International Journal of Computer Sciences and Engineering Open Access
Research Paper Vol.-6, Issue-7, July 2018 E-ISSN: 2347-2693

Applying Sentiment Analysis to Predict Rating and Classification of Text Review

M. M. Sutar [1*], T. I. Bagban [2]

[1]Dept. of Computer Science and Engineering, D.K.T.E‘S TEI (An Autonomous Institute), Ichalkaranji, India.
[2]Dept. of Information Technology, D.K.T.E‘S TEI (An Autonomous Institute), Ichalkaranji, India.

[*]*Corresponding Author: mlindsutarsm@gmail.com*

Available online at: www.ijcseonline.org

Accepted: 18/July/2018, Published: 31/July/2018

Abstract— In past few years prevalence of internet usage is increasing. Different online shopping sites have many options of purchasing the product while shopping. Users share experiences in the form of reviews. The number of reviews shared by people are increasing. So it is difficult to find the right information about the product. Traditional recommender systems (RS) makes use of different factors, such as users purchase records, geographical location etc. We propose sentiment based recommender system. Based on the sentiment word in the CPRST system, the review has been rated by finding sentiment score. Also textual reviews are categorized into different feature of product using text classification technique. Experimental results of CPRST system show that user preference can be characterized by the sentiment from text review and it can improve the performance of recommendation system. Using LDA method we classified text review and this resulted in good results in nickel.

Keywords— Item reputation, Text Reviews, Rating prediction, Recommender system, Sentiment influence, User sentiment, Sentiment analysis, Text classification.

I. Introduction

Nowadays every peoples have smart device with high speed Internet connectivity, everyone connected to social network. People’s access and share informational data on Internet, peoples are become information consumer as well as information producer. Today there are huge amount of data are available on internet, so the information overloading problem arises. Peoples confuse which information is reliable and which is fake. Users cannot easily trust on online service/product, each users have different thinking about same product. So that peoples shared data (we will say that text review) plays an important role in decision making. For example consumer decides what to buy after reading the valuable text reviews shared by others. But we not easily find out valuable review because of information overloading. Consumer believe on other people’s review because they help to predict the product reliability. Star rating is key point to find valuable reviews. Prediction of rating based on sentiment CPRST is based on the idea that high-star ratings mean it is related with the good reviews. Reviews contain detailed information along with user opinion, which is important for a user to choose a product to be purchased or not. Some people are think about price, quality and other comparative factors. All these factors describe the user’s interests according to their comments on the product.

Sentiment analysis mainly use in natural language processing, text analysis to extract, identify and study subjective information. Sentiment analysis widely used in customer’s domain, such as online shopping, social media, review and survey response. Sentiment analysis is the most fundamental task in extracting user’s interest from text review. The sentiment is used to predict the rating score of review. Before that, there are directly star rating options available by which user select number of stars on its own experience of the product, but this result may not be valid because it not clear plus or minus points of product. To make a more accurate rating user sentiment takes an important role. Reviews are in two types positive or negative. To make a purchase decision, customers not only need to know whether the product is good, but also need to know how good the product is. For example, some users prefer to use “good” to describe an “excellent” product, while others may prefer to use “good” to describe a “just good, not a best” product. Item’s reputation depends on customer’s text reviews. The reputation of the product is obtained from the sentiment words. A product is of good reputation if it is having positive sentiment whereas negative sentiment represents bad reputation. So those reviews are to be explored who have objective attitude on items. When a reviewer gives like and dislike for a particular product other user's attention is drawn.

Rating prediction is based on sentiment from text review Framework of matrix factorization is used along with the method of sentiment-based by RPS system. In this work, sentiment analysis is used to predict the ratings. Features of products are to be extracted from the reviews after that sentiment words are retained from text review. Sentiment dictionaries is to be used to calculate the sentiment of a reviews. In Fig.1, based on the previous user preference of item from text reviews, the last item will be recommended to the last user. The work is given in [1], [2], [3], [4], [5], here user sentiment are classified into two main polarities positive sentiment and negative sentiment. RPS does two thing side by side finding interpersonal sentiment influence and item's reputation along with mining the user sentiments. At last, everything is taken into the recommender system.

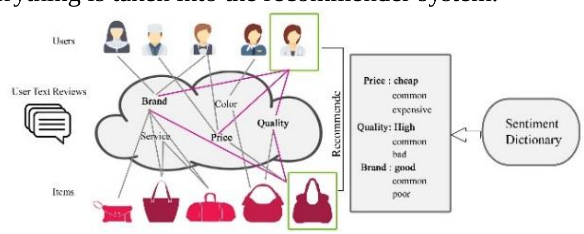

Figure 1. Sentiment-based rating prediction method

With star rating if reviews are classified based on specific object, user can so easy get a valuable reviews which are he want.

If we predict the rating for text is easy to find valuable reviews to customer. But there is on more complexity is to find review for some specific feature of product like product specification, price, service, hardware configuration, software configuration, brand etc. The classification make more attraction of customer to product by providing this kind of information easily or quickly as they required.

The remainder of this paper is organized as follows. In Section 2, we present the related work about rating prediction in recommender systems. Section 3 explain the methodology, Section 4 is about Result and discussion. Conclusion and Future work are drawn in Section 5.

II. RELATED WORK

A. Collaborative Filtering

Collaborative filtering (CF) is technique in which automatically prediction is done on the interest of user by collecting preference and choices of many different users. The amount of information on internet going to increase very quickly. The ability to process them CF work on database of user's preference for item. CF is success to filter the information, however there are two fundamental challenges first one is scalability: if information going to more than ten thousands, the CF have many challenges of filtering. The existing algorithm of CF have performance problem with individual users for whom that site has huge size of information. Second challenge is of improve the accuracy of recommendation for user. As large size of information CF take more time and have no accuracy as expected. B. Sarwar et al. [6] Item-based collaborative filtering algorithm was introduced to overcome the challenges. The CF technique which is item based analyse the matrix of User-Item to identify the relationship between different items and these relations are used to calculate the recommendations of user. B. Sarwar et al. analyse different types of techniques for computing item-item similarities [6]. Overcome limitation of k-nearest neighbour approach and give better performance. K.H. L. Tso-Sutter et al. [7] propose generic method which allow tag to be three-dimensional correlation to three two-dimensional correlations. And then apply fusion method to reassociate correlation of dimension.

B. Matrix Factorization based Approaches

CF has performance issue if there are large size of database. To overcome this issue R. Salakhutdinov. et al. [8] presents the probabilistic matrix factorization (PMF) model. The PMF model scales linearly with number of observation. PMF perform well with large dataset as compare to CF. Salakhutdinov. et al. added adaptive prior and shown that system can control capacity automatically this was done to extend the PMF model. A new version of PMF was introduced that does assumption based on similar preferences. Means that user who have similar set of item/product having similar preference. PMF compared their system with NetFlix and it was found that PMF gave 7% better performance and achieved error rate up to 0.8861. Salakhutdinov. et al. present two derivation with PMF are PMF with learnable prior and constrained PMF.

C. Reviews based Applications

Increasing web of social network peoples are connected to each other and people like to share their day to day experience, such as rating, review and blogs. X. Qian. et al. [9] propose three social factors, personal interest, interpersonal interest similarity and interpersonal influence. This three factors are combined and built into the unified personalize recommendation. Personal interest means rating the items by different individual users and after that combining their factors to improve the accuracy and applicability of the rating prediction method. X. Qian. et al. [9] conduct experiment of three large size of dataset. L. Qu. et al. [10] here a concept of bag-of-opinions concept was introduced where the opinions of reviews consist of 3 main factors like root word, component, a set of modified words and negation words. L. Qu. et al. present ridge regression algorithm for learning opinion scores and n-gram features [10].

Due to automated mining of product reviews ranking score was re-calculated and this was a valuable tool because of which customers were able to make more information decisions. K. Zhang. et al. present product ranking model that applies weights to product review factor to calculate a product ranking score [11]. K. Zhang. et al. experiment his work on amazon.com, they present novel approach (model) to rank products by analysing the sentiment of review. K. Zhang. et al. consider various product review factors such as quality of product, review time, durability of product, and historical positive review of customers.

D. Sentiment based Applications

Sentiment analysis conducted at Review level, sentiment-level and phrase-level. B. Pang. et al. [12] propose a context insensitive evaluation lexical method. They classify document based on overall sentiment. Naive Bayes, maximum entropy classification, and support vector machines this machine learning methods not perform as well on sentiment classification as on traditional topic based classification [12]. D. Tang. et al. [13] find issue by incorporating user-level information and product-level information into neural network method for classification of document level sentiment. Vector space model is used to modelled user and products. Which capture important clues of product like in individual user's performance or quality of product. D. Tang. et al. achieve state-of-the-art performance by combining evidence at user-level, product-level and document-level in unified framework of neural. D. Tang. et al. introduce user-product neural network for document level sentiment classification. T. Nakagawa. et al. presence dependency tree based methods for sentiment classification of Japanese and English subjective sentences [14]. Content words often by subjective sentence, reverse the sentiment polarities of other words. So interaction between words in sentiment classification need to consider by using bag-of-words approach, it is difficult to handle. T. Nakagawa. et al. exploited syntactic dependency structure of subjective sentence. By hidden variable sentiment polarity of each dependency sub-tree in sentence which is not observable in training data is represented. The polarity of sentence is calculated.

To mine important information from users review and recommended it to determine user preference is difficult task. User purchase record, product category, and geographical location this factors are considered in traditional recommendation system. Xiaojiang Lei. et al. propose sentiment based rating prediction method (RPS) to improve accuracy of prediction[15]. Xiaojiang Lei et al. propose three factors to prediction, first one is social user sentiment, second users own sentiment attribute with interpersonal sentiment influence and last is product reputation. This thee factors are fused into unified matrix factorization framework to achieve task of rating prediction [15].

III. METHODOLOGY

The main intention of the approach is to predict the rating from sentiment words and classify the text review based on different features of products. Figure 2 describes the system architecture of the system. First product features are extracted from text reviews using Latent direchlit allocation LDA method, after that the basic sentiment score is measured to predict the ratings, at the final using classification technique and product feature table the reviews are classified. The following sub-sections describe the details of the approach.

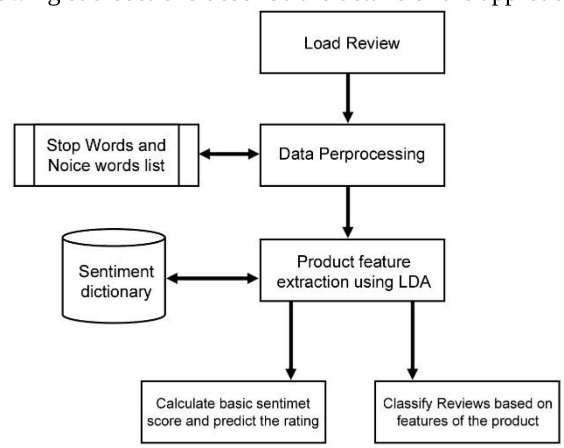

Figure 2. System architecture.

A. Product feature extraction using latent dirichlet allocation method

I. Data Pre-processing

Consider each user's review as a collection of words without considering the order and construct vocabulary then filter all "Stop words", "Noise words", Sentiment words, Sentiment degree words and negation words from text review. After word filtering, get input text clear and without much interference. Then construct vocabulary V with all this unique words. This V is used to input for LDA. We need to construct another vocabulary Vs to measure the score of text review, in which we filter only common words, stop words and noise words.

II. Generate process for LDA

LDA is used in natural language processing model to classify the words into different topics. LDA is generative topic bag of words model that automatically discover topic in text document [16]. In LDA model each word in document belongs to one of topics. To assign number of topics T and user document D is given as input to LDA. The output is topic preference distribution for each topic list and user. In each topic there may be more than two words.

LDA Input: Text documents V

LDA Output: Topic List

- **Generative process of LDA**

LDA assume that new document are created in the following way:

1. Determine number of words in document
2. Choose topic mixture for document over a fixed set of topics (i.e 20% topic A, 30% topic B, 50% topic C)
3. Generate the words in document by :

a. Find word to topic distribution = p(topic t/ document d)
b. Find document to topic distribution = p(word w/ topic t)

4. Multiply above to portions and assign w a new topic based on that probability.

LDA model and a synonym lexicon to extract product features from online product reviews addresses the limitations of existing feature extraction approaches, namely the dependence on manual work and domain-specific nature [17].

III. Product Feature Extraction

In each topic there are some frequent words, based on co-occurrence with adjective words and their frequency in baground corpus they need to filter features from candidate set. Table 1 shows an example of topic and product features of smartphone product.

Topics	Example of Product Features
Topic 0	Price, discount, bought, worth, cash, buy, buying, pay, sold, money, cost, dollars...
Topic 1	service, shipping, Seller, people, review, arrived, receive, return, received, supply, customer, warranty, service centre...
Topic 2	Display, Charger, speakers, screen, keypad, microphone, battery, mic, microphone, storage, gsm, camera, Bluetooth, port, card...
Topic 3	Height, screen, handle, thin, features, slim, colour, resolution, weight, size, looks, looking, smaller, compact, user friendly, Sharp...
Topic 4	set up, settings, charging, power, upgraded, android, surfing, works, working, volume, slow, faster, ringtones, download, applications, heat...

Table 1. Frequent product features of topic on smartphone.

B. Sentiment measurement

We use sentiment dictionary[1] to calculate user sentiment on product. We create mainly two list of sentiment dictionary (SD). POS-Words, NEG-Words. POS-Words list containing positive sentiment words list and positive evaluation words list, NEG-Words list containing negative sentiment words list and negative evaluation words list. We use sentiment degree dictionary (SDD), which has manly five degree levels. Level-1 means highest degree of sentiment, Level-2 means higher degree of sentiment likewise Level-3 having lower degree than level-2, Level-4 lower degree than level-3 and Level-5 has lowest degree of sentiment. Also, we use negation word dictionary (ND) containing all negation words. Negation words reverse the polarity of sentiment word. The respective sentiment dictionary word list as shown in table 2.

Table 2. The sentiment dictionaries

Dictionaries	Representative words list
SD	**POS-Words:** attractive, beautiful, comfy, convenient, delicious, delicate, exciting, happy, nice, ok, … **NEG-Words:** annoyed, awful, bad, poor, boring, complain, expensive, hostile, terribly, unfortunate, worse …
SDD	**Level-1:** most, best, greatest, absolutely, extremely, highly, excessively, completely, entirely, 100%, highest, sharply, superb… **Level-2:** awfully, better, lot, very, much, over, greatly, super, pretty, unusual… **Level-3:** even, more, far, so, further, intensely, rather, relatively, slightly more, insanely, comparative. **Level-4:** a little, a bit, slight, slightly, more or less, relative, some, somewhat, just. **Level-5:** less, not very, bit, little, merely, passably, insufficiently.
ND	no, nor, not, never, nobody, nothing, none, neither, few, seldom, hardly, haven't, can't, couldn't, don't, didn't, doesn't, isn't, won't,…

To predict rating from sentiment word we first need to calculate score of the review. To measure the score we need to divide the original text review into number of sentence or we say that clauses. Then for each clause we find sentiment word and match to our SD dictionary, for positive word we assign +1.0 and for negative we assign -1.0. After that we find sentiment degree word form SDD and assign value as per its level of degree. As per empirical rule in [18], [19] We assign 5 value for level-1, 4 value for level-2, 2 for level-3, for level-4 we assign 0.5 and for level-5 we assign value 0.25.. Finally we check negation coefficient that has by default value of +1.0. If sentiment word preceded with odd number of negative prefix, we set +1.0 value to -1.0. Then for review r we get sentiment score as follows:

$$S(r) = \frac{1}{N_c} \sum_{c \in r} \sum_{w \in c} Q \cdot D_w \cdot R_w \tag{1}$$

Where c denotes number of clauses. N_c denotes total number of clauses in text review. Q denotes the negation check coefficient. D_w determines level of sentiment degree word from SD. R_w demotes initial score of sentiment word w.

To improve the precision of mapping of sentiment, there are conjunctive rules are used as:

- **"and" rule:** In the review clauses that are connected with "and" like conjunctive rule usually having same polarity. For example, "It does everything I need and works great". Other "and" like term includes: "as well as", "likewise".
- **"but" rule:** In the review clauses that are connected with "but" like conjunctive rule usually having opposite polarity. For example, "Mobile works good but it goes slow sometimes" Other "but" like term includes: "however", "thought" and etc.

- "and" conjunctive rule which connect two nouns which also we detect, because two noun connected with "and" is not a two different sentences or clauses. So that we can achieve better score of review. Examples: "Me and my friend use this Nokia mobile from last few months.", "you and me", "brother and sister", "he and me", "he and his cousin" etc.

After calculating the basic sentiment score and improving sentiment mapping from text review *r*, we need to normalize the score as.

$$E_r = \frac{10}{1+e^{-S(r)}} - 5 \tag{2}$$

Based on this methods, we get all sentiment score and rating.

C. Document classification using feature words

Using table 1. of feature words, find the frequency or number of occurrence of topic words (feature words) in text review and then classify review according to maximum number of occurrence of respective topic words.

IV. RESULTS AND DISCUSSION

A. Sentiment analysis and rating prediction

To analyse the performance of sentiment analysis and rating prediction we focus on precision. We use 85% training dataset and 15% we create as required for our dataset. We evaluate sentiment by transforming sentiment value E_r into binary number, we categories $E_r>0$ as positive review, $E_r<=0$ as negative review. We label all 5-star value of dataset review as positive and 1-star for negative. E_r further normalised using following equation.

$$R(s) = \frac{E_r+5}{2} \tag{3}$$

First we evaluate system on total of 1895 review of Amazon reviews unlocked mobile phones [21] on which 1475 have positive review and 315 negative. Secondly we evaluate system on 541 review of Amazon Reviews [23] on which we get 450 as positive review and 65 negative. Further we evaluate our system on Yelp dataset [15]. The statistics and evaluation result of our system for both dataset are shown in following table 3. We can see that precision for positive review for Yelp is 78.15%, for Amazon reviews unlocked mobile phones dataset is 74.43% and for Amazon Reviews dataset is 76.21%.

Dataset	Scale	Precision of Positive	Precision of Negative	Avg.
Yelp	2000	94.42%	61.88%	**78.15%**
Amazon reviews unlocked mobile phones	1895	87.62%	61.24%	**74.43%**
Amazon Reviews of Mobile	541	92.88%	59.54%	**76.21%**

Table 3. Statistic and evaluation result for our system.

From above table we can observed that there is a more difference in negative precision as compared to positive precision value, hence we need to improve more linguistic rules to obtain better negative precision value.

B. review classification

Using the feature word extracted from topic word table 1. We classify the review based on topic or feature of product. Before that we need to give manual touch to topic word, we add and remove some words form topic table as per require to our dataset Amazon reviews unlocked mobile phones. So we achieve some accuracy in classification. Then we analyse same feature word list for second dataset Amazon Reviews.

From first and second dataset we classify 1895 and 541 review respectively and we get the result as in flowing in following table 4.

Class name	Amazon reviews unlocked mobile phones	Amazon Reviews of mobile
Performance	16.36%	18.11%
Features & accessories	8.81%	18.30%
Hardware	18.58%	25.51%
Service	12.77%	25.51%
Price	13.03%	8.87%
Unclassified	30.13%	3.7%

Table 4. Classification result.

By analysing the result of both dataset we get following reason for unclassified reviews:

- Some review contain single word only like "good", "Excellent" etc. which is not in feature word list.
- Some review contain different language and symbols.
- Some review contain incorrect spelling.

V. CONCLUSION AND FUTURE SCOPE

In this paper, a rating prediction model is proposed which performs mining of sentiment information from text reviews. The sentiment score and rating is measured using sentiment dictionary. In the system "and" and "but" linguistic rules are used which improves the result accuracy. To classify the text review feature word list is used which is the output of LDA model. The rating prediction and classification experiment are conducted by three different real time dataset. In the future this work can be enhanced to improve the training dataset of classification i.e. feature word list should also consider the incorrect spelling of the reviews, so that more accurate classification results can be achieved. For sentiment analysis further more linguistic rules can be applied along with the sentence detection method to get the more accurate sentiment score.

REFERENCES

[1] F. Li, N. Liu, H. Jin, K. Zhao, Q. Yang, X. Zhu, "*Incorporating reviewer and product information for review rating prediction*" , in Proceedings of the Twenty-Second international joint conference on Artificial Intelligence, 2011, pp. 1820-1825.

[2] G. Ganu, N. Elhadad, A Marian, *"Beyond the stars: Improving rating predictions using Review text content"*, in 12th International Workshop on the Web and Databases (WebDB 2009). pp. 1-6.

[3] Y. Ren, J. Shen, J. Wang, J. Han, and S. Lee, *"Mutual Verifiable Provable Data Auditing in Public Cloud Storage"*, Journal of Internet Technology, vol. 16, no. 2, 2015, pp. 317-323.

[4] Y. Zhang, G. Lai, M. Zhang, Y. Zhang, Y. Liu, S. Ma, *"Explicit factor models for explainable recommendation based on phrase-level sentiment analysis"*, in proceedings of the 37th international ACM SIGIR conference on Research & development in information retrieval, 2014.

[5] X. Lei, and X. Qian, "*Rating prediction via exploring service reputation"*, 2015 IEEE 17th International Workshop on Multimedia Signal Processing (MMSP), Oct 19-21, 2015, Xiamen, China. pp.1-6.

[6] B. Sarwar, G. Karypis, J. Konstan, and J. Reidl, *"Item-based collaborative filtering recommendation algorithms",* in Proc.10th International Conference on World Wide Web, 2001, pp. 285-295.

[7] K.H. L. Tso-Sutter, L. B. Marinho, L. Schmidt-Thieme, *"Tag-aware recommender systems by fusion of collaborative filtering algorithms",* in Proceedings of the 2008 ACM symposium on Applied computing, 2008, pp. 1995-1999.

[8] R. Salakhutdinov, and A. Mnih, *"Probabilistic matrix factorization"* , in NIPS, 2008.

[9] X. Qian, H. Feng, G. Zhao, and T. Mei, *"Personalized recommendation combining user interest and social circle"*, IEEE Trans. Knowledge and data engineering. 2014, pp. 1763-1777.

[10] L. Qu, G. Ifrim, G. Weikum, *"The bag-of-opinions method for review rating prediction from sparse text patterns",* in Proc. 23rd International Conference on Computational Linguistics, 2010, pp. 913–921.

[11] K. Zhang, Y. Cheng, W. Liao, A. Choudhary, *"Mining millions of reviews: a technique to rank products based on importance of reviews",* in Proceedings of the 13th International Conference on Electronic Commerce, Aug. 2011, pp. 1-8.

[12] B. Pang, Bo, L. Lee, and S. Vaithyanathan, *"Thumbs up? sentiment classification using machine learning techniques",* in Proc. EMNLP, 2002, pp. 79-86.

[13] D. Tang, Q. Bing, T. Liu, *"Learning semantic representations of users and products for document level sentiment classification",* in Proc. 53th Annual Meeting of the Association for Computational Linguistics and the 7th International Joint Conference on Natural Language Processing, Beijing, China, July 26-31, 2015, pp. 1014–1023.

[14] T. Nakagawa, K. Inui, and S. Kurohashi, *"Dependency tree-based sentiment classification using CRFs with Hidden Variables",* NAACL, 2010, pp.786-794.

[15] Xiaojiang Lei, Xueming Qian, Member, IEEE, and Guoshuai Zhao, *"Rating Prediction based on Social Sentiment from Textual Reviews",* IEEE TRANSACTIONS ON MULTIMEDIA, MANUSCRIPT ID: MM-006446.

[16] D.M. Blei, A. Y. Ng, and M. I. Jordan, *"Latent Dirichlet Allocation",* Journal of machine learning research3. 2003, pp. 993-1022.

[17] Baizhang Ma, Dongsong Zhang, Zhijun Yan, Taeha Kim *"AN LDA AND SYNONYM LEXICON BASED APPROACH TO PRODUCT FEATURE EXTRACTION FROM ONLINE CONSUMER PRODUCT REVIEWS"* Journal of Electronic Commerce Research, VOL 14, NO 4, 2013

[18] W. Xiong, Y. Jin, Z. Liu, *"Chinese sentiment analysis using appraiser-degree-negation combinations and PSO",* Journal of Computers, 2014, 9(6).

[19] H. Nie, Z. Rong, *"Review helpfulness prediction research based on review sentiment feature sets",* New Technology of Library and Information Service, 2015, 31(7): 113-121.

[20] https://www.cs.uic.edu/~liub/FBS/sentiment-analysis.html #lexicon

[21] http://jmcauley.ucsd.edu/data/amazon

[22] https://blog.algorithmia.com/lda-algorithm-classify-text-documents/

[23] http://times.cs.uiuc.edu/~wang296/ Data/LARA/Amazon/AmazonReviews.zip

Authors Profile

Mr. Milind M. Sutar pursed Bachelor of Engineering from Shivaji University, Kolhapur in year 2015, He is currently pursuing Master of Technology from D.K.T.E'S Textile and Engineering Institute, Ichalkaranji. His main research work focuses on classification of sentiment analysis, machine learning, network and security.

Mr. Tanveer I. Bagban pursed Master of Engineering from Shivaji University, Kolhapur in year 2007.His area of research is Web Mining.He is currently pursuing Ph.D in Computer Science and Engineering.and currently working as Associate Professor in Department of Information Technology, D.K.T.E'S Textile and Engineering Institute, Ichalkaranji. He has published more than 20 research papers in peer reviewed reputed International Journals. He worked as expert faculty to teach course Operating System at Busitema University Uganda, Africa. He worked as Board of Studies Member for Rajarambapu Institute of Technology, Islampur and member of Departmental Academic Advisory Board for Annasaheb Dange college of Engg. and Tech. Ashta.