ASIAN JOURNAL OF PHARMACEUTICAL AND CLINICAL RESEARCH

INNOVARE ACADEMIC SCIENCES
Knowledge to Innovation

Special Issue (April)

Online - 2455-3891
Print - 0974-2441

Full Proceeding Paper

A SURVEY ON BUSINESS FACILITIES ANALYSIS

JOSHIKA S, PARVATHI R

School of Computing Science and Engineering, VIT University, Chennai, Tamil Nadu, India. Email: joshika.subburaj@gmail.com

Received: 7 March 2017, Revised and Accepted: 8 March 17

ABSTRACT

The objective is to recommend the feature on which a restaurant has to concentrate for becoming a top rated company in the future. The Yelp dataset has been used in this work. The N-gram method is used to find the features from the reviews of the user, reviews are classified using k-means clustering algorithm, and the features are ranked based on the reviews sentiment. The content-based recommendation will be used for making the recommendation of feature that has to be concentrated. The result will contain the feature on which the 1 or more restaurant has to concentrate that increases the chances of becoming a top rated company in the future. This will help a restaurant to concentrate on a specific feature for getting preferred by the customers.

Keywords: Yelp, Business analysis, Sentiment classification, Sentiment analysis, Valence aware dictionary and sentiment reasoner intensity, Prediction, Suggestion, Best facility.

 DOI: http://dx.doi.org/10.22159/ajpcr.2017.v10s1.19990

INTRODUCTION

Yelp is an American multinational corporation that developed Yelp.com social networking site which contains the details regarding local businesses in the USA. User reviews play a vital role in web services such as Amazon, Yelp, Epinions users can post their own opinion over the business, services, and products provided by Yelp.com. Yelp provides free-form text for reviews and also allows the user to rate the business, product, and services. The various forms of analysis and predictions can be made from the reviews. The recommendation engine can be built, or sentiment analysis (SA) can be made to find the users sentiment on any business category. Various papers have already been written by the authors based on Yelp data challenges, and it has been awarded by the Yelp. Most of the papers explain about the restaurant reviews and the review versus rating. Other papers were mostly based on the distribution of the reviews, relation between the reviews, reviews in 90s versus recent reviews, intensity changes in the reviews, etc.

LITERATURE SURVEY

Max Woolf [1] has explained that the positive reviews will have more positive words and the negative reviews will have less positive words. The rating and review of the Yelp users will have a high correlation. The comparison between the two, three-word phrases of 1-star and 5-star rating has an apparent difference. Positivity and negativity over the past few years are also explained with graphs.

Fig. 1 shows the distribution of reviews in the dataset. The huge number of reviews will give a good accuracy for prediction. This makes the participants to select the restaurant business for the analysis and prediction instead of any other business categories. The technical report by the team Sajnani *et al.* [2] from Yelp challenge have explained that the classification can be made for any business category which helps user to make a clear decision on choosing a product or service based on their personal preferences.

Chandrakala and Sindhu [3] have explained that the sentiment classification is different from the SA where the SA will only give sentiment polarity (whether the opinion is positive or negative) whereas the sentiment classification will give both sentiment polarity and sentiment intensity (whether the positive or negative sentiment is mild or strong) of a sentence. In the Yelp dataset, many analyses have been made on the restaurant reviews than any other business categories. The review count of restaurant business is high than any other business category. Asghar [4] has explained that in the all over the distribution of business categories the restaurant category takes the highest percentage compared with other business categories and the distribution of review is 68.32% overall for restaurants.

Fig. 2 shows the normal view of the Yelp site in which [2] have applied multi-label classification to the reviews of restaurant data where the facilities expected for the restaurant business are identified, and these facilities can be added to the site as a filter for the results. Fig. 3 shows the sample page after adding the facilities as filter which can help the user to easily access the products, service or business based on their personal preference. Li and Jiang [5] have proposed that semi- supervised sentiment classification method will build a classifier based on the co-training framework which consists a minimum number of labeled and unlabeled instances. They also found that this method performs better than the self-learning support vector machine (SVM) and Naive Co-training SVM.

Sahu and Ahuja [6] have extracted extra new features from the movie reviews with the polarity of the reviews and by applying computational linguistic for pre-processing the review data. Among Naive Bayes, random forest, decision tree, COCR, Bagging, and KNN, Random Forest performed with high accuracy. Ashok *et al.* [7] have worked on limiting the results retrieved for a query. A social framework was proposed for processing the reviews to personalize and rank suggestions based on user's preference which performed faster in service and removing irrelevant data results. Emelda [8] applied Sentiment classification based on the keywords, emoticons, and SentiWordNet. The classifiers with sentiment analyzers are compared and finally verified with Weka tool's result. Maruthi *et al.* [9] topic modeling, topic modeling using Latent Dirichlet Allocation (LDA), aspect-based sentiment scoring methods on the consumer reviews about food businesses are used to analyze the association between star ratings and reviews which show that some aspects significantly have influence over the ratings. Asghar *et al.* [10] have performed a review on existing methods of feature extraction in SA and mining opinion that analyzes over the areas on which the researchers are most and least interested. Marx and Yellin-Flaherty [11] addressed the problem of aspect-specific SA,

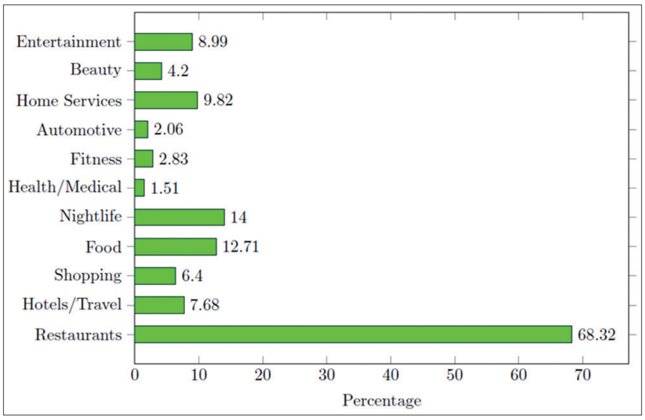

Fig. 1: Distribution of review in Yelp dataset

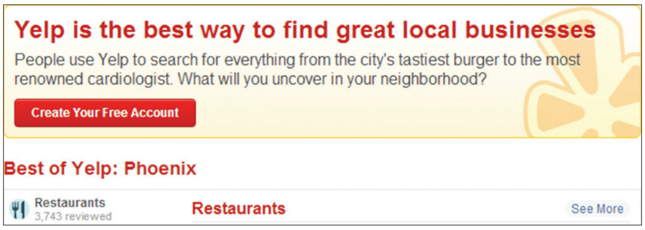

Fig. 2: The normal Yelp site for restaurant category

Fig. 3: The modified Yelp site for restaurant category

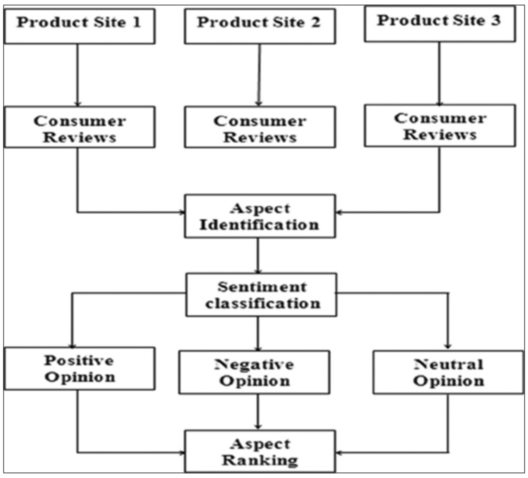

Fig. 4: Ranking product aspects architecture

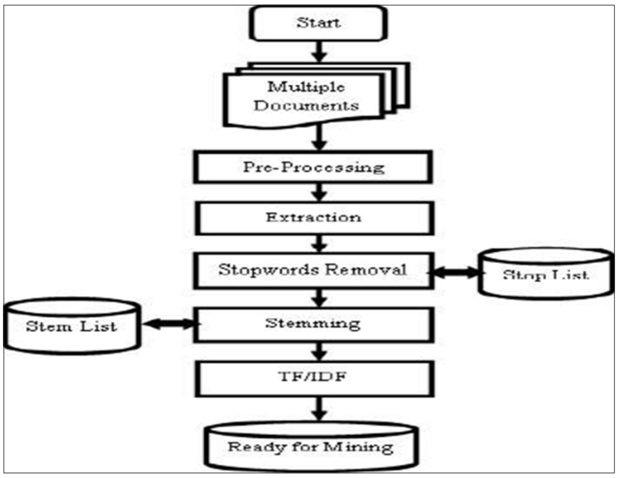

Fig. 5: The steps of pre-processing and text mining

and the methods have to be applied not only to find the aspects but also should find the sentiment of that aspect in the user's point of view. Join aspect sentiment model, and separate aspect sentiment model methods are used where the recurrent and recursive neural networks are compared. When a less amount of data is considered recurrent neural network performs better than recursive neural networks. Nicholls and Song [12] have explored different methods for feature selection, and a new method for SA was also proposed. The document frequency difference method proposed in this paper was observed to perform better for SA. Suganya [13] has implemented the product aspect ranking architecture for a single product from the customer reviews of different sites. Fig. 4 shows the architecture used in Londhe [13] for the process of product aspect ranking.
Tripathi and Naganna [14] have proposed a new method which combines NLP and supervised learning techniques. For feature selection, N-gram algorithm is used where 4-grams (i.e., n=1, 2, 3, 4) is used in this work. It shows that bigram is providing efficient results with accuracy and precision as measures, but the recall of the bigram is same to that of unigram. Okeefe and Koprinska [15], the classifiers Naive Bayes and SVM is used in evaluating feature selectors and feature weights. New feature selection and feature weighting methods are also introduced in this work.

Vijayarani *et al*. [16] have discussed the pre-processing and text mining techniques. The processing techniques are depicted in Fig. 5. Gupte *et al*. [17] have carried out the analysis to decide which algorithm has to be used in a particular preference and also analysis the accuracy of the algorithms. Some of the widely used algorithms in SA are Naive Bayes, max entropy, boosted trees; random forest when compared if accuracy is given priority then random forest can be preferred. Naive Bayes gives a better result when compared to other machine learning classifiers when considering processing power and memory. To process in less training time with a powerful processing system; max entropy can be used otherwise for average performance in all aspects; boosted trees will suit better than any other algorithms.

Wawre and Deshmukh [18] applied sentiment classification techniques to movie review dataset, and two supervised machine learning approaches SVM, and Naive Bayes are compared on the same data. The result in this paper states that Naive Bayes performed better than SVM. Elawady *et al.* [19] tried to provide an optimized tool for feature selection. The methods used in this process are a rough set theory (RST), decision rules, minimum redundancy maximum relevance, and machine learning algorithms (SVM, Naive Bayes). All methods are tested on Movie, Book, DVD, and electronic datasets. A Hybrid method which is developed based on the RST and information gain performed better than the other methods. Kong *et al.* [20], have identified the key features people in different countries look for in their dining experience by applying classification models such as Naive Bayes, SVM, decision trees, logistic regression, and Gaussian discriminant analysis on Yelp dataset. The test accuracy of natural language processing method was high than any other methods.

Gamallo and Garcia [21] describe the polarity of English tweets using naive Bayes. The experiments gave the best performance in binary classifier when there are two polarities. Londhe [22] proposed a product aspect ranking framework to identify the aspects of a product from the on-line consumer reviews which helps in making an easy decisions on product purchase. The reviews are classified and ranked with the probability ranking algorithm, and the result is represented in graphical format. Huang *et al.* [23] points out the demand of customers from the reviews having high dimensionality. This increases the Yelp ratings when concentrated on the demands expected by the user, which directly affects their revenue. Salinca [24] proposed several approaches for automatic sentiment classification. It uses 2 feature extraction methods loop entire data-set to for tokenizing and preprocesses with 4 machine learning models Naive Bayes, Linear support vector classification (SVC), logistic regression, and stochastic gradient descent classifier. It showed the effectiveness of the ensemble methods for reviews sentiment classification. In the performance point of view, Naive Bayes and logistic regression have bad results than linear SVC.

Medhat *et al.* [25] have done a survey on SA techniques which also includes categorizations of methods in SA. The most techniques used in sentiment classification are depicted in Fig. 6. To achieve better performance in both precision and recall Wang *et al.* [26] proposed a method by combining a dynamic and iterative process. In the propagation process, syntactic patterns are used as opinion relations. The experiment is tested on both English and Chinese on-line reviews which proved to be better performing than the other methods.

Sajnani *et al.* [27] have analyzed 10,000 review datasets and divided into 5 bins based on the facility which the review focuses. This was done manually which lead to 225 man hours to complete the reviews which was in the confusion of which bin it falls in was neglected and finally, 9019 reviews were taken for the further process.

Yu *et al.* [28] have tried linear regression, random forest and latent factor model with some text mining techniques to build the features and compared settings of the parameter of every model to get a better performance and to avoid the complex model. The models are then tested on different datasets for comparison of the performance variance and found that random forest performed the best. Table 1 shows the comparison of mean squared error using various algorithms.

de Oliveira *et al.* [29] have done a systematic study on deep learning utilization of humor detection which then derived to recurrent neural networks, and then they examined convolution networks. They have also provided a simple extension. The convolution network built by them has given a best result across all the methods reaching a maximum of 81.57% accuracy.

Tables 2 and 3 show the accuracy achieved by normal convolution network and the Maxout convolution network [29]. Huang *et al.* [30] described latent subtopics which was derived from Yelp restaurant reviews with an online LDA algorithm with this the demand of customers is identified from reviews. This provides useful information regarding the opinion of users on restaurants for increasing the Yelp ratings that also have an impact on the revenue.

The paper [30] has been awarded the first prize by Yelp and this is applied on their Yelp.com in view of better performance to the users as well as for the companies' better improvement. Angulakshmi and Chezian [31] proposed an approach for domain-specific word identifying. The polarity of a word is calculated by using disambiguate process. The experiment achieved a sentiment classification with the unsupervised approach. Naive Bayes is used on review depending on features.

Fig. 7 shows the approach proposed in Spring; 2013 [31] in which the first level of feature extraction mentions seed extraction, the second level mentions conjunction based extraction and the third level mentions double propagation. The preprocessing will contain the stemming as the major process.

Kotzias *et al.* [32] presented a general framework for transforming groups of instances to individual instances which depend on a cost function that can classify by leveraging both instance and group level label information. This shows that embedded process will give better result which was also applied for different datasets.

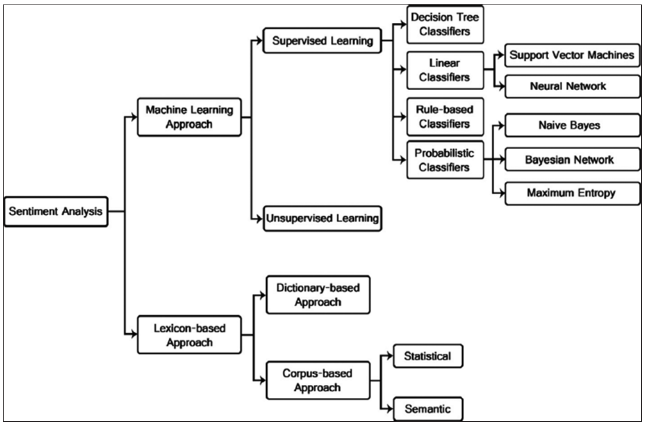

Fig. 6: Sentiment analysis techniques survey

Table 1: Comparison between the algorithms

Model	MSE (test data)
Baseline	1.40981598434
Linear regression	0.79202671108
Random forest regression	0.63987291881
Latent factor model	1.26561688673

Table 2: Accuracy of normal convolution network

Method	Train accuracy	Dev accuracy	Test accuracy
Convolution, word vectors fixed	0.8322	0.8001	0.8010
Convolution, word vectors optimized	0.8331	0.8093	0.8101
Convolution, random word vectors	0.8193	0.7941	0.7986

Table 3: Accuracy of Maxout convolution network

Method	Train accuracy	Dev accuracy	Test accuracy
Maxout convolution, word vectors fixed	0.8329	0.8102	0.8114
Maxout convolution, word vectors optimized	0.8539	0.8115	0.8157
Maxout convolution, random word vectors	0.8010	0.7977	0.7912

RESULT AND DISCUSSION

This survey paper gives a clear picture of insights about the Yelp business data-set, the steps to be involved in text pre-processing and sentiment classification. This proposed work which is a hybrid algorithm that will classify the reviews into different bins based on the facility it focuses on, to predict the future top rated company, and to suggest the facility on which a company has to concentrate. The steps to be followed in the work are as follows:

The review data have to be first processed using NLTK valence aware dictionary and sentiment reasoner (VADER) intensity analyzer which will give the value of the intensity of review that varies between 0 and 1 where,

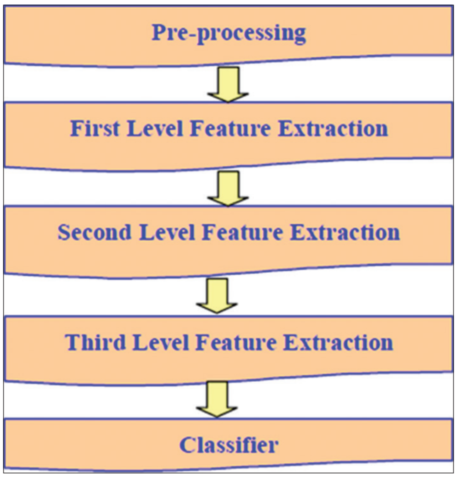

Fig. 7: Three level feature extraction

Restaurant	3.5	pos
Restaurant	3.5	pos
Restaurant	3.5	pos
Restaurant	3.5	neg
Restaurant	2.0	pos
Restaurant	2.0	neg
Restaurant	2.0	s_pos
Restaurant	2.0	neg
Restaurant	2.0	s_neg
Restaurant	2.0	neg
Restaurant	2.0	s_pos
Restaurant	2.0	s_pos
Restaurant	2.0	pos
Restaurant	2.0	pos

Fig. 8: Valence aware dictionary and sentiment reasoner intensity result for a set of reviews

Value >0.5 -> "positive"
Value <0.5 and >0.1 -> "somewhat positive"
Value <0.1 and >–0.1 -> "neutral"
Value <–0.1 and >–0.5 -> "somewhat negative"
Value <–0.5 -> "negative."

Fig. 8 shows the reviews classified based on the value generated by the VADER intensity analyzer. Then the data are pre-processed using the following steps:

1. POS tagging.
2. URL removal.
3. Repetition removal.
4. Punctuation removal.
5. Stop words removal.
6. Stemming.
7. Lemmatization.

Fig. 9 shows the result of the Unigram generated for the preprocessed reviews from 13001 to 16000. The steps for the classification of reviews into different bins based on the facility are as follows:

1. Identifying the facilities that matters a lot from the reviews by SA.
2. Generate the common term of different words that points to one facility.
3. Classify the reviews into different bins based on the common terms available in the review and neglect the reviews that do not have a proper common term.

The prediction of the future top rated company has the following steps:

1. Separate the reviews of every company into different bins.
2. Rank the facility based on the reviews of the top rated companies.
3. Find the company that has good rating and performs better on all the facilities at the most equivalent to the current top rated company.

The suggestions of the facility on which the company should concentrate have the following steps:

1. Compare the top-rated companies' facility rank with the predicted companies facility rank.
2. Identify the facility from the suggestions provided by the consumer.
3. Compare the facility identified from the suggestion with current top rated company facility rankings and identify the facility that should be concentrated.

CONCLUSION AND FUTURE WORK

The literature survey gives a clear picture on the steps to be followed for working on the proposed work. Ranking the facilities from reviews using SA was achieved. The future top rated company prediction and the suggestion of facility on which the company have to concentrate will be done in future with the number of times a user visited the restaurant and the location of the restaurant as additional considerations which helps in increasing the performance.

```
processed 13000: 2016-11-05 08:29:38.329184
[('food', 963), ('place', 858), ('time', 490), ('sandwich', 414), ('order', 371), ('sauc', 357), ('restaur', 337), ('servic', 326), ('menu', 320), ('chicken', 293), ('p
izza', 282), ('lunch', 260), ('dish', 253), ('rice', 196), ('bean', 186), ('taco', 184), ('salad', 180), ('peopl', 178), ('day', 176), ('bianco', 175), ('way', 172), ('
tabl', 165), ('meal', 162), ('review', 160), ('thing', 156), ('burrito', 155), ('shrimp', 154), ('lot', 153), ('meat', 152), ('phoenix', 150), ('drink', 150), ('star',
148), ('salsa', 147), ('flavor', 147), ('bread', 146), ('friend', 144), ('chees', 137), ('side', 132), ('bit', 132), ('someth', 131), ('minut', 130), ('price', 130), ('
chip', 126), ('year', 126), ('tast', 125), ('tomato', 124), ('spici', 122), ('locat', 122), ('experi', 121), ('water', 119), ('everyth', 117), ('hour', 113), ('bbq', 10
8), ('beef', 105), ('server', 102), ('pork', 99), ('buffet', 98), ('style', 96), ('home', 95), ('soup', 94), ('seafood', 94), ('area', 94), ('dinner', 93), ('plate', 93
), ('noth', 92), ('staff', 92), ('famili', 90), ('qualiti', 90), ('mozzarella', 89), ('tortilla', 88), ('kind', 86), ('item', 85), ('night', 84), ('bar', 80), ('pane',
79), ('macayo', 79), ('custom', 78), ('cevich', 78), ('anyth', 78), ('waitress', 76), ('one', 75), ('hous', 75), ('wing', 74), ('fish', 73), ('week', 72), ('option', 72
), ('park', 66), ('ingredi', 66), ('pepper', 65), ('spot', 64), ('bite', 64), ('door', 64), ('reason', 63), ('line', 63), ('street', 62), ('appet', 60), ('waiter', 60),
 ('fan', 60), ('onion', 60), ('everyon', 59)]
processed 14000: 2016-11-05 08:29:38.371570
[('place', 920), ('food', 914), ('time', 606), ('servic', 410), ('beer', 349), ('pizza', 343), ('order', 304), ('restaur', 281), ('chicken', 270), ('lunch', 254), ('men
u', 252), ('sandwich', 242), ('salad', 242), ('dish', 240), ('friend', 235), ('tabl', 226), ('bar', 223), ('thing', 203), ('peopl', 190), ('day', 180), ('drink', 179),
('sauc', 175), ('bread', 174), ('night', 166), ('price', 166), ('star', 162), ('year', 158), ('minut', 157), ('server', 157), ('meal', 156), ('way', 155), ('someth', 15
3), ('chip', 148), ('staff', 148), ('hour', 145), ('phoenix', 143), ('lot', 143), ('soup', 139), ('wait', 131), ('dinner', 127), ('review', 126), ('side', 122), ('roll'
, 121), ('bit', 120), ('noodl', 117), ('chees', 114), ('beef', 113), ('oregano', 111), ('pub', 110), ('experi', 110), ('locat', 109), ('spici', 109), ('home', 106), ('n
oth', 103), ('everyth', 101), ('flavor', 99), ('area', 99), ('pasta', 97), ('tast', 91), ('chili', 88), ('famili', 88), ('appet', 86), ('rice', 84), ('meat', 83), ('kin
d', 83), ('burger', 82), ('portion', 82), ('egg', 80), ('fri', 78), ('qualiti', 78), ('shrimp', 76), ('park', 76), ('wing', 75), ('hous', 74), ('plate', 73), ('visit',
72), ('potato', 72), ('tomato', 72), ('name', 71), ('sushi', 71), ('one', 69), ('cooki', 68), ('custom', 68), ('room', 67), ('bowl', 67), ('waitress', 67), ('bianco', 6
6), ('anyth', 66), ('select', 65), ('today', 64), ('onion', 64), ('pie', 64), ('spot', 63), ('stuff', 62), ('everyon', 61), ('coupl', 60), ('crust', 60), ('option', 60)
, ('week', 59), ('fan', 58)]
processed 15000: 2016-11-05 08:29:38.414894
[('place', 907), ('food', 890), ('time', 490), ('restaur', 395), ('servic', 380), ('chicken', 380), ('menu', 283), ('steak', 275), ('pizza', 266), ('order', 249), ('sau
c', 239), ('meal', 221), ('salad', 219), ('thing', 206), ('tabl', 192), ('dish', 188), ('night', 185), ('year', 184), ('price', 181), ('way', 176), ('lunch', 176), ('da
y', 173), ('dinner', 172), ('friend', 170), ('tart', 169), ('star', 163), ('potato', 157), ('side', 157), ('sandwich', 145), ('someth', 143), ('phoenix', 141), ('review
', 141), ('ramen', 141), ('flavor', 135), ('meat', 132), ('soup', 132), ('wing', 131), ('lot', 130), ('chees', 130), ('peopl', 129), ('server', 128), ('home', 126), ('b
eef', 123), ('everyth', 120), ('noodl', 120), ('wine', 116), ('bit', 115), ('plate', 114), ('dessert', 111), ('staff', 109), ('waitress', 108), ('portion', 108), ('brea
d', 102), ('fri', 101), ('experi', 101), ('noth', 100), ('onion', 100), ('bar', 99), ('roll', 99), ('minut', 94), ('deli', 91), ('drink', 86), ('husband', 85), ('gravi'
, 84), ('owner', 84), ('pork', 83), ('appet', 83), ('kind', 83), ('top', 82), ('tast', 81), ('scott', 80), ('hour', 77), ('hous', 77), ('anyth', 75), ('qualiti', 75), (
'famili', 74), ('everyon', 73), ('rice', 71), ('sushi', 71), ('spici', 71), ('entre', 71), ('cream', 70), ('option', 69), ('crust', 68), ('cours', 68), ('rib', 68), ('i
tem', 67), ('area', 66), ('egg', 65), ('wall', 65), ('burger', 65), ('texa', 64), ('one', 64), ('week', 64), ('texaz', 64), ('pastrami', 61), ('bowl', 60), ('grill', 60
), ('coup', 60), ('room', 59)]
processed 16000: 2016-11-05 08:29:38.458970
```

Fig. 9: Unigram generated for the pre-processed reviews

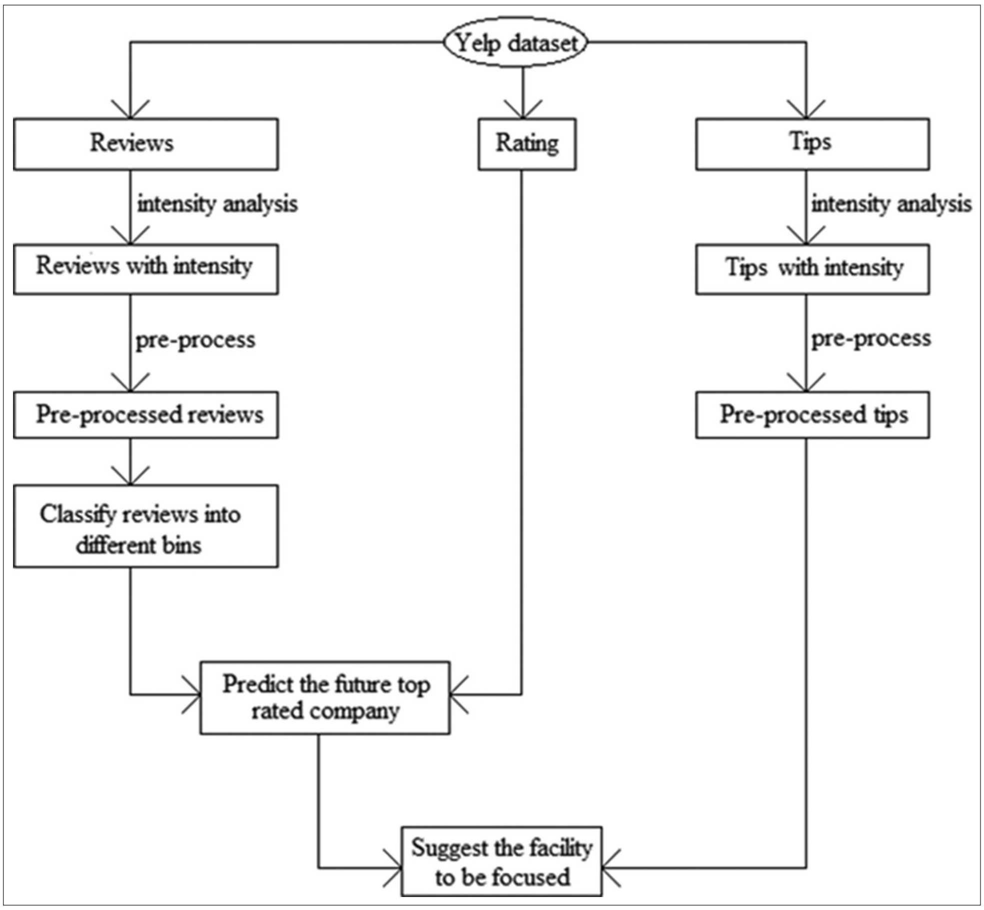

Fig. 10 : The flow of the proposed work

REFERENCES

1. Chandhrakala S, Sindhu C. Opinion mining and sentiment classification: A survey. ICTACT J Soft Comput 2012;3(1):420-5.
2. Asghar MZ, Khan A, Ahmad S, Kundi FM. A review of feature extraction in sentiment analysis. J Basic Appl Sci Res 2014;4(3):181-6.
3. Li S, Jiang Y. Semi-supervised sentiment classification using ranked opinion words. Int J Database Theory Appl 2013;6(6):51-62.
4. Emelda C. A comparative study on sentiment classification and ranking on product reviews. Int J Innov Res Adv Eng 2014;1(10):367-71.
5. Nicholls C, Song F. Comparison of feature selection methods for sentiment analysis. Advances in Artificial Intelligence. Ottawa: LNCS; 2010. p. 286-9.
6. Suganya R. Volume 3 Issue I, January; 2015, ISSN: 2321-9653.
7. Tripathi G, Naganna S. Feature selection and classification approach for sentiment analysis. Mach Learn Appl Int J 2015;2(2):1-16.
8. Vijayarani S, Ilamathi J, Nithya. Preprocessing techniques for text mining - An overview. Int J Comput Sci Commun Netw ???;5(1):7-16.
9. Gupte A, Joshi S, Gadgul P, Kadam A. Comparative study of classification algorithms used in sentiment analysis. Int J Comput Sci and Inf Technol 2014;5(5):6261-4.
10. Wawre SV, Deshmukh SN. Sentiment classification using machine learning techniques. Int J Sci Res 2016;5(4):819-21.
11. Elawady RM, Barakat S, Elrashidy NM. Different feature selection for sentiment classification. Int J Inf Sci Intell Syst 2014;3(1):137-50.
12. Gamallo P, Garcia M. Citius: A Naive-Bayes Strategy for Sentiment Analysis on English Tweets; 2014. p. 171-5.
13. Londhe DR. Product aspect ranking on the consumers reviews. Int J Adv Res Innov Ideas Educ 2016;2(2):???.
14. Medhat W, Hassan A, Korashy H. Sentiment analysis algorithms and applications: A survey. Ain Shams Eng J 2014;5(4):1093-113.
15. Angulakshmi G, Chezian RM. Three level feature extraction for sentiment classification. Int J Innov Res Comput Commun Eng 2014;2(8):5501-7.
16. Sahu TP, Ahuja S. Sentiment Analysis of Movie Reviews: A Study on Feature Selection Classification Algorithms, International Conference on Microelectronics, Computing and Communications, July; 2016.
17. Ashok M, Rajanna S, Joshi PV. A Personalized Recommender System using Machine Learning Based Sentiment Analysis over Social Data, Students conference on Electrical, Electronics and Computer Science; 2016.
18. Prithivirajan M, Lai V, Shim KJ, Shung KP. Analysis of Star Ratings in Consumer Reviews, IEEE International Conference on Big Data (Big Data); 2015.
19. OKeefe T, Koprinska I. Feature Selection and Weighting Methods in Sentiment Analysis, Proceedings of the 14th Australasian Document Computing Symposium, Sydney, Australia, 4 December; 2009.
20. Salinca A. Business Reviews Classification Using Sentiment Analysis, 17th International Symposium on Symbolic and Numeric Algorithms for Scientific Computing, 2016 IEEE.
21. Wang H, Zhang C, Yin H, Wang W, Zhang J, Xu F. A Unified Framework for Fine-Grained Opinion Mining from Online Reviews, 49th Hawaii International Conference on System Sciences; 2016.
22. Kotzias D, Denil M, De Freitas N, Smyth P. From Group to Individual Labels using Deep Features, Proceedings of the 21th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, August; 2015.
23. Asghar N. Caornell University, May, 2016.
24. Marx E, Yellin-Flaherty Z. Aspect specific sentiment analysis of unstructured online reviews. ???: Stanford University; 2015.
25. Kong A, Nguyen V, Xu C. Stanford University; 2016.
26. Huang J, Rogers S, Joo E. Stanford University; 2013.
27. Yu M, Xue M, Ouyang W. UC San Diego Jacobs School of Engineering, Jan; 2016.
28. de Oliveira L, Alfredo LÌ, Rodrigo A. Stanford university, June; 2015.
29. Available from: http://www.minimaxir.com/2014/09/one- starfive-stars.
30. Available from: http://www.ics.uci.edu/vpsaini.
31. Spring; 2013. Available from: https://www.yelp.com/html/pdf/Yelp DatasetChallengeWinner_ImprovingRestaurants.pdf.