labeled corpora where every document and sentences have been labeled as positive or negative prior to the training.

The disadvantage is that the ML approaches did not perform as well on sentiment classification as on traditional title based classification.[4] Fujita et al., have proposed a combination WSD method which consisting of automated label data oriented expansion and semi supervised learning methods. These two methods are both effective but in this; the training data set is one of the severest problems.

Xia et al., focused on the polarity shift problem and proposed an approach called DTDP, which addressed it [7]. The basic idea is to first generate artificial samples that are polarity-opposite to the original sample by polarity reversion. The disadvantage is that it does not maintain integrity and it takes a long time for polarity reversion. It can able to detect only explicit expression of statements and it is difficult to find whether the document is review or not [9].

Globalized User Review Analysis (GURA)

We have proposed from the derivation of the existing approach to another efficient model, named Globalized User Review Analysis (GURA). Sentiment identification and classification have two opposite classes of labels, Initially we have proposed a data based expansion method by making use of sentiment switched comments. Then the stored original and estimated reversed comments are constructed in a peer-to-peer correspondence. From that point we proposed a dual training (DT) approach and a dual forecasting (DF) approach in a single algorithm for making utilization of the given and switched samples in the pairs for the given training samples with a statistical classifier and made expectations.

In Dual Training, the classifier method is studied by augmenting a grouping of probability of the stored unique and estimated switched exercise data set. In Dual Forecasting, forecasts are estimated by bearing in mind the dual sides of the given input assessment. That is, we have measured not just how much optimistic or undesirable the actual assessment is, additionally how negative or positive reviews the switched review is. Then the algorithm facilitates our work from extremity (positive versus negative) classification into 3-classes of classification (positive vs. negative vs. neutral) of sentiment, by taking the considerations from the neutral reviews in both dual training approach and dual forecasting method [10].

In order to reduce the systems dependency on an external dictionary of antonym verification and validation, we at last build up a approach for making a new pseudo antonym vocabulary [11]. The self-styled antonym glossary is designed with the features of language-independency and domain- versatility. This makes our model much conceivable to be connected with a wide range of real time applications.

3. Detailed Description

The proposed GURA model is very effective and maintains integrity over the various levels of Users. We extend the existing Sentiment Analysis algorithm by including neutral sentiment keywords and sample set for training. We have also proposed a method for assembling a self-styled antonym wordlist that gives better performance than using WordNet antonym dictionary.

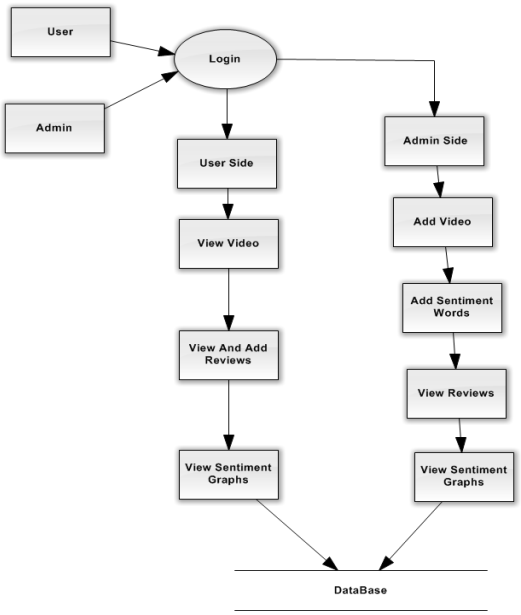

Figure 1: Data Flow Diagram for Sentiment Analysis

System Architecture

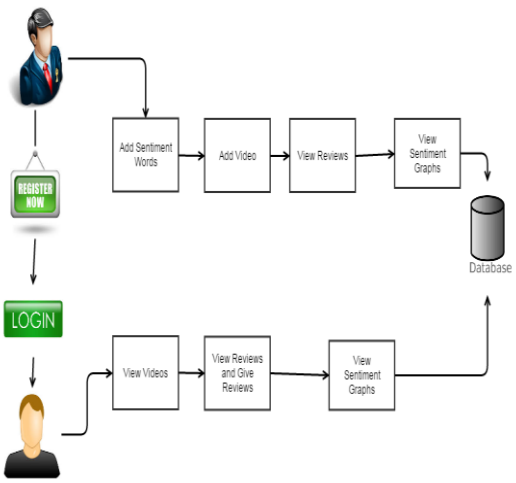

Figure 2: Architecture Diagram

The architecture for "Globalized user review analysis based on sentiments " is given above. There are two

May-June 2020
ISSN: 0193-4120 Page No. 3653-3657

users in this application. One will be the admin and another will be the customer. Both the admin and the user should undergo the registration process and they both will be given separate login Id's

Dual Training (DT)

In this we have to train original words and its reversed words. Let us consider x to be the original sample and $\bar{x}$ to be the reversed sample. The output of the original sample y lies between (0, 1) i.e., y € (0,1) and the output of the reversed sample $\bar{y}$ is 1-(original sample). i.e., 1-y. Dual training algorithm is explained below with an example:

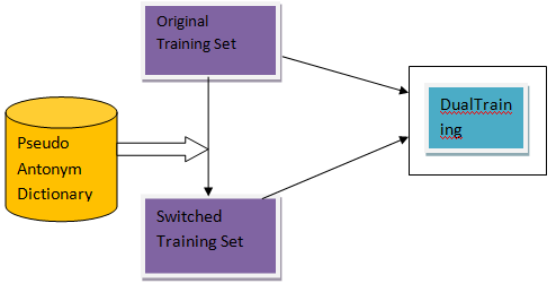

Figure 3: Dual Training

Original

Don't like this product. This is not fashionable (NEGATIVE)

Switched

I liked this product. It looks very trendy (POSITIVE)

Step 1: The first step is to detect the scope of negation. E.g. lie.

Step 2: Except the scope of negation all other words should be switched in accordance with the meaning found in dictionary. i.e., not fashionable to trendy, boring to interesting. The word like should not be reversed because it is the scope of negation.

Step 3: Remove negative words from the user input like no, not, don't.

Step 4: Finally the class label should also be switched inversely. i.e., from positive to negative or from negative to positive.

If the original sample x is positive, the standard probability with respect to x becomes log h(x). In dual training the probability becomes log [h(x) (1-h ($\bar{x}$)] i.e., it not only considers how x is positive but also how xl is negative. If the original sample x is negative, the standard probability with respect to x is log (1-h(x)). In dual training the probability becomes log [(1-h(x)) h ($\bar{x}$)] i.e., it not only considers how x is negative but also how $\bar{x}$ is positive.

Dual Forecasting (DF)

The words which are trained during training process will be used in future. Consider,

P(- / x) – subsequentpossibility of x.

P(./ $\bar{x}$) – subsequentpossibility of $\bar{x}$.

In this method for x we create $\bar{x}$. The objective of this work is not only to guess the type of $\bar{x}$, but with the use of $\bar{x}$ to assist the estimate of x. To measure the sample test review x to be positive, we have not only considered how much constructive the original test comment (p(+ / x)) is, but also how much negative the reversed test reviews is (p(- / $\bar{x}$)).

To measure a test review x to be negative, we not only consider how negative the original test review (p (+ / x)) is, but also how positive the reversed test reviews is (p (- / $\bar{x}$)). When dual forecasting is not confident, then go for original assumption

$P_f(y/x) = \{p_d(y/x, \bar{x})$

$\{p_o(y/x)$ if $\Delta p >= t$

Here it is called as threshold parameter and we set t closer to zero. i.e., the prediction of the possible value thru a sophisticated following possibility is elected as the concluding outcome.

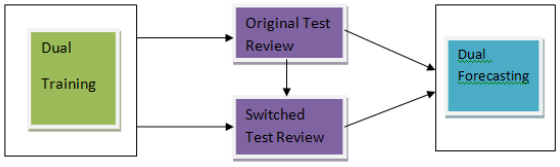

Figure 4: Dual Forecasting

Classification

There may still be happen many other neutral assessments. The algorithm we discussed in literature review above may not have the capability to organize neutral evaluations. So we are going to outspread the frame work for 3-class (constructive-neutral-destructive) sentimentality classification. There exist two situations for neutral reviews:

1. Objective of the texts without expressing the sentiment explicitly
2. Texts that articulating assorted or conflicting sentiment.

Hence to overcome this problem, we have reversed the reviews text but we will possess its type label as it was.

Table 1: Sample Analysis

	REVIEWEDCOMMENT	TYPE
Sample Stored Comment	This boy is very tall. But not looking much handsome	Neutral
Switched estimated Comment	This boy is not so tall. But look somewhat handsome	Neutral

In the second stage, we have examined for each and every original test sample x collected from the dataset, we then created inverted one $\bar{x}$. In order to

May-June 2020
ISSN: 0193-4120 Page No. 3653-3657

take them into account for the neutral evaluations, we finally fill in the preceding likelihood algorithm in the following equation.

$P(+ / x, \overline{x}) = (1-\alpha) . p_d(+ / x) + \alpha . p (- / \overline{x})$,
$P(- / x, \overline{x}) = (1-\alpha) . p_d(- / x) + \alpha . p (+ / \overline{x})$,
$P(* /x, \overline{x}) = (1-\alpha) . p(* / x) + \alpha . p(* / \overline{x})$.

where {+,-,*} denote the classes of labels and in this case $p(+ / x, \overline{x}) + p(- / x , \overline{x}) + p(* / x, \overline{x}) = 1$.

In this stage we know that, whenever we extent how much optimistic or undesirable a test review is, we have not only considered how confident or harmful the actual comment is, but also how much harmful or helpful the reversed review is. We have introduced the 3-class corn classification scheme, then we are trying to quantity how unbiased a test examination is, we have not only considered how impartial the inventive analysis is, but also how impartial the overturned sample comment is. The performance analysis of accuracy in sentiment classification is given in the figure 5.

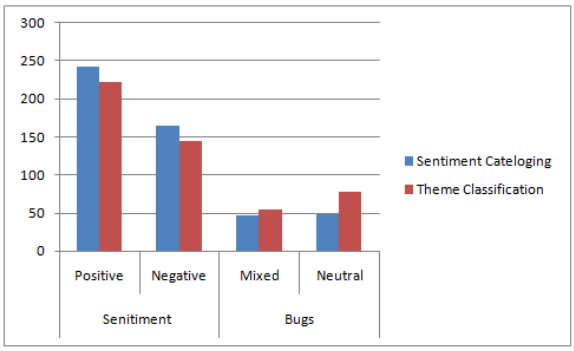

Figure 5: Performance Analysis

4. Conclusion

We conclude this paper with a proposed algorithm of text sentiment classification, called Globalized User Sentiment Analysis (GURA) by using the property of the text based sentimentality cataloging which has two contradictory extremes of class labels (i.e., constructive and damaging), and a data based emotional expansion method by forming sentiment toggled comments. We have estimated impact of the original user comment is, also how these adverse and helpful reviews may have impact on the other forms of, that the reversed review is. Further we have expanded our methodology from binary classification to the cataloging into 3-class (optimistic vs. undesirable vs. impartial) sentiment cataloging, by considering the original neutral comments in both. In order to diminish the impact of craving on an outer vocabulary for verification and validation of antonym, we developed a glossary method for building a pseudo-antonym vocabulary. In this model, we have also included a sentiment graph feature by which the overall cataloguing of the user comments and feedbacks can be viewed using the sentiment graph.

References

[1] Z. Hai, K. Chang, J. Kim, and C.C. Yang, "Identifying Features in Opinion Mining via Intrinsic and Extrinsic Domain Relevance, "IEEE transactions on knowledge and data engineering (TKDE), vol. 26,no.3, pp. 447 – 462, 2014.

[2] R. Xia, T. Wang, X. Hu, S. Li, and C. Zong, "Dual Training and Dual Forecasting for Polarity Classification, "Proceedings of the Annual Meeting of the Association for Computational Lingustics (ACL), pp. 521 – 525,2013

[3] X. Yu, Y. Liu, X. Huang, anfd A. An, "Mining online reviews for predicting sales performance: A case study in the movie domain, "IEEE transactions on knowledge and Data Engineering (TKDE), vol. 24, no. 4, pp. 720 – 734,2012.

[4] C. Lin, Y. He, R. Everson, and S. Rugar, "weakly supervised joint sentiment-topic detection from text, "IEEE Transaction on knowledge and Data Engineering (TKDE), vol. 24, no. 6, pp. 1134-1145,2012.

[5] B. IU, "Sentiment Analysis and Opinion Mining, " synthesis Lectures on Human Language Technologies, Morgan & Claypool, pp. vol. 5,no. 1,pp. 1-165,2012.

[6] S. Orimaye, S. Alhashmi, and E. Seiw, "Buy it – don't buy it: sentiment classification on Amazon reviews using sentence polarity shift, "Proceeding of the Pacific Rim International Conference on Artificial Intelligence (PRICAI),2012.

[7] ThangarajS.John Justin, M.RajeshKhanna, R. Balamanigandan "Failure Node Detection And Recovery In Wireless Sensor Networks", International Journal of Innovative Technology and Exploring Engineering (IJITEE) ISSN: 2278-3075, Volume-8, Issue- 12, October 2019.

[8] S. Fujita and A. Fujino, "Word sense disambiguation by combining labeled data expansion and semi-supervised learning method, "proceedings of the International Joint Conference on Natural Language Proceeding (IJCNLP), pp. 676 – 685,2011.

[9] Mukund R. , John Justin Thangaraj S , S P. Chokkalingam, "Semi-Supervised Learning Using Procreative Modelling Techniques", TEST Engineering & Management, ISSN:

0193-4120, Page No. 5554 – 5559, Vol: 81, Issue: Nov/Dec 2019.

[10] ThangarajS.John Justin, Rengarajan A, Selvanayaki S, "Comprehensive Learning On Characteristics, Applications,Issues And Limitations In MANETS", International Journal of Innovative Technology and Exploring Engineering (IJITEE) ISSN: 2278-3075, Volume-8, Issue- 9S2, July 2019.

[11] Rajasekaran S, Rengarajan A, John Justin Thangaraj S," Reliable Transmission With Flawless Power Proportion In Low Defined Signal Disseminate Systems" Global Journal of Pure and Applied Mathematics, vol. 11(5) pp(3337-3347), 2015

[12] Arvindh B, MadhuBalaji B, karthick Sudhakar G, Thangaraj S. John Justin, "Social Behaviour Based Effort Tracker System For Human Resourcing", International Journal of Pure and Applied Mathematics (IJPAM), PP 667-687, ISSN: 1314-3395, Volume 119, Number 15, 2018.

[13] Balamanigandan Ramachandran, Kamalraj Subramaniam, "Secure and efficient data forwarding in untrusted cloud environment", Cluster Computing, Springer Science+Business Media, LLC, part of Springer Nature 2018.

[14] DharmapuriSaiVamshi, V. Karthick, S. Magesh, "Text Mining and Sentimental Analysis on Comments using Machine Learning and NLP" Test Engineering and Management., vol. 82, pp. 6649 - 6652, January-February 2020

[15] YellampallySamyuktha, V. Karthick, S. Magesh, "Effective Analysis and Classification of News with Progressive Rating System" Test Engineering and Management., vol. 82, pp. 6653 - 6655, January-February 2020

[16] TejaynaBandi, John Justin Thangaraj S, S. Rinesh, K. Logu, "Efficient Client-Slide Deduplication of Encrypted Data with Public Auditing Cloud Storage", TEST Engineering & Management, ISSN: 0193-4120, Page No. 10425 -10430, Vol: 82, Issue: Jan/Feb 2020

[17] Jeyakumar D, Rajabhushanam C, "A Virtual Network Placement Algorithm for Radio Access Network", International Journal of Engineering and Advanced Technology (IJEAT), Volume-8, Issue-3S, February 2019