Opinion mining framework using proposed rb-bayes model for text classification

Rajni Bhalla, Amandeep Bagga
Department of Computer Application, Lovely Professional University, India

Article Info	ABSTRACT
	Information mining is a capable idea with incredible potential to anticipate future patterns and conduct. It alludes to the extraction of concealed information from vast data sets by utilizing procedures like factual examination, machine learning, grouping, neural systems and genetic algorithms. In naive baye's, there exists a problem of zero likelihood. This paper proposed RB-Bayes method based on baye's theorem for prediction to remove problem of zero likelihood. We also compare our method with few existing methods i.e. naive baye's and SVM. We demonstrate that this technique is better than some current techniques and specifically can analyze data sets in better way. At the point when the proposed approach is tried on genuine data-sets, the outcomes got improved accuracy in most cases. RB-Bayes calculation having precision 83,333.

Corresponding Author:

Amandeep Bagga,
School of Computer Application,
Lovely Professional University,
Jalandhar-Delhi, G.T. Road, Phagwara, 144411, Punjab, India.
Email: amandeep.bagga@gmail.com

1. INTRODUCTION

Information Discovery from Data (KDD) is the objective of information mining process [1]. Everybody is active on social media and their motive is not only too active on social media but also to generate information. Before purchasing anything, we always check reviews on social media. That reviews not only useful for the consumer but also for the manufacturer. With this pattern, there are more examinations on the programmed investigation and blend of data from client audits gathered from online networking. Because of the valuable data gave by these investigations, makers can enhance their items, the specialists can change strategies in like manner, and in addition, clients can pick the item most appropriate to their conditions. They can improve the features and can increase sales of their product.

The advancement of innovation alongside the request of investigating obstinate data has prompted another exploration subject in normal dialect preparing and information mining named "conclusion mining and notion examination". Concentrates on this issue began from the 2000s, tending to a few issues including extremity characterization [2], [3], subjectivity characterization [4], [5], [6], [7], and conclusion spam location [8], [9], [10]. Early examinations concentrated on the basic data sources which for the most part contain the sentiment on one subject and the errand is the manner by which to arrange this supposition into the classes negative, unbiased or positive [11], [12], [13].

Late issues with more entangled data sources have pulled in numerous specialists. An audit frequently contains assessments on various item perspectives, or contains practically identical feelings. A few issues of concern incorporate identifying equivalent sentences [14], [15], deciding viewpoints [16], [17], [18], rating angles [19], [20], [21] or deciding viewpoint weights [22], [23], [24]. Viewpoint based feeling examination as of late turns into a vital issue, in which we have to give the integrated slant on each item viewpoints. Viewpoint based feeling is more important because both manufacturer and customers want to

know that which features are more popular and on which features they need to do improvement. For example: Before purchasing mobile phone or laptop, customer always ask for their feature like camera,video,Bluetooth,brand,product price, special offer, dual sim,charging hour, battery backup, operating system etc. After finding which features are important, manufacturer can improve on that particular aspect.

Some past examinations, for example, [25], [26] have proposed a model called the Latent Rating Regression (LRR) which is a sort of Latent Dirichlet Allocation to break down both perspective appraisals and viewpoint weights, or [27] utilized the Maximum A Posterior (MAP) procedure to handle the angle sparsity issue. This paper [25] depicted a non-parametric Bayes technique to perform parallel order on graphs. In perspective of computational many-sided quality it may be more worthwhile to think about different strategies to adaptively discover the tuning parameter, for example, observational. An iterative estimation technique in light of the slightest item relative mistake (LPRE) [26] misfortune is developed.

An experimental probability deduction on the evaluated parameter vector is made. In this paper, our essential objective is to proposed new strategy to expel issue of probability of zero in naive bayes and to explore the estimation for parameters also, nonparametric capacity both theoretically and practically. The rest of paper is composed as takes after. In Section 2, we present the dataset. In Section 3, we present proposed RB-Bayes calculation. In Section 4, we speak to usage of RB-bayes calculation and furthermore demonstrated the examination with other algorithms. We give some finishing up comments and future work in Section 5.For the purpose of evaluation set one of variable as class label variable to convert text into numeric.If our dataset contains any text value,we need to convert into binary.Hot encoder class of sklearn will solve this purpose. RB-Bayes algorithm algorithm will apply after splitting dataset into testing and training.

2. RELATED WORKS

The best test result for music feeling grouping was the utilization of Random Forest strategies for verses and sound features.Some mixture method can be manufacture utilizing irregular backwoods moreover [28]. Assessment mining models are assembled which competent for extraction printed information into structure so deliver supposition and ordered to decide people in general reaction to the exercises in network improvement programs [29]. [30] In this paper author tested the Multinomial Naïve Bayesian classifier, Support Vector Machine and Artificial Neural Network in this exploration. In our setup, SVM beat other two classifiers with noteworthy precision for the assignment. Out work can be utilized to take out right now accessible repetitive frameworks of making feeling power vocabulary. [31] This exploration produced a Decision Tree establishes in the element "aktif" in which the likelihood of the component "aktif" was from positive class in Multinomial Naive Bayes strategy. The assessment demonstrated that the most astounding precision of arrangement utilizing Multinomial Naïve Bayes.

3. RESEARCH METHOD
3.1. Dataset

From dataset, we can predict whether customer will purchase computer or not. We have five parameters. On basis of these parameters we will predict. Parameters are age, income; student and credit rating. As we can see from Figure 1, Age wise maximum response are from youth and senior. To check accuracy and for comparison we test proposed algorithm on small dataset. Similarly we have parameter income that consists of three values high, medium and low. Student type will be binomial consist of two values either yes or no.Credit rating will also be binomial type and consist of two values either fair or excellent. We want to predict variable buys computer consist of two values either yes or no. Similarly we have other parameters i.e. income' which consist of three values High, medium and low; student which consist values-Yes or No; Credit rating which consist values fair or excellent.
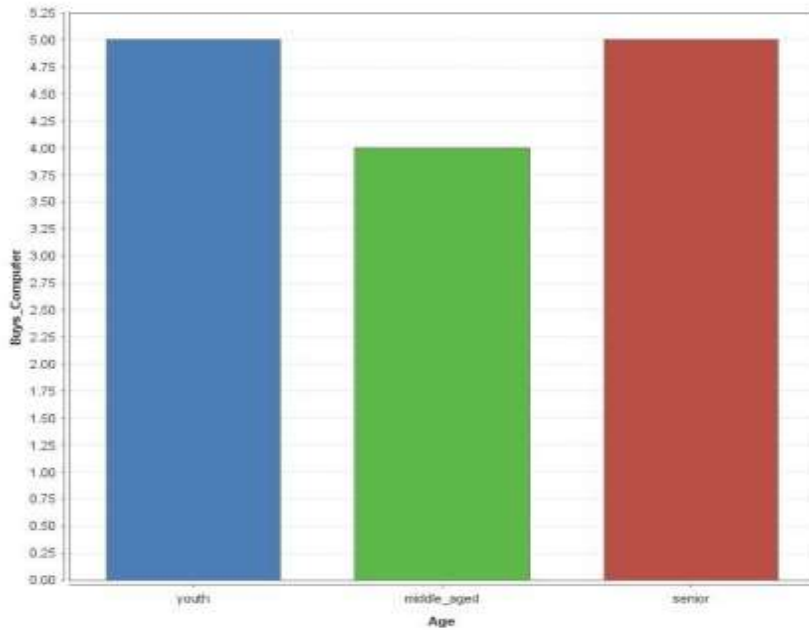
Figure 1. Age-wise response

3.2. RB-Bayes Algorithm

RB-Bayes is one of simplest supervised technique. It is a classification system in light of bayes theorem. It is mostly used in text classification. Naive bayes is also based on bayes theorem. But unable to handle problem of likelihood of zero possibility. RB-Bayes is proposed to solve this problem.RB-Bayes algorithm provides a way of calculating prediction. Look at equation below.

$$P_y\,F = \frac{T_y}{Total_{Sampleset}} * (\frac{Ty_i + \cdots \dots \dots \dots \dots. + Ty_n}{T_F * T_y})$$

RB-bayes Algorithm steps
1. Each tuple that we wish to classify is represented by X=(x_1 , x_2 ……….x_n)
2. There are n numbers of labels. Given a tuple, X, the classifier will anticipate that X has a place with the label having the highest value from all labels.
3. Checking highest value for labels
   ($P_y\,F > P_n F$) Where y ≠ n
   Value of y and n are different labels. Maximum value from all labels will do prediction.
4. Maximize

   $$Mean = \frac{T_y}{Total_{Sampleset}}$$

   $$P_y F = Mean * (\frac{T_y a + T_y b + T_y c + T_y d \dots \dots \dots \dots T_y n}{T_F * T_y})$$

5. T (Y_i), for i=1, 2, 3….n, is a prior probability value depends on labels. Prior probability of each class can be computed based on training tuples. We calculate $T_y a$, $T_y b$, $T_y c$ and $T_y d \dots \dots \dots. T_y n$ and this needs to be maximized. Ty_a is calculated by comparing value with P(y).Count will store in $T_y a + T_y b + T_y c + T_y d \dots \dots \dots \dots T_y n$ wherever both values are active. Similarly for $T_y b$, $T_y c$ and $T_y d$.

   $$T_y a = \prod_{k=1}^{n} P(x_{k} \mid Y_i)$$

   $$= P(x1 \mid Y_i) * P(x_2 \mid Y_i) * \dots \dots ..P\,(x_n \mid Y_i)$$

6. Predicting class label depending upon highest value after comparing value of $P_y F$, $P_n F$
7. Factors affecting y and n can be n number of factors i.e. $T_y a$, $T_y b$, $T_y c$ and $T_y d$ ... ... $T_y n$
8. The classifier predicts that the class label of tuple X is the class $P_y F$ or $P_n F$
   If and only if

$$\frac{T_y}{Total_{SS}} * \left(\frac{T_y a + T_y b + T_y c + T_y d \ldots \ldots T_y n}{T_F * T_y} \right) > \text{Or} < \frac{T_n}{Total_{SS}} * \left(\frac{T_n a + T_n b + T_n c + \cdots \ldots \ldots \ldots + T_n n}{T_F * T_n} \right)$$

RB-Bayes classifier have minimum error rate as compared to other algorithms. All factors are taken into consideration.

4. RESULTS AND ANALYSIS

Python is used for implementation of methodology. We build RB-based algorithm based on baye's theorem. Preprocessing steps done before applying proposed algorithm is shown in Figure 2. After we have dataset on which we want to implement this algorithm. We need to perform some preprocessing steps. Seperate tuple from label on which we want to do prediction. Single row represent tuple. We check it on small datasets as shown in Figure 3, compare with naive baye's also. Dataset contains text data also. So need to convert this data into numeric form. Some categorical variables consist of more than two values. So after converting dataset into numeric, dataset consist of values more than 0 and 1 also depending on category. We need to do dummy encoding of those variables which consists more than two values.

Our dataset is as great blend of categorical and continuous qualities and fills in as a helpful case that is moderately simple to understand. Thus, the examiner is looked with the test of making sense of how to transform these content traits into numerical qualities for encourage processing. Label encoding has the preferred standpoint that it is direct however it has the drawback that the numeric qualities can be "confounded" by the calculations. A typical elective approach is called one hot encoding. In spite of the diverse names, the fundamental methodology is to change over every classification esteem into another segment and appoints a 1 or 0 (True/False) esteem to the segment.Dataset is shown in Table 1.
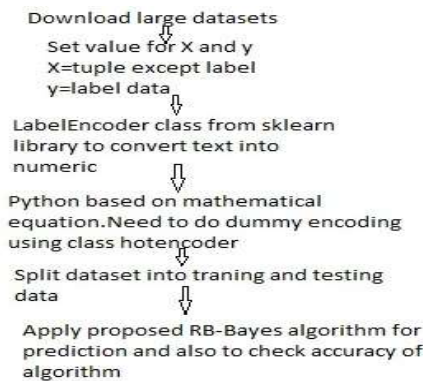

Figure 2. Methodology used for prediction

Table 1. Dataset

Age	Income	Student	Credit ratimg	Buys_Computer
Youth	High	No	Fair	No
Youth	High	No	Excellent	No
Middle_aged	High	No	Fair	Yes
Senior	Medium	No	Fair	Yes
Senior	Low	Yes	Fair	Yes
Senior	No	Yes	Excellent	No
Middle_aged	Low	Yes	Excellent	Yes
Youth	Medium	No	Fair	No
Youth	Low	Yes	Fair	Yes
Senior	Medium	Yes	Fair	Yes
Youth	Medium	Yes	Excellent	Yes
Middle_aged	Medium	No	Excellent	Yes
Middle_aged	High	Yes	Fair	Yes
Senior	Medium	No	Excellent	No

Table 2. Dummy Encoding

Age	Middle_aged	Senior	Youth	Income	High	Low	Medium
Youth	0	0	1	High	1	0	0
Youth	0	0	1	High	1	0	0
Middle_aged	1	0	0	High	1	0	0
Senior	0	1	0	Medium	0	0	1
Senior	0	1	0	Low	0	1	0
Senior	0	1	0	Low	0	1	0
Middle_aged	1	0	0	Low	0	1	0
Youth	0	0	1	Medium	0	0	1
Youth	0	0	1	Low	0	1	0
Senior	0	1	0	Medium	0	0	1
Youth	0	0	1	Medium	0	0	1
Middle_aged	1	0	0	Medium	0	0	1
Middle_aged	1	0	0	High	1	0	0
Senior	0	1	0	Medium	0	0	1

As we know python is based on mathematical equations, so all dataset must in binary form. But our dataset contains text data. First step is to convert all dataset into numeric. We import class Label Encoder from sklearn to change data from text into numeric. Age and income contains three categories. Under age categories youth converted into 2, middle-aged converted into 0 and last senior converted into 1.Under income categories high is converted into 0,medium converted into 2 and low is converted into 1. OneHotEncoder class will convert this data from numeric to binary because python understand only binary data. Dummy encoding will be generated by OneHotEncoder class as shown in Table 2. Data has been converted into binary form.

We compare the result with naive Bayes. Suppose we wish to predict value for below tuple whether this tuple will purchase computer or not. X = (age=youth, income=medium, student=yes, credit rating=fair) Using RB-Bayes algorithm, we are going to predict the possibility for above tuple.

$$Mean_{yes} = \frac{T_y}{Total_{sample}Set}$$

$$Mean_{no} = \frac{T_n}{Total_{sample}Set}$$

Total yes consist of all records that purchases computer and Total no consist of those who do not purchase computer. So, we are calculating mean.
mean_yes=0.64 and mean_no=0.36 and total number of samples =14
After calculating mean, now we take summation of all factors who said yes to purchase computer and similarly calculating for those who said no for purchasing computer.

$$P_yF = \frac{T_ya + T_yb + T_yc + T_yd}{F_y}$$

$$P_nF = \frac{T_na + T_nb + T_nc + T_nd}{F_n}$$

Tya, Tyb, Tyc and Tyd are total number of factors where Tya is one and even value for label is also one. Where both of the condition is true that gives value for these factors. Totalyes count number of yes from total number of samples.Totalno count number of no from total number of samples. To calculate value for probability for yes or no we multiply value with meanyes and meanno respectively after calculating summation of PyF and PnF.

$P_yF=18$
$P_nF=8$

$$P_{yes}=P_yF * mean_{yes}$$
$$P_{no} = P_nF * mean_{no}$$

$P_{yes}=0.32$
$P_{no}=0.14$

Compare values for probability of yes and probability of no to find greater value. Greater value will decide whether particular tuple will purchase a computer or not. Value for P_yF is greater than P_nF.

So we can predict that this tuple will purchase computer.Our algorithm removes the problem of zero probability and also improves accuracy. To find this we divide our data into training and test set. We set test size=0.37.We test same dataset using naïve bayes and RB-bayes algorithm also to check accuracy.After calculating value for probability of yes i.e. P_{yes} and probability of no i.e.P_{no}, we compare these two values and highest value does prediction. We test the accuracy score in python using class accuracy score. From sklearn.metrics import accuracy_score print ('Accuracy score:', accuracy_score(y_test,y_pred))

Using RB-Bayes algorithm, value of accuracy =83.3Using Naive bayes algorithm, value of accuracy=50. Reason is in Naive Bayes when we end up with probability of zero, we lose effect for other factors also. Although we use Laplace correction that each one value in each account but in actual value is zero. IN RB-Bayes, we remove possibility of zero. In machine learning, Support vector machines (SVMs, additionally bolster vector systems) are directed learning models with related learning calculations that break down information utilized for grouping and relapse examination. We apply same dataset on SVM also for comparison. SVM methodology is implemented in Rapid miner as shown in Figure 3.
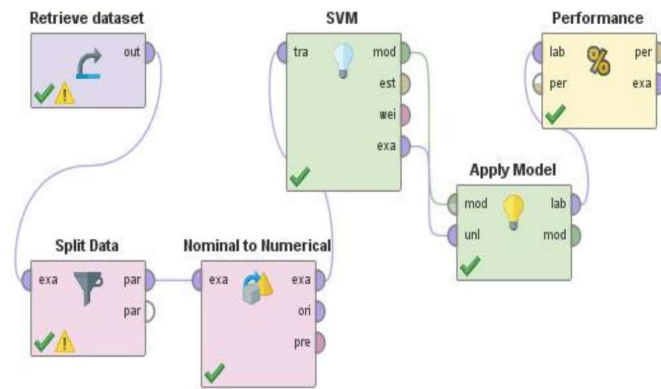

Figure 3. Implementation with SVM

Nominal to numerical operator is used to convert text data into numerical.Before applying SVM; it is require converting data into numerical form. Performance operator is used to test the accuracy of model.Confusion matrix is generated as shown in Figure 4. This Operator is utilized to measurably assess the qualities and shortcomings of a double order, after a prepared model has been connected to named data. A paired arrangement makes forecasts where the result has two conceivable qualities: call them positive and negative.

In addition, the forecast for every Example might be correct or wrong, prompting. TP - the quantity of "genuine positives", positive Examples that have been accurately distinguished. FP - the quantity of "false positives", negative Examples that have been inaccurately recognized. FN - the quantity of "false negatives", positive Examples that have been inaccurately recognized. TN - the quantity of "genuine negatives", negative Examples that have been accurately recognized. 83.3 is not bad accuracy of RB-Bayes algorithm. So we can characterize exactness measures of model as a component of the check accurately anticipated records.Accuracy using SVM shown in Figure 5.
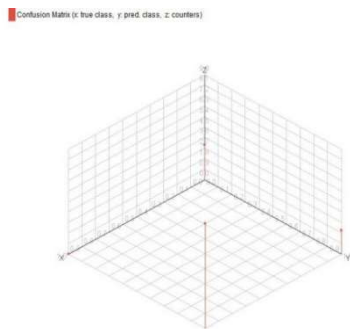

accuracy: 85.71%

	true no	true yes	class precision
pred. no	3	0	100.00%
pred. yes	2	9	81.82%
class recall	60.00%	100.00%	

Figure 4. Confusion matrix   Figure 5. Accuracy using SVM

5.   CONCLUSION AND FUTURE WORK

In this paper, we study the supervised techniques which allow doing prediction based on training data. Naive bayes algorithm for data mining has been reviewed and a new approach is proposed. It is important to stress that the proposed algorithm consider all factors even if probability of likelihood is zero. Apart from the existing supervised techniques, this model may also be of interest in market where manufacturers or seller wants to know why their sale is up or down. On what factors they need to give importance or work. They can improve sales performance. We can know what the factors affect the buying decision of customer are. Tests are directed to confirm this calculation for small datasets and promising outcomes are acquired. In future, this idea of amalgamation of bunching and characterization can be connected over enormous information influencing utilization of guide to diminish method to deal with vast databases.

REFERENCES

[1]   Xu LEI, Jiang C, Wang J. "Information Security in Big Data : Privacy and Data Mining," 2014;1149–76.
[2]   Turney PD. "Thumbs up or thumbs down? Semantic Orientation applied to Unsupervised Classification of Reviews," *Proc 40th Annu Meet Assoc Comput Linguist.* 2002;(July):417–24.
[3]   Pang B, Lee L, Vaithyanathan S. "Thumbs up?: Sentiment Classification Using Machine Learning Techniques," *Proc Conf Empir Methods Nat Lang Process* [Internet]. 2002;79–86. Available from: http://portal.acm.org/citation.cfm?id=1118693.1118704
[4]   Mihalcea R, Banea C, Wiebe J. "Learning Multilingual Subjective Language Via Cross-Lingual Projections," *Proc 45th Annu Meet Assoc Comput Linguist* [Internet]. 2007;(June):976–83. Available from: http://acl.ldc.upenn.edu/p/p07/p07-1132.pdf
[5]   Su F, Su F, Markert K, Markert K. "From words to senses: a case study of subjectivity recognition," *Proc 22nd Int Conf Comput Linguist 1* [Internet]. 2008;(August):825–832. Available from: http://portal.acm.org/citation.cfm?id=1599081.1599185
[6]   Pang B, Lee L. "A Sentimental Education: Sentiment Analysis Using Subjectivity Summarization Based on Minimum Cuts," 2004; Available from: http://arxiv.org/abs/cs/0409058
[7]   Pang B, Lee L. "Opinion Mining and Sentiment Analysis. Found Trends® InformatioPang," B, Lee, L (2006) Opin Min Sentim Anal Found Trends® Inf Retrieval, 1(2), 91–231 doi101561/1500000001n Retr [Internet]. 2006;1(2):91–231. Available from: http://www.nowpublishers.com/product.aspx?product=INR&doi=1500000001
[8]   Lim E-P, Nguyen V-A, Jindal N, Liu B, Lauw HW. "Detecting product review spammers using rating behaviors," Proc 19th ACM Int Conf Inf Knowl Manag - CIKM '10 [Internet]. 2010;939. Available from: http://portal.acm.org/citation.cfm?doid=1871437.1871557
[9]   Jindal N, Liu B. "Opinion spam and analysis," Proc Int Conf Web search web data Min WSDM 08 [Internet]. 2008;219. Available from: http://portal.acm.org/citation.cfm?doid=1341531.1341560
[10]   Jindal N, Liu B, Street SM. "Review Spam Detection," 2007;1189–90.
[11]   Liu B, Street SM. 'Opinion Observer:Analyzing and Comparing Opinions on the Web," Proc 14th Int Conf World Wide Web [Internet]. 2005;342–51. Available from: http://dl.acm.org/citation.cfm?id=1060797
[12]   Morinaga S, Yamanishi K, Tateishi K, Fukushima T. "Mining Product Reputations on the Web," Proc eighth ACM SIGKDD Int Conf Knowl Discov data Min-KDD '02 [Internet]. 2002;341. Available from: http://portal.acm.org/citation.cfm?doid=775047.775098
[13]   Li F, Han C, Huang M, Zhu X, Xia Y-J, Zhang S, et al. "Structure-aware Review Mining and Summarization," Proc 23rd Int Conf Comput Linguist [Internet]. 2010;(August):653–61. Available from: http://dl.acm.org/citation.cfm?id=1873781.1873855
[14]   Jindal N, Liu B. "Identifying Comparative Sentences in Text Documents," Proc 29th Annu Int ACM SIGIR Conf Res Dev Inf Retr-SIGIR '06 [Internet]. 2006;244. Available from: http://portal.acm.org/citation.cfm?doid=1148170.1148215
[15]   Kim HD, Zhai C. "Generating Comparative Summaries of Contradictory Opinions in text," Proceeding 18th ACM Conf Inf Knowl Manag-CIKM '09 [Internet]. 2009;385. Available from: http://portal.acm.org/citation.cfm?doid=1645953.1646004
[16]   Hu M, Liu B. "Mining and Summarizing Customer Reviews," Proc 2004 ACM SIGKDD Int Conf Knowl Discov data Min KDD 04 [Internet]. 2004;4:168. Available from: http://portal.acm.org/citation.cfm?doid=1014052.1014073
[17]   Jo Y, Oh AH. Aspect and Sentiment Unification Model For Online Review Analysis," Proc fourth ACM Int Conf Web search data Min - WSDM '11 [Internet]. 2011;815. Available from: http://portal.acm.org/citation.cfm?doid=1935826.1935932
[18]   Wu Y, Zhang Q, Huang XX, Wu L. "Phrase Dependency Parsing for Opinion Mining," Proc 2009 Conf Empir Methods Nat Lang Process Vol 3 EMNLP 09 [Internet]. 2009;(August):1533–41. Available from: http://portal.acm.org/citation.cfm?doid=1699648.1699700
[19]   Snyder B, Barzilay R. "Multiple Aspect Ranking using the Good Grief Algorithm," 2005.
[20]   Titov I, McDonald R. "A Joint Model of Text and Aspect Ratings For Sentiment Summarization," Proc ACL08 HLT [Internet]. 2008;51(June):308–16. Available from: http://www.aclweb.org/anthology/P/P08/P08-1036.pdf
[21]   Pham DH, Le AC, Le TKC. "A Least Square Based Model for Rating Aspects and Identifying Important Aspects on Review Text Data" Proc 2015 2nd Natl Found Sci Technol Dev Conf Inf Comput Sci NICS 2015. 2015;265–70.

[22] Pham, Duc-Hong and Le A-C and others. "Determining Aspect Ratings and Aspect Weights from Textual Reviews by Using Neural Network with Paragraph Vector Model," In: Nguyen HT, Snasel V, editors. International Conference on Computational Social Networks [Internet]. 1st ed. Vietnam: Springer International Publishing; 2016. p. 309--320. Available from: https://link.springer.com/bookseries/558

[23] Zha ZJ, Yu J, Tang J, Wang M, Chua TS. "Product Aspect Ranking And Its Applications," IEEE Trans Knowl Data Eng. 2014;26(5):1211–24.

[24] Pham D-H, Le A-C. "A Neural Network based Model for Determining Overall Aspect Weights in Opinion Mining and Sentiment Analysis," Indian J Sci Technol [Internet]. 2016;9(18):1–6. Available from: http://www.indjst.org/index.php/indjst/article/view/93164

[25] Hartog J, van Zanten H. "Nonparametric Bayesian Label Prediction on a Graph," Comput Stat Data Anal [Internet]. Elsevier B.V.; 2018;120:111–31. Available from: https://doi.org/10.1016/j.csda.2017.11.008

[26] Muthulakshmi S, Dash CS, Prabaharan SRS. "Memristor Augmented Approximate Adders and Subtractors for Image Processing Applications: An Approach," AEU-Int J Electron Commun [Internet]. 2018;91:91–102. Available from: https://doi.org/10.1016/j.aeue.2018.05.003

[27] Abdulrazzaq MM, Yaseen IFT, Noah SA, Fadhil MA. "Multi-Level of Feature Extraction and Classification for X-Ray Medical Image," 2018;10(1):154–67.

[28] Hastarita Rachman F, Sarno R, Fatichah C. "Music Emotion Classification based on Lyrics-Audio using Corpus based Emotion," Int J Electr Comput Eng [Internet]. 2018;8(3):1720–30. Available from: http://iaescore.com/journals/index.php/IJECE

[29] Yuliyanti S, Djatna T, Sukoco H. "Sentiment Mining of Community Development Program Evaluastion based on Social Media," TELKOMNIKA (Telecommunication Comput Electron Control [Internet]. 2017;15(4):1858. Available from: http://journal.uad.ac.id/index.php/TELKOMNIKA/article/view/4633

[30] Azim MA, Bhuiyan MH. "Text to Emotion Extraction Using Supervised Machine Learning Techniques," TELKOMNIKA (Telecommunication Comput Electron Control), 2018;16(3):1394–401.

[31] Susanti AR, Djatna T, Kusuma WA. Twitter's Sentiment Analysis on Gsm Services using Multinomial Naïve Bayes. TELKOMNIKA (Telecommunication Comput Electron Control) [Internet]. 2017;15(3):1354. Available from: http://journal.uad.ac.id/index.php/TELKOMNIKA/article/view/4284

BIOGRAPHIES OF AUTHORS

Rajni Bhalla is pursuing PHD from lovely professional University. She is working as assistant professor in Lovely Professional university. She is interested in data mining, data analysis, feature extraction, prediction, machine learning and clustering techniques.

Amandeep bagga is an Associated Professor in Department of Computer Application at Lovely Professional university, India. She has completed his studies for B.C.A, M.C.A, Ph.D. from LPU.Her research interest include cloud computing, security, network security and cryptography.