Opinion Mining of Tourism Review Using Hybrid Technique of Support Vector Machine and Animal Migration Optimization

Pratishtha Parashar
Department of Computer Science & Engineering and Information Technology
MITS, Gwalior, M.P, India

Sanjiv Sharma
Department of Computer Science & Engineering and Information Technology
MITS, Gwalior, M.P, India

Abstract- Nowadays, online media become one of the powerful sources of information. People visit many places and share their experience on the tourism websites. Tourism websites is a most capable and easy way of information.Opinion mining is a process to extract knowledge from user generated reviews.For proposed method, reviews of visiting place of Gwalior are collected from tourism website and manually pre-classified into class of positive, negative, doubt reviews. Naive bayes is a probabilistic method of classification. Support Vector Machine (SVM) is a classifier that classify into two class with maximum margin. The hybrid method Which combines Animal migration optimization (AMO) and SVM to perform in an optimized way. Performance measures like accuracy, recall, precision are evaluated for justification.

Keywords–Tourism, Opinion Mining, SVM, Naïve Baye's, AMO

I. INTRODUCTION

In earlier day's individuals were just data consumers yet since the appearance of Web 2.0 they assume a more vital part in distributed data on the Web as remarks and surveys. Today's people from each corner of the world share their experiences and views through social media. The web has made it possible for people to discover what others are saying about the brands, places and various other means online, either in mainstream media like online newspapers and magazines, or on social media [1]. Tourists now search for reviews on the web before visit any place. The next step for travellers is finding out whether people are talking positively or negatively about the places to visit, and why. Some online rating system provide a number to represent review but it does not provide actual reason behind the review. People travel different places and express their views in front of millions of people without physical interaction. This is the finest way to express their thoughts and provide initial information to the people who are interested in visiting these places. Tourism websites provide the information about the historical places of different states and countries. Tourism sites advertise the places that are sometimes hidden from the common people. Basically, on this website people give reviews in three ways, positive, negative and doubt or neutral reviews.

Opinion Mining or Sentiment analysis [2] includes building a framework to investigate user's opinions made in blog entries, remarks, audits or tweets, about the item, strategy or a theme. Opinion mining is only finding the feeling of individual from sentences and orders them on the premise of extremity.

Opinion mining is a process to extract knowledge from the opinion of users about objects, entity, item, and event. Opinion mining [3] is a way to analyze the feelings, thoughts of peoples about real world things. Sentiment analysis is trying to analyze the polarity of the sentence. That mean's sentence expressed positive or negative review. There are various sources from where opinions are collected like websites, micro-blog, Twitter etc.

Basically, there is two types of sentences in a document, objective and subjective. Objective sentences contain facts about the world. These sentences do not contain opinion information. Subjective [4] sentences contain reviews, feelings, emotions. An example "Taste of soft drink is good" is a subjective sentence. In this sentence, opinion is expressed about the soft drink. It is not necessary that subjective sentence always shows the opinion. Take an example "I think; she is not well today". This is a subjective sentence but not give any opinion. On the basis of sentiment, reviews express various responses. A review which expresses good sound and positive attitude is called positive review. The negative review contains bad experiences and information about that experience. Doubt reviews are that which expresses a confusing behaviour that's why they are not properly distinguished in a class of positive and negative. Doubt reviews have a potential but because of confusing behaviour they considered as a doubt. Sometimes these reviews have important information that is beneficial for common ones.

The main part of opinion mining task is finding a set of features that are used to describe sentiment information in the document. Some common features are:

- Terms and their frequency- Term is expressed as a 1-gram or n-gram and frequency is described as how many times term occurs in a document. According to the study, simple opinion word in a document is better than term frequency.
- Part-of-Speech- Part-of-speech is an important part of natural language. It includes nouns, pronouns, adjectives, verbs, adverbs, conjunctions etc. These are the best indicators of an opinion.
- Opinion words and phrases- Words which shows opinion itself is considered as an opinion word. Common words like good, best, awesome, wonderful, amazing shows positive response and words like bad, cheap, ugly, horrible represent the negative opinion. Opinion phrases have a potential to express review. Phrases are more pinpointed and need prior knowledge about it.
- Negations- Negations are the word which reverses the meaning of an opinion word. Negation word changed the polarity of review, so there is a need to be handled carefully. For example, “I like this place” vs. “I do not like this place”.
- Other features include rules and regulations of opinions that express an opinion in a multilevel linguistic than a single opinion word.

1.1 N-gram word model

The n-gram word model is a method of finding a set of n-gram words from a given document. Unigram, Bigram, Trigram models have come under n-gram model and these are important features for knowledge extraction. In categorization, classification and text prediction n-gram model are used. Classification task classifies opinion into the class of positive and negative opinion and also for another task of sentiment analysis.

EXAMPLE 1. Unigrams

Text: A young lady looks lovely in a red dress.
Unigrams: *“A”, “young”, “lady”, “looks”, “in “, “a”, “red”,” dress”*

EXAMPLE 2. Bigrams

Text: A young lady looks lovely in a red dress.
Bigrams: *“A young”,” young lady”,” lady looks”, “looks lovely”, “lovely in”, “in a”, “a red”, “red dress”*

EXAMPLE 3. Trigrams

Text: A young lady looks lovely in a red dress.

Trigrams: *"A young lady", "young lady looks", "lady looks lovely", "looks lovely in", "lovely in a", "in a red", "a red dress"*

Unigram model is a simplest model. Document is split into individual words and each individual word is considered as a 1-gram as shown in example 1. In example 1, sentence break into individual word and each word is unigram. Bigrams contain the two words that are adjacent to each other in a sentence. It is a set of all pairs of two words as seen in example 2. Bigrams works as same as unigrams but in a pair of two words [5]. Trigrams contain three words that are adjacent to each other. It is a set of all pair of three words as seen in example 3. Unigrams are important feature but it not gives information properly because sometimes single word is not enough to express information. Bigrams and trigram are more informative than unigram. Bigram and trigram is pair of words so it gives better and meaningful information for sentiment analysis.

1.2 POS Tagging

POS tagging is a method of extracting part-of-speech from document and labelling them according to the category [5]. There are eight part-of-speech categories include; noun, pronoun, adjectives, verb, adverb, conjunctions, prepositions etc. Part-of-speech are the important part of English. They are helpful in better representation of opinion because a single word is not enough to represent an opinion. So sentences are needed to give a brief description of review that means why a user like or dislike a particular entity. So part-of-speech is helpful for this expression. Several techniques are available for detection of part-of-speech from a document. The statistical method is one of the popular methods of detection in which probabilities are calculated according to the words exists next to each other.

II. LITERATURE REVIEW

About some element of interest is a critical bit of data for most clients amid the basic leadership process discovering what others think [6]. Opinion mining is valuable for customers, as well as helps associations to assess sentiments and conduct of customers towards their company and its item, the associations can get surveys about its item straight from customers in long range interpersonal communication, for example, Tweets.

A few supposition mining strategies were discharged either to figure out if input sentences in a characteristic dialect are subjective or objective, or whether they are certain or negative. The primary paper on such an opinion mining end-to-end sentiment mining examination was at that point discharged [7]. These techniques were discharged in light of genuine needs, for example, motion picture surveys, which lead to a business accomplishment of this examination, and the greater part of the principle organizations in item creation taking into account these opinion mining strategies.

Sentiment mining is a most recent study in the piece of Text Mining (TM) that has been particular by various conditions like conclusion investigation, subjectivity examination, or estimation introduction [8]. Sentiment classification is a binary classification method that classifies into two classes [6]. Turney [12] recommended making sense of the bearing of conditions by bootstrapping from a few two negligible arrangements of seeds conditions by monitoring the quantity of hits discounted from internet searcher. Pang et al. [9] found out that machine learning considering overwhelms human-proposed baselines. They utilized Naive Bayes, Maximum Entropy Classification, and Support Vector Machine (SVM) for classification. These are the classification techniques which classify reviews into positive and negative class. These are supervised classifiers so there is need to train for classification. Movie reviews are classified by using these three techniques. After applying classification technique results are come into consideration. On the basis of research Support Vector Machine (SVM) perform best out of all three classification technique. SVM have been utilized effectively as a part of numerous content order studies because of their significant advantages, for example, they are hearty in high point of view zones, and any capacity is suitable, hearty when there is a sporadically set of test and most content order issues are sprightly free. Besides, SVM has

gotten awesome results in assessment mining and these strategies have overpowered other machine learning techniques [8].

Peter D. Turney [12] proposed an unsupervised learning algorithm for classifying reviews as recommended or not recommended. The algorithm used POS tagging to identify adjectives and adverbs, and then uses an algorithm employing Point-wise Mutual Information along with information retrieval to measure the similarity of found opinion-bearing words with known opinion reference words (e.g. excellent for positive).

Morinaga et al. [13] proposed a framework to ease opinion mining, based on automatic opinion labelling in order to assess an entity's reputation. The system was built around extracting characteristic words using stochastic complexity, in order to gain a sense of overall features. For each of these features, words generally found to be co-occurring with the feature word are added to the feature (e.g. \display", \text", \email"). A set of user specified categories is then tagged with meta-scores indicating how likely an opinion is to be expressed in a typical sentence. This is done via Bayesian theory. Finally, the found features are mapped to the opinions of the specified categories via principal component analysis.

Cristian Brucur [14] proposed a system which used unsupervised learning method to classify tourism reviews into positive, negative and neutral. In this system different modules are used like acquisition module, processing module, opinion mining module. At the last performance measures accuracy, precision, recall, F-score are evaluated.

Zhu et al. [15] proposed a system to classify text into a class of positive and negative review. They use an i-model prototype based on ANNs. This research contains prior knowledge, feature weighting etc. Results of this system is better than Support Vector Machine and hidden Markov Model classifiers on corpora of a movie review.

Various researchers used hybrid techniques for enhancement of the classifiers. Some optimization techniques also come into consideration for better results of classifiers. Abd. Samad Hasan Basari et al. [16] proposed a technique in which Particle swarm optimization (PSO) technique and Support Vector Machine are combined. SVM is best classifier and PSO is an optimization technique hence the combination of these technique give better result. Result of this system as comparison to SVM is good.

Doubt reviews have a greater potential and it contain important information that are beneficial for common ones. Doubt reviews have a confusing behaviour so they are miscategorised by classifier. So, sometimes performance is degraded because of unstable behaviour. Hence there is need to categorize doubt reviews into positive and negative so the information inside it is useful for users.

III. CLASSIFICATION METHODS

The classification method is used to classify text into pre-defined classes on the basis of features present in a text. Classification is an important task in the field of opinion mining. Opinions are classified in a class by using classification technique. Major classes of classification technique are supervised learning method and unsupervised learning method. The supervised technique uses training and testing phase. There is various technique come under the category of supervised learning like Naïve Baye's, Support Vector Machine, Maximum Entropy, Decision tree and so on.

Naïve Baye's

Naïve Baye's [10] classifier is a supervised technique of classification. It works on the concept of Baye's rule. Firstly, a classifier is trained by using training dataset. By using training dataset classifier learns the patterns and after training classifier predicts results for an unknown problem. Naïve Baye's classifier is a probabilistic model and works on the concept of probability. Probability decides that how much chance has a text to belong to a particular

class. This classification technique separates reviews into two categories namely positive and negative. This is an efficient technique and easy to handle.

Support Vector Machine

Support Vector Machine (SVM) [11] is a machine learning algorithm come in the supervised category of classification methods. SVM looks for the maximum marginal hyper plane which separates text between two class. SVM is trained by training dataset and after training SVM function is created. SVM function is used to for classification. SVM function assign class to review namely positive class and negative class. Support vectors are the instances which lie on the boundary of a marginal plane. SVM perform better than other classification technique and accuracy of classification is better than other algorithms. SVM classify reviews in positive and negative class but there are some instances that lie on the boundary and not separate because of confusing behaviour. Reviews that are on the boundary and miscategorised are called doubt reviews. There is a need to separate doubt reviews and give the category to it.

Decision Tree

It is a tree in which inner hubs are spoken to by highlights, edges speak to tests to be done at highlight weights also, and leaf hubs speak to classifications which results from above tests. It sorts an archive by beginning at the tree root and moving effectively descending by means of the branches (whose conditions are fulfilled by the archive) until a leaf hub is come to. The archive is then grouped in the classification that marks the leaf hub. Choice Trees have been utilized as a part of numerous applications in discourse and dialect preparing.

IV. RESEARCH DESIGN

Data is collected from tourism website https://www.tripadvisor.in/. All the reviews and comments of historical places of Gwalior like Gwalior Fort, Sun Temple, Sahastrabahu Temple, Jai Vilas Palace, Gurudwara etc. are available on the tourism website. Visitors come from various places to visit these historical places and share their views and experiences on tourism websites. From tourism website 700 reviews are collected. Now, in pre-processing data cleaning and filtering of data is held. All the unwanted spaces, commas, emoticons are removed from data. Reviews are manually pre-classified into positive, negative and doubt class.

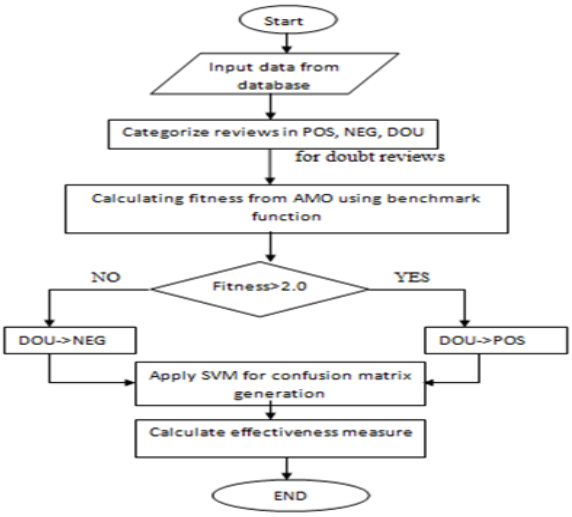

Figure 1 Flowchart of Proposed Algorithm

Proposed Pseudo-code:

Input: Reviews file

Output: Effectiveness Measures

1. Read the review file.
2. Count POS, NEG and DOU tags.
3. Calculating fitness using AMO()
    a. Generating the population using formula:
        i. D=10;
        ii. lu=[0*ones(1,D);1*ones(1,D)]
        iii. p=repmat(lu(1,:),popsize,1)+rand(popsize,D).*(repmat(lu(2,:)-lu(1,:),popsize,1))
        iv. fit= benchmark_func(x, 'standard function')
        v. Update GlobalMin and GlobalParam
        vi. Finding positions lseq=[i-2 i-1 i i+1 i+2]
        vii. Updating new position, GlobalMin and GlobalParams

4. If fitness > 2.0

    Convert DOU == POS

    Else

    Convert DOU == NEG

5. Update the file and apply SVM for confusion matrix generation.
6. Calculate the effectiveness measures i.e. accuracy, recall, precision, and kappa.
7. End.

Doubt reviews have various type of information. Sometimes it contains general information like a history of any place, thoughts, a combination of positivity and negativity of mind and so on. So according to a survey, these are neither purely positive nor negative. But there is need to optimize these reviews and categorize into positive and negative so information behind it becomes useful.

For doubt reviews, Animal Migration Optimization (AMO) [17] technique is used to calculate the fitness of doubt review. AMO works on the concept of animal, birds, mammals, fish etc. behaviour of migration. Migration is a common phenomenon in animals due to inappropriate climatic condition, lack of food and so on. Animal migration is the movement of individuals in a long distance, usually when the season changes. Basically, three rules are needed to be obeyed in migration process of animals, move in neighbour's direction, always close to a neighbour and do not collide with a neighbour.

Animal migration process [17] is divided into two sub-processes; animal migration and animal updating process. Animal groups are move from one place to another place in the animal migration process. For animal migration three rules are required; avoid collision with neighbour, move in the direction of neighbour, remain close to neighbour. In this process, concept of the local neighbourhood is coming into consideration. For the neighbourhood, a ring topology is used. Neighbourhood topology is static and is described on the set of indices of vectors. If an animal is at position k, then its neighbours have position k-2, k-1, k, k+1, k+2 if the individual has position is 5 then its neighbours are 3, 4, 5, 6, 7 and so forth. When topology of neighbourhood is created, then individual randomly

select one neighbour and update position according to neighbour by using formula:

$$X_{i,g+1} = X_{j,g} + \delta\,(X_{neiborhood,g} - X_{j,g}) \quad \text{.....(2.1)}$$

Where, $X_{neiborhood,g}$ is the neighbourhood present position, δ is random number generator, $X_{j,g}$ is the current position of the ith individual, and $X_{j,g+1}$ is the new position of ith individual.

In population updating phase, algorithm deals with how animals join new group or population and leave old group. Animals are moved and replaced by new individual by probability Pa. Probability is used on the basis of fitness value. Now, randomly choose integers and evaluate new solution and compared with old and choose better objective fitness:

$$X_i = \begin{cases} X_i, & \text{if } f(X_i) \text{ is better than } f(X_{i,g+1}), \\ X_{i,g+1}, & \text{otherwise} \end{cases} \quad \text{.....(2.2)}$$

By using AMO, population is generated as same as number of doubt reviews present in database. Now fitness of all individual of population is calculated by benchmark function. Benchmark functions are used to check the fitness value of an individual. There are 23 benchmark functions in AMO for fitness value evaluation. When fitness of all individual are calculated then global minimum and global parameters are calculated. Global minimum is a smallest value of fitness of individual from population and whose parameters are called as global parameter. Now, calculate neighbors of each individual and find the maximum value from these neighbors and update by using formula 2.1. After updating new population of individual is created by replacing with neighbor.

For new population, benchmark function is used for fitness calculation of each individual in population. Then compare new population with old population. Select 5 random values from population and calculate probability of each individual on the basis of fitness value. By using formula 2.2 compare and update the population and generate new individuals. If fitness value is greater than 2.0, then doubt review convert into positive otherwise negative. After this process new database is generated and Support Vector Machine is applied on this database.

SVM is a binary classifier and classify review into two classes of positive and negative. It has marginal hyperplane which separated text into two classes. SVM classify data and generate confusion matrix. On the basis of confusion matrix data performance measures precision, recall, accuracy, kappa is calculated.

V. RESULT ANALYSIS

For performance evaluation of proposed system, we use a dataset of reviews extracted from the government websites for Gwalior and Madhya Pradesh Tourism and manually pre-classify it. The corpus contains 700 reviews posted by users on websites. Reviews have been manually classified in positive, negative and doubt. Each review was extracted in a separate file tagged as one of two classes. The files have been processed and content was introduced into another file named “new reviews.xlxs”.

Performance Measures

System classification performance was evaluated using text mining specific measures. For current system are calculated: precision, accuracy, recall and kappa. These values are calculated based on confusion matrix. Confusion matrix contain values namely true positive. True negative, false positive and false negative, that are required for calculation of precision, recall, accuracy, kappa. Performance measures are discussed below:

- Precision: Precision refers to the closeness of two or more measurements to each other. Precision is defined as the number of true positives over the number of true positives plus the number of false positives.

$$\text{Precision} = \frac{t_p}{t_p + f_p}$$

- Recall: Recall is defined as the number of true positives over the number of true positives plus the number of false negatives.

$$\text{Recall} = \frac{t_p}{t_p + f_n}$$

- Accuracy: Accuracy refers to the closeness of a measured value to a standard or known value.

$$\text{Accuracy} = \frac{t_p + t_n}{t_p + t_n + f_p + f_n}$$

- Kappa: Cohen's kappa is a measure of the agreement between two raters who determine which category a finite number of subjects belong to whereby agreement due to chance is factored out. The two raters either agree in their rating (i.e. the category that a subject is assigned to) or they disagree; there are no degrees of disagreement (i.e. no weightings).

$$\Pr(a) = \frac{t_p + t_n}{t_p + t_n + f_p + f_n}$$

$$\Pr(e) = \frac{(f_p + t_n) * (t_p + t_n)}{(f_p + f_n) * (t_p + f_n)}$$

$$k = \frac{(\Pr(a) - \Pr(e))}{1 - \Pr(e)}$$

Where,

tp is true positive, fp is false positive, tn is true negative, fn is false negative

True Negative (TN): These are the negative tuples that were correctly labelled by the classifier.

True Positive (TP): These are the positive tuples that were correctly labelled by the classifier.

False Positive (FP): These are the negative tuples that were incorrectly labelled as positive.

False Negative (FN): These are the positive tuples that were mislabelled as negative.

Comparison Result

The classification algorithms Naïve Baye's and SVM are analyzed on the basis of performance measures namely precision, recall, accuracy and kappa. The proposed method also analyzed on these measures. Table1 shows the comparison between Naïve Baye's, SVM and SVM+AMO. All the results of precision, recall, accuracy and kappa are illustrated in a Table 1. On the basis of different number of reviews performance is evaluated.

Table 1 Comparison of Effectiveness measures for various numbers of reviews

RESULTS					
No. of Reviews	**Algorithm**	**Precision**	**Recall**	**Accuracy**	**Kappa**
100 Reviews	Naive Bayes	0.8690	0.9605	0.76	0.0596
	SVM	0.8666	0.975	0.8	0.0566
	SVM+AMO	0.9247	1	0.86	-0.0753
200 Reviews	Naive Bayes	0.8437	0.9440	0.715	0.0468
	SVM	0.8757	0.9155	0.77	0.2223
	SVM+AMO	0.9180	0.9655	0.87	0.2448
300 Reviews	Naive Bayes	0.9352	0.8850	0.87	0.5287
	SVM	0.9595	0.8681	0.91	0.6737
	SVM+AMO	0.9507	0.9608	0.9367	0.5043
400 Reviews	Naive Bayes	0.8906	0.8962	0.795	0.3212
	SVM	0.8993	0.8939	0.825	0.394
	SVM+AMO	0.9361	0.9777	0.9	0.2331
500 Reviews	Naive Bayes	0.9034	0.8946	0.816	0.3721
	SVM	0.9188	0.8891	0.8666	0.5089
	SVM+AMO	0.9446	0.9736	0.912	0.3064
600 Reviews	Naive Bayes	0.9057	0.8668	0.8133	0.4217
	SVM	0.9028	0.8796	0.845	0.473
	SVM+AMO	0.9359	0.9722	0.9017	0.2852
700 Reviews	Naive Bayes	0.9014	0.8777	0.8186	0.4128
	SVM	0.9114	0.888	0.8314	0.4218
	SVM+AMO	0.9334	0.9762	0.9029	0.2562

Graphical representation of comparison

In graphical representation, calculated measures represented by graphs. Each graph shows performance measure separately. Figure 2 show the comparison between SVM and SVM+AMO on the basis of precision. Precision of proposed method is better than SVM. Figure 3 shows comparison between SVM and SVM+AMO on the basis of recall. Figure 4 shows comparison on the basis of accuracy measure. Figure 5 shows comparison on the basis of kappa.

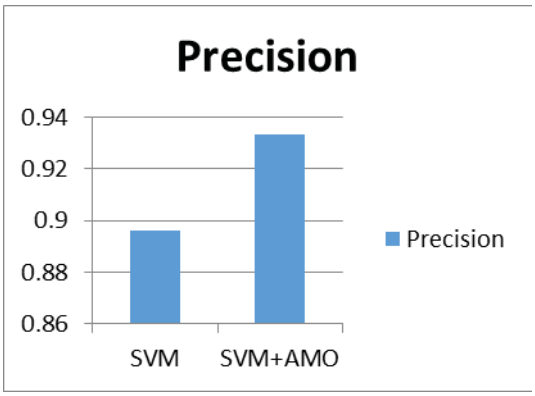

Figure 2.Comparison of Precision

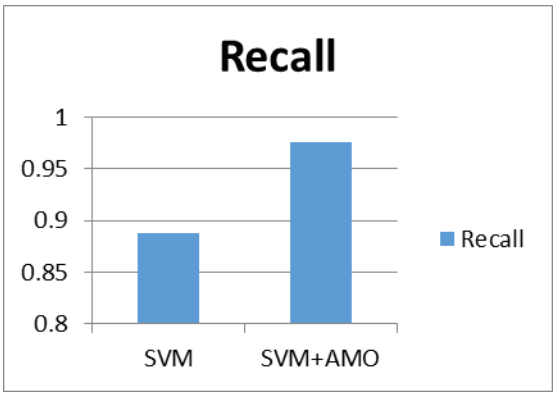

Figure 3.Comparison of Recall

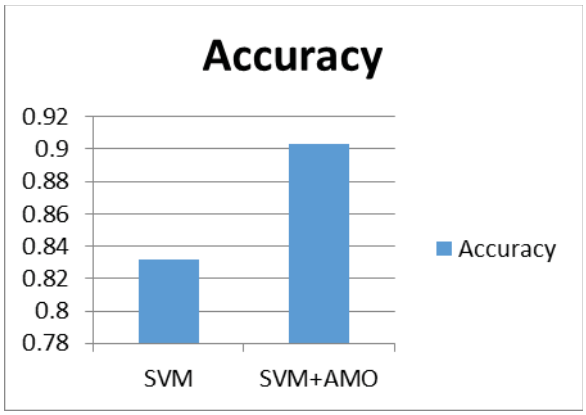

Figure 4.Comparison of Accuracy

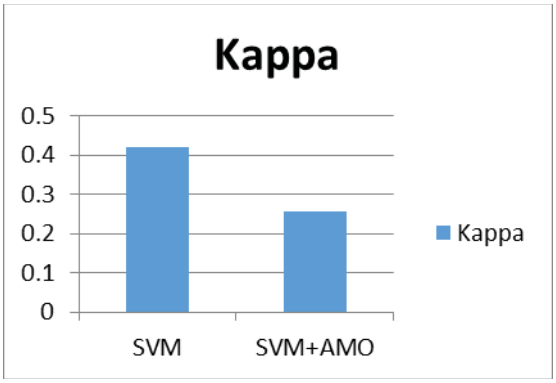

Figure 5.Comparison of Kappa

On the basis of different number of reviews

By taking different number of reviews, calculate performance measures of Naïve Baye's , SVM and SVM+AMO for comparison.

100 reviews

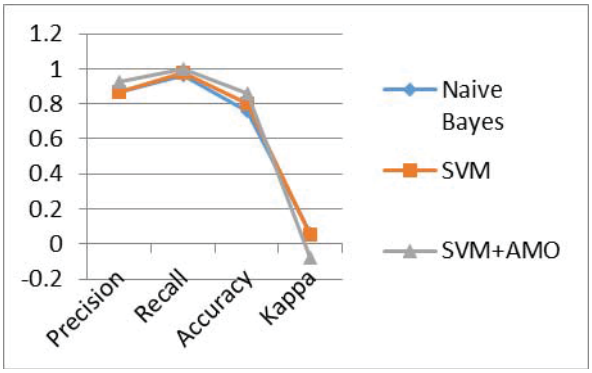

Figure 6.Comparison of Performance Measure of 100 reviews

200 reviews

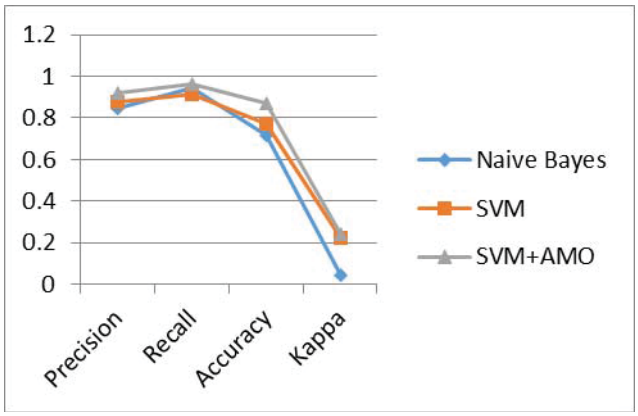

Figure 7.Comparison of Performance Measure of 200 reviews

300 reviews

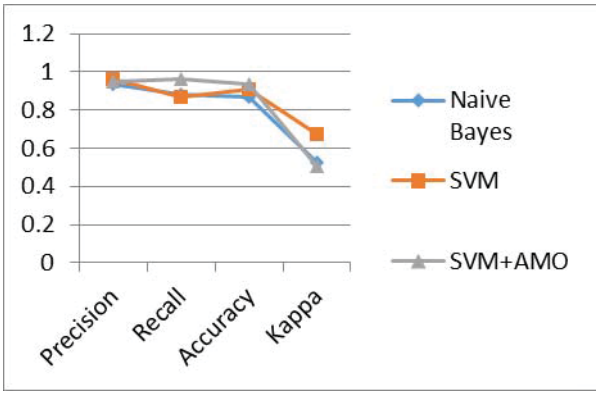

Figure 8.Comparison of Performance Measure of 300 reviews

400 reviews

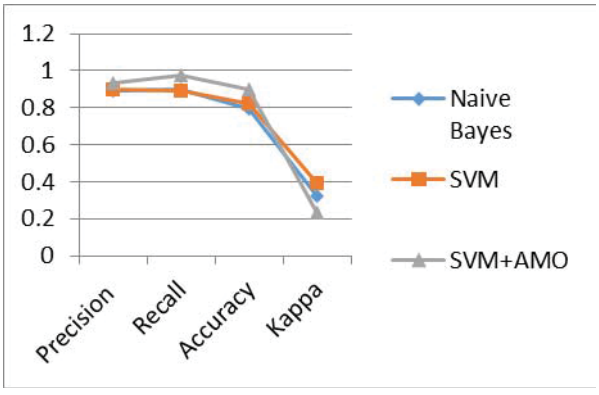

Figure 9.Comparison of Performance Measure of 400 reviews

500 reviews

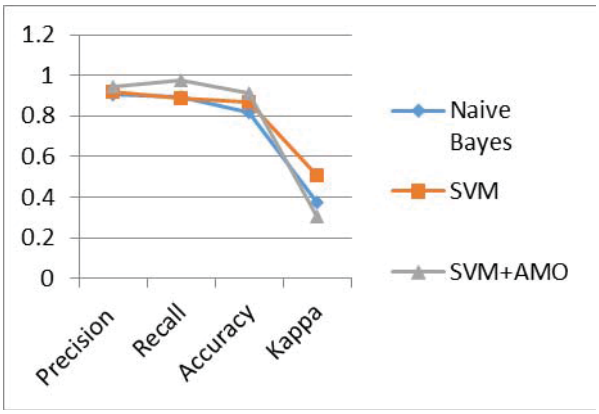

Figure 10.Comparison of Performance Measure of 500 reviews

600 reviews

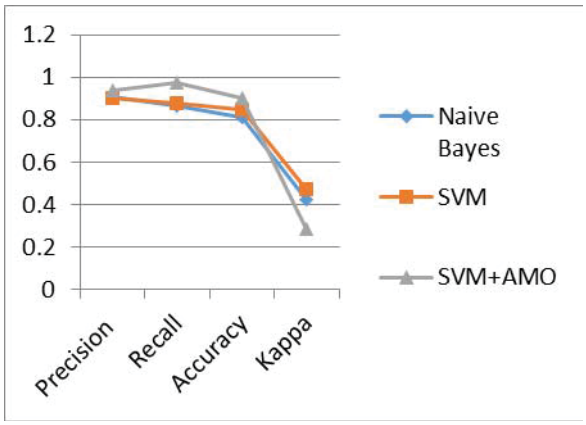

Figure 11.Comparison of Performance Measure of 600 reviews

700 reviews

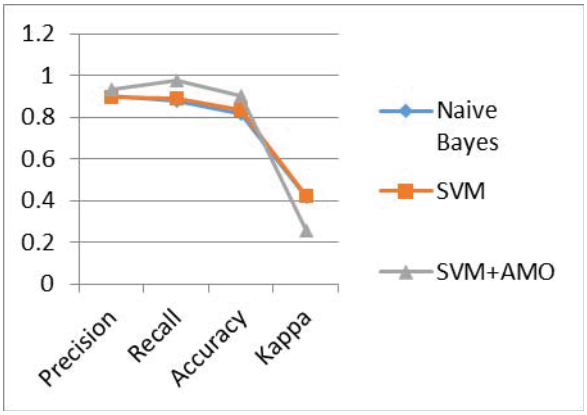

Figure 12.Comparison of Performance Measure of 700 reviews

VI. CONCLUSION

This research presents an efficient way of opinion mining by analyzing Gwalior tourism reviews, which are given by users that visit historical place of Gwalior and give their opinion on the tourism platform. For proposed method, reviews are collected from tourism website and pre-classified into positive, negative and doubt. Support Vector Machine performs better but because of doubt reviews performance become degraded. The proposed method combines Animal migration optimization technique with Support Vector Machine for optimized results. Animal migration optimization technique used to optimize doubt reviews and classify into positive and negative on the basis of fitness value. Performance measures give good result as comparison to SVM and accuracy is good.

REFERENCES

[1] G.Vinodhini, RM.Chandrasekaran: "Sentiment Analysis and Opinion Mining: A Survey" , IJARCSSE, Volume 2, Issue 6, June 2012.

[2] Yi and Niblack, "Sentiment Mining in Web Fountain",Proceedings of 21st international Conference on Data Engineering, pp. 1073-1083, Washington DC,2005.

[3] Arti Buche, Dr.M.B.Chandak, Akshay Zadgoanakar "Opinion Mining and Analysis: A Survey", International Journal on Natural Language Computing (IJNLC) Vol 2 No 3 June2013Pg No 39-48.

[4] Blessy Selvam, A. Abirami, "A Survey on Opinion Mining Framework", International Journal of Advanced Research in Computer and Communication Engineering, Vol 2, Issue 9, Sep 2013Pg No 3544-3549.

[5] Vijay .B. Raut et al, "Survey on Opinion Mining and Summarization of User Reviews on Web", International Journal of Computer Science and Information Technologies (IJCSIT), Vol 5(2), 2014. 1026-1030.

[6] B. Pang and L. Lee, "Opinion Mining and Sentiment Analysis" Foundations and Trends® in Information Retrieval, vol. 2, no. 1 2, pp. 1-135, 2008.

[7] V.Hatzivassiloglou and K. R. Mckeown, "Predicting the Semantic Orientation of Adjectives", in 35th Annual Meeting of the Association for Computational Linguistics, 1997, pp. 174-181.

[8] M. Rushdi Saleh, M. T. Martín-Valdivia, A. Montejo-Ráez, and L. a. Ureña-Lopez, "Experiments with SVM to classify opinions in different domains", Expert Systems with Applications, vol. 38, no. 12, pp. 14799-14804, Nov. 2011.

[9] B. Pang, L. Lee and S. Vaithyanathan, "Thumps Up? Sentiment Classification using Machine Learning Techniques",ACL-02 Conf. Empirical Methods Natural Lang. Process, 2002, pp. 79-86.

[10] Nidhi Mishra et al, "Classification of Opinion Mining Techniques", International Journal of Computer Applications, Vol 56, No 13, Oct 2012Pg No 1-6.

[11] G. Angulakshmi and Dr. R. ManickaChezian, "An Analysis on the Opinion Mining: Techniques and Tools", International Journal of Advance Research in Computer and Communication Engineering, Volume 3, Issue 7, July 2014.

[12] Peter D Turney. Thumbs up or thumbs down?: semantic orientation applied to unsupervised classification of reviews. Proceedings of the 40th Annual Meeting on, (July):417-424, 2002.

[13] S. Morinaga, K. Yamanishi, K.Tateishi and T. Fukushima,"Mining Product reputation on the Web", International Conference on Knowledge Discovery & Data mining, 2002

[14] Cristian Bucurab*,"Using Opinion Mining Techniques in Tourism", 2nd Global Conference on Business, Economics, Management and Tourism, 30-31 October 2014, Prague, Czech Republic

[15] J. Zhu, C. Xu, and H.-shi Wang, "Sentiment Classification using the theory of ANNs" The Journal of China Universities of Posts and Telecommunications, vol. 17, no. July, pp. 58-62, Jul. 2010.

[16] Abd. Samad Hasan Basari, "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.

[17] Xiangtao Li, Jie Zhang and Minghao Yin, "Animal migration optimization: an optimization algorithm inspired by animal migration behavior", Neural Computation And Applications Springer-verlag, 2014.

[18] L Dini and G Mazzini, "Opinion classification through information extraction", Intl. Conf. on Data Mining Methods and Databases for, 2002.