sources by using probability based selection process. As nectar amount in food source increases, probability value with which it is preferred by onlooker's increases similar to natural selection process in evolutionary algorithms.

The artificial bee colony contains three groups: scouts, onlooker bees and employed bees. The bee carrying out random search is known as scout. The bee which is going to the food source which is visited by it previously is employed bee. The bee waiting on the dance area is an onlooker bee. The onlooker bee with scout also called unemployed bee.

In the ABC algorithm, the collective intelligence searching model of artificial bee colony consists of three essential components:

- Employed bees
- Unemployed foraging bees
- Food sources

The employed and unemployed bees search for the rich food sources around the hive. The employed bees store the food source information and share the information with onlooker bees. The number of food sources is equal to the number of employed bees and also equal to the number of onlooker bees. Employed bees whose solutions cannot be improved through a predetermined number of trials (that is "limit") become scouts and their solutions are abandoned. Analogously in the optimization context, the number of food sources in ABC algorithm represents the number of solutions in the population. The position of a good food source indicates the position of a promising solution to the optimization problem and the quality of nectar of a food source represents the fitness cost of the associated solution.

In the ABC algorithm, the maximum number of cycles was taken as 2000. The percentages of onlooker bees and employed bees were %50 of the colony and the number of scout bees was selected to be one. The increase in the number of scouts encourages the exploration process while the increase of onlookers on a food source encourages the exploitation process.

In ABC algorithm six benchmark functions are used to optimize the numerical value.

i. Schaffer function.
ii. Rosenbrock function.
iii. Sphere function.
iv. Griewank function.
v. Rastrigin function.
vi. Ackley function.

ABC algorithm being a powerful optimization technique and is widely used for solving combinatorial optimization problems. Hence, this method is incorporated for optimizing the feature subset selection in this investigation. Using sentiment analysis, features from text are extracted, and classified those providing opinions/sentiments about text/data/documents through Support Vector Machine classifier. The ABC algorithm is used for classification to improve the accuracy of the classifier with BOW and BON feature.

Sphere benchmark function is used for optimization of best result of the classification.

$$f(x) = \sum_{i=1}^{D} x_i^2 \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots (1)$$

Where the initial range of x is $[-100,100]^D$. The minimum solution of the sphere function is $\vec{x} = [0,0...,0]$ and $f(x)$. Fitness functions have been designed for the extraction of best classification results of weight of sentiments.

4.6 Evaluation of Performance

Evaluation of performance can be calculated by precision, recall, F-measure, accuracy.

4.6.1 Precision

In a classification task, the precision for a class is the number of true positives (i.e. the number of items correctly belonging to the positive class) divided by the total number of elements belonging to the positive class (i.e. the sum of true positives and false positives, which are items incorrectly labeled as belonging to the class).

$$Precision = \frac{TruePositive}{TruePositive + FalsePositive} \quad \ldots(2)$$

4.6.2 Recall

Recall in this context is defined as the number of true positives divided by the total number of elements that actually belong to the positive class (i.e. the sum of true positives and false negatives, which are items which were not labeled as belonging to the positive class but should have been).

$$Recall = \frac{TruePositive}{TruePositive + FalseNegative} \quad \ldots\ldots.(3)$$

4.6.3 F-measure

A measure that combines precision and recall is the harmonic mean of precision and recall, the traditional F-measure or balanced F-score.

$$F - measure = 2 \times \frac{Precision \times Recall}{Precision + Recall} \quad \ldots\ldots(4)$$

4.6.4 Accuracy

The accuracy is the proportion of true results among the total number of cases examined. To make the context clear by the semantics, it is often referred to as the "rand accuracy". It is a parameter of the test.

$$Accuracy = \frac{TP + TN}{TP + FP + TN + FN} \quad \ldots\ldots\ldots\ldots\ldots(5)$$

5. Result and Discussion

The proposed software efficiently calculates the weight of sentiments. The software will give the better accuracy, result with weighted sentiment, improved classification of sentiment.

Hence very limited work has been done the field of increasing

the accuracy of the classification of sentiment opinion's or reviews. Optimal feature selection is used for reducing feature subset size and computational complexity thereby increasing the classification accuracy. The ABC algorithm being a powerful optimization technique and is widely used for solving combinatorial optimization problems.

Hence, this method is incorporated for optimizing the feature subset selection in this investigation. Experiment results evaluated feature selection techniques based on BOW, BON and for classification ABC algorithm with sphere benchmark function is used. Experimental results show that the classification results of the sentiment using ABC algorithm as compared to SVM in figure 2.

System classifies the weight of sentiments from both SVM binary classifier and ABC nature classifier. Actual polarity or weights of sentiments or reviews are known by us. We test the system with 20 sentiments will give the correct weight or not. Out of 20 sentiments SVM classifier with BOW features and ABC classifier with BOW features gives the 12 sentiments are correctly classified, SVM classifier with BON feature give the 11 sentiments are correctly classified and at last ABC classifier with BON features classified the 14 sentiments correctly.

Table 1: Classification Results

Classifier	SVM (BOW)	SVM (BON)	ABC (BOW)	ABC (BON)
Precision	1	1	1	1
Recall	0.60	0.55	0.60	0.70
F-measure	0.75	0.70	0.70	0.82
Accuracy	0.60	0.55	0.60	0.70

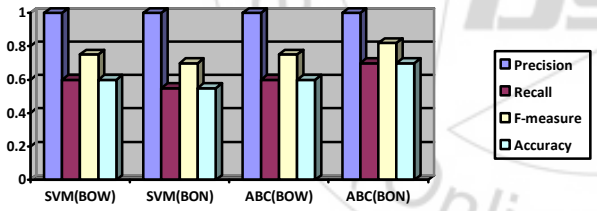

Figure 2: Performance of ABC and SVM classifier on Evaluation

Develop software improves the accuracy of classification using ABC algorithm with BOW and BON features that improve classification accuracy. At last the comparison between SVM (BOW), SVM (BON), ABC (BOW) and ABC (BON) the BON feature selection with ABC classifiers give the better accuracy than the BOW with ABC and SVM classifier feature selection.

6. Conclusion and Future Work

6.1 Conclusion

Sentiment analysis is the process identifying customer sentiments and emotional states. The feelings of the customer can be expressed in positive, negative and neutral ways. Mostly, parts of speech are used as feature to extract the sentiment of the text but we use BOW and BON as a feature of Datasets.

The system computes sentiments polarity. These sentence splits in BOW and for more accurate BON are extracted from the parse data set. Classification with feature BON give the best result as compare to BOW in any classification algorithm.

ABC algorithm is nature inspired algorithm and it give the best optimum solution for the polarity detection of sentiments of reviews of customer. The system has been tested on 20 sample reviews. It has been observed that system has identified sentiments for reviews of people.

The Artificial Bee Colony (ABC) algorithm is a new searching algorithm under Swarm Intelligence technology. Many approaches, methods and goals have been tried out for SA. From the best of our knowledge, previous researches have never applied the ABC algorithm for SA classification. We have compared the proposed algorithm with SVM classifier algorithms which selected from data mining software tools: "Matlab". It has been proved that the proposed ABC data mining algorithm can obtain the better result for weighted (polarity) sentiments. Therefore, we can conclude that the proposed ABC algorithm for SA can obtain competitive result against SVM algorithms and can be considered as useful and accurate classifier.

6.2 Future Work

The proposed system works for simple and short reviews. It can be extended to work for big reviews also. BOW and BON file and their frequencies can be improved by adding more sentiments and reviews in data set. The proposed system use Stanford Parser. If any better parser is available in future classification accuracy of SA has been improved. The system works for short data set. It can be extended to for all type of reviews that is challenges of SA. The proposed system is not web based, so it can be extended as web based application in future. Limitation of the system is not work properly because the data base not so strong, it based on small knowledge base system according to our knowledge base system give can classify the weight of sentiments. It does work well for our data base.

References

[1] Farhadloo. M, Rolland. E, "Multi-Class Sentiment Analysis with Clustering and Score Representation" 13th International Conference on Data Mining Wokshop, IEEE, 2013.
[2] S. Njolstd. P, S. Hoysaeter. L, Wei. W and Atle Gull. J. "Evaluating Feature Sets and Classifiers for Sentiment Analysis of Financial News, " IEEE/WIC/ACM International Joint Conferences on Web Intelligence(W1) and Intelligent Agent Technologies(IAT), IEEE, 2014.
[3] Valakunde. N, Patwardhan. M, "Multi-Aspect and Mutli-Class Based Document Sentiment Analsis of Educational Data Cater in Accreditation Process" International

conference on Cloud & Ubiquitous Computing & Emerging Technologies, IEEE, 2013.
[4] Liu. L, Nie. X, Wang. H, "Toward a Fuzzy Domain Sentiment Ontology Tree for Sentiment Analysis" 5th International Congress on Image and Signal Processing(CISP 2012), IEEE, 2012.
[5] Sumathi. T, Karthik. S, Marikkannan. M, "Artificial Bee Colony Optimation For Feature In Opinion Mining ", Journal of Theoretical and Applied Information Technology, Elsevier, 2012.
[6] Basari. A, Hussin. B, Ananta. I, Junta Zeninarja, "Opinion Mining of Movie Review using Hybrid Method of support Vector Machine and Particle Swarm Optimization" Malaysian Technical Universities Conference on Engineering & Technology 2012, MUCET 2012 Part 4-Information And Communication Technology, Elsevier, 2013.
[7] ManjuS. R, Kalaiman. E, R. Bhavani, "Product Aspect Ranking Using Sentimentic Oriented Sentiment Classifier" IJERA, 2014.
[8] Singh. N, Ghalib. M. R, "An Effective E-Commerce Management using Mining Techniques" International Journal of scientific and Research Publications, Volume 3, Issue 8, IJSPR, 2013.
[9] R. Shikalgar. N, Badgujar. D, "Online Review Mining For Forecasting Sales" IJRET, 2013.
[10] Vigneshkumar K, Gnanavel S, "Mining Online Reviews for Predicting Sales Performance in Movie Domain" TIJCSA, 2013.
[11] S. Modha. J Prof & Head S. Pandi. G , J. Modha. S, "Automatic Sentiment Analysis for Unstructured Data", IJARCSSE, 2013
[12] Patil. G, Galande, V, Kekam. V, Dange. K, "Sentiment Analysis Using Support Vector Machine" IJIRCCE, 2014.
[13] Basiri. M. E, Naghsh-NilchiA. R, AND Ghasem-Aghaee. N, "Sentiment Prediction Based On Dempster-Shafer Theory Of Evidence" HINDAWI, 2014.
[14] Tripathi. G, S. N, "Opinion Mining: A Review" IJICT, 2014.
[15] Joshi. N, Itkat. S, "A Survey on Feature Level Sentiment " IJCSIT, 2014.
[16] Saif. H, He. Y and Alani. H, "Semantic Sentiment Analysis of Twitter", ISWC, 2012.
[17] Patni. S, Wadhe. A, "Reviews Paper on Sentiment Analysis is - Big Challenge" IJARCSMS, 2014.
[18] Varghese. R, "A Survey On Sentiment Analysis and opinion mining" association rules", IJRET, 2013.
[19] Shukran. M, Yeh. W, Wahid. N, Zaidi. A, "Artificial Bee Colony based Data Mining Algorithms for Classification Task" Vol. 5, No 4, August 2011.
[20] Akay. B, Karaboga. D, "A modified Artificial Bee Colony algorithm for real-Parameter optimization " ELSEVIER, 2010.
[21] Murugan. R, Mohan. M, "MODIFIED ARTIFICIAL BEE COLONY ALGORITHM FOR SOLVING ECONOMIC DIPATCH PROBLEM" ARPN, 2012.
[22] Khaze. S, Maleki. I, Hojjatklhah. S, Bagherinia. A, "Evaluation The Efficiency of Artificial Bee Colony and The Fireefly Algorithm in Solving The Continuous Optimization Problem " IJCSA, Vol. 3, No. 4, 2013.
[23] Joshi. N, Itkat. S, "Feature Selection with Chaotic Hybrid Artificial Bee Colony Algorithm based on Fuzzy (CHABCF)" ISPACS, 2013.
[24] Guppu Zhu, Sam Kurong, ELSEVIER, Applied Mathematics and computation 217(2010) 3166-3173, "Gbest- guided artificial bee colony algorithm for numerical function optimization."

Author Profile

Ruby Dhurve received the B.E. degree from Chhattigarh Swami Vivekanand Technical University, Bhilai (C.G.) India in Computer Science & Engineering in the year 2013. She is currently pursuing M.Tech. Degree in Computer Science Engineering with specialization in Computer Science & Engineering from CSVTU Bhilai (C.G.), India. Her research area includes Data Mining and Text Mining etc.

Megha Seth is currently Assistant professor in Department of Computer science & Engineering RCET, Bhilai (C.G.) India. She completed her B.E and M.Tech. in Computer Science and Engineering Branch , experience nine year read RCET Bhilai. Her research area includes Data mining, Image processing, Computer Network, AI & NN etc. She has published many Research Papers in various reputed National & International Journals, Conferences, and Seminars.