# Sentiment Classification using Neural Network and Ensemble Model based on Genetic Algorithm

**Kalaivani P, Logeshwari D, Tamizhselvi A**

*Abstract: The fast development of web sites and the number of product on these websites are available. The purpose of classification of sentiment is to efficiently identify opinion expressed in text. This paper compares three different optimized models including genetic optimized feature selection method, Genetic Algorithm (GA), ensemble approach that uses information gain and genetic algorithm as feature selection methods incorporated SVM model, Genetic Bagging (GB) and the next method uses optimized feature selection as feature selection technique incorporated back propagation model, Genetic Neural Network (GNN) models are compared. We are tested in sentiment analysis using sample multi-domain review datasets and movie review dataset.. These approaches are tested using various quality metrics and the results show that the Genetic Bagging (GB) technique outperforms in classifying the sentiment of the multi domain reviews and movie reviews. An empirical analysis is performed to compare the level of importance of the classifiers GB, GNN methods with McNemar's statistical method.*

*Keywords: sentiment classification, machine learning, feature selection, review, information gain, genetic algorithm, ensemble method, back propagation model.*

## I. INTRODUCTION

Sentiment analysis analyzes a person's emotions, feelings, and behaviors from an enormous amount of subjective data in digital format. This information is used for the identification and classification of sources. It is designed to determine a writer's attitude towards a certain product subject or the general polarity of field studies. The attitude might be his decision, his situation or the desired emotional contact. Data collection and user opinion identification is thus an important task that the research community has been focused on over the last decades.

Sentiment analysis not only benefits individual, but also allows businesses and organizations to determine their feelings or opinions. During the decision making process the customer's actions and an opinion regarding the product encourage the organization. A fundamental task in the classification of opinions is to determine the polarity of the information contained in the text file, sentence or function level for the product or service analysis. Whether a positive opinion, negative opinion articulated in the review text file, sentence or entire feature ( Turney, 2002).

**Kalaivani P**\*, Department of IT, St.Joseph's College of Engineering, OMR, Chennai-600119. Email: kalaivanip@stjosephs.ac.in
**Logeshwari D**, Department IT, St.Joseph's College of Engineering, OMR, Chennai-600119. Email: logesh_gd@yahoo.com
**Tamizhselvi A**, Department IT, St.Joseph's College of Engineering, OMR, Chennai-600119. Email: atamizhselvi@gmail.com

The WWW provides a review of consumer views, emotions and service views that have been stored in websites, blogs, a nd web forms.Today, there are a rapidly increasing number o f articles, websites, and blogs.The blogs collect important te xts and users use blogs to express their emotions, feelings and opinions. We need to classify the polarity of the product or service review data set in the document.

There are several datasets in the domain of movie reviews and multi-domain reviews that have been annotated as opinions and/or sentiment contained in the text. Various experiments are conducted on five domain reviews for the proposed work. The movie-review datasets include movie reviews of approximately thousand positive reviews dataset and thousand negative reviews dataset. The datasets contains product reviews such as book reviews, kitchen appliances and product reviews, DVD reviews; Electronics product reviews each containing thousand positive reviews dataset and thousand negative reviews dataset. Movie reviews play a tough task during sentiment classification due to the existence of various comparisons, abbreviations, presence of slang, and unclear languages. As opposed to the user-generated content, multi-domain reviews provide a more structured and less emotional text in terms of style, yet more subtle in opinion expression, thus rendering difficulty for analysis.

This paper studies and analyzes about the ensemble classification algorithm, back propagation algorithm and genetic algorithm for sentiment classification of multi domain reviews and movie reviews.

- To develop a new approach to sentiment classification using genetic algorithm, hybrid genetic algorithm incorporated with a multilayer Feed-Forward Neural Network and ensemble algorithm. The sampling technique like stratified sampling is used with the bagging, method in which comparison is done to find the one which results better.
- To apply this approach to multi domain reviews and movie reviews.
- To evaluate the efficiency of these new approaches in the multi domain reviews and movie reviews.

## II. LITERATURE SURVEY

Several methods have been applied to review the documents, such as Support Vector Machine (SVM), Naive Bayes (NB), K-Nearest Neighbor (kNN), Decision tree (DT) algorithms incorporated with feature / attribute selection methods which are used to guess user's feedback, thoughts, feelings for example, positive, negative and neutral.

(Mullen and Collier, 2004; Zhang, Ye, and Li., 2011; Tan and Zhang, 2008; Xia, Zong and Li., 2011; Ye, Zhang and Law, 2009).

Researchers investigated diverse ensemble strategies for classification tasks is to improve the performance of the base learners by combining different feature sets and different classification algorithms. (Wilson, Wiebe and Hwa, 2006; Tsutsumi, Shimada and Endo, 2007).

Previous studies showed that an ensemble method has worked well than single machine learning techniques for sentiment classification (Abbasi, Chen and Salem, 2008).

The new method for affect analysis suggested by Abbasi, Chen and Salem, 2008 is the Support Vector Regression Correlation Ensemble (SVRCE). The hybrid approach for sentence-level sentiment analysis, focused on sentiment lexicon and fuzzy set, was proposed by Appel, Chiclana, Carter and Fujita, 2016.

## III. MOTIVATION AND CONTRIBUTION

Ghiassi, Olschimke, Moon and Arnaudo, 2012 published first practical implementation of a neural network based classification and many other supervised learning tasks. There were many models for the field of neural networks; several type of research have successfully studied and implemented various classification problems in the application field. However previous studies have shown that the classification of text sentiment based on neural network is rare . Learning using a multilayer feed-forward neural networks is the basic concept of back- propagation algorithm. At every iteration, a set of sentiment attribute weights are learned for prediction of class labels. A neural network is a group of input units and output units in which each link has a sentiment attribute weight associated with it. To predict correct class labels for each sentiment attribute in the input, the networks learn by adjusting the weight (Zhu. Jian, Xu. Chen, Wang, 2010; Chen, Liu,Chiu, 2011; Sharma and Dey, 2012; Moraes, Valiati, 2013).

In this article, we have applied the supervised method for the classification of sentiments. The solution suggested is presented here. The proposed models of GNN and GB are defined in the following way:

**Input:**

As a learning scheme, the evaluation dataset D, collection of d training review datasets and a classifier will be used.

**Output:**

Prediction model.

**Method:**

- Alter all characters into lower case characters; do stemming and filtered stop words, tokenization of carrying out.
- TF-IDF calculation to transform the text representation matrix, the test uses unigram and bigram.
- A random analysis subdivision of the entire document is provided by the stratified sample.
  - Calculate and assign attribute weights to the importance of an attribute based on IG.
  - Select from the input terms, compare the weight of which meets the criteria (with the top 7 percent weight value) in terms of input weight.

- For each test IG feature selection and optimized feature selection that incorporates SVM with the weighting scheme of TF-IDF are used for each test.

The suggested used as a training dataset for learning models.

- The first task is based on the method of machine learning which uses IG attribute selection method and an optimized attribute selection (GA). The weight value of each attribute in the movie reviews dataset and multi-domain reviews data set is calculated by using information gain attribute selection method.
- To improve the efficiency of the process, Genetic Bagging (GB), the second task is based on IG attribute selection method and an optimized feature selection, integrating an ensemble technique, along with a well-known base learner classification algorithm for enhancing technical performance.
- The third approach is the Genetic Neural Network (GNN), which uses optimized feature selection and backpropagation model,.
- Such models are tested on five classification datasets for public sentiments. Our primary objective is to build and develop a system to improving the classification efficiency.
- The results of the three models (GA, GB, and GNN) are compared and the statistical test Mc Nemar's is applied to assess model efficiency. The research is carried out to predict the sentiment of movie reviews and multi- domain reviews.
- Assess the hybrid model performance, compared to the baseline process.
- Compute the quality parameters like positive precision value, negative precision value, positive recall value, negative recall value, and f-score value.

## IV. DATA SOURCE

The perceptions of the customers are important data sources that contribute to improving the quality of service. Some of the outlets where people are sharing their views include forms, review sites, and blogs. Movie reviews and multi-domain datasets are considered here to conduct here to carry out this analysis. The Cornell movie-review corpora[1] consists of movie review dataset. The multi domain dataset[2] includes Book reviews dataset, DVD product reviews dataset, Electronics product reviews dataset, and Kitchen product reviews dataset. A stratified sampling is applied for each domain to get a reduced attribute our problem. Table I shows the number of stratified sample reviews, number of positive reviews dataset, number of negative reviews dataset, the total number of attributes, attributes reduced after applying information gain, reduced attributes after applying an optimized feature reduction. The properties of the data source are developed and the test is represented in unigram and bigram attributes as a word vector.

## V. RESULT AND DISCUSSION

The performance of the genetic neural network (GNN) and hybrid genetic bagging (GB) algorithm for positive and negative reviews are evaluated and analyzed using various performance measures.

The performance of three approaches, genetic algorithm, genetic algorithm incorporated NN and hybrid genetic algorithm incorporated bagging technique has been tested and discussed. In the classifiers precision values, recall values and f-scores value of the classifiers are measured with the accuracy value.

Positive Precision refers to an accurately classified evaluation as positive. Negative Precision refers to an accurately classified evaluation as negative. Positive recall refers to the proportion of actual positive reviews are correctly classified as positive. It is also known as Sensitivity. Negative recall refers to the proportion of actual negative reviews that are correctly identified; it also known as specificity. F-score refers to the combination of precision and recall values, where f- score value the best 1 value and 0 worst score. The precision, recall, and f-score of the classifiers on each single class label are measured.

Tables II-VII describes all the hybrid classification approaches used in the study for different output measurements. Tables II and III for the multi-domain dataset show the results achieved for positive and negative precision. The results show that in classifying the positive reviews dataset and negative reviews dataset with high precision, the output of the GB method is better than the GNN model. In the case of positive precision and negative precision, GB performs better among all methods. Tables IV and V for the multi domain dataset show the results achieved for positive recall and negative recall. The results show that in classifying the positive reviews and negative reviews with high recall, the performance of the GB method is better than the GNN method. In terms of positive recall and negative recall, GB is better among all the methods. However, traditional measures such as precision, recall, f-score, and accuracy do not discriminate between the numbers of properly classified samples of various classes in the context of multi-domain datasets. Further, suitable metrics must be considered for the movie review dataset and multi-domain data sets. A confusion matrix gives results for four metrics, true positive value, true negative value, false positive value and false negative value.

Since these measures separately predict the outcome of each class these measures are combined to achieve good results for both classes. The Receiver operating characteristic (ROC) arc makes the representation of tradeoff from TPR (True Positive Rate) to FPR (False Positive Rate).

Furthermore, by increasing false positives, it is clear that any classifier cannot increase the number of true positives. The ROC curves in multi-domain datasets of the three models are shown in Figs. 1- 5. Four ROC curves are plotted for multi-domain reviews dataset for better illustration. Fig 5 for movie reviews dataset. The ROC curve was a little worse than GB and GNN for GA classifier. Among the entire five ROC curve, GB classifier performed better than others.

Table VIII presents the results obtained for AUC for GNN and GB models. It should be noted that the behavior in terms of AUC is different for all feature weight classification methods. Also, the result of GB for all feature weight has higher AUC values. Thus a high AUC value is obtained with the GB approach. The GNN result stated to achieve minimum AUC value of 0.77, 0.82, 0.80, 0.80, and 0.81 respectively for attribute weight greater than or equal with values 0.500,

0.400, 0.300, 0.200 and 0.100. While the result of GB for Book reviews has higher AUC values of 0.79, 0.85, 0.88, 0.91, 0.91 respectively for attribute weight greater than or equal with values 0.500, 0.400, 0.300, 0.200 and 0.100.

The minimum AUC values of 0.77, 0.78, 0.73, 0.79 and 0.77 are obtained for DVD reviews. While the result of GB for DVD reviews has higher AUC values of 0.83, 0.87, 0.88, 0.92, 0.89 is obtained. The minimum AUC values of 0.84, 0.82, 0.82, 0.87, and 0.86 are obtained for electronics reviews. While the result of GB for electronics reviews has higher AUC values of 0.81, 0.86, 0.90, 0.94, 0.93 is obtained. The minimum AUC values of .70, 0.79, 0.78, 0.79 and 0.73 are obtained for kitchen reviews. While the result of GB for kitchen reviews has higher AUC values of 0.71, 0.81, 0.85, 0.87, 0.86 is obtained. The minimum AUC values of 0.75, 0.82, 0.94, 0.95 and 0.95 are obtained for movie reviews. While for movie reviews GB has higher AUC values of 0.77, 0.85, 0.95, 0.99, 0.99 are obtained.

## A. Statistical Significance Test

The statistical test of McNemar's is used to compare the efficiency of the classifiers. The statistical analysis shows that GB is better than other classification methods. In Table IX to Table XI, symbol (←) means that classifier B is better than classifier A because Cfs value is lower than Csf value. In Table IX to Table XI, the sign (↑) indicates that classifier A performed better than classifier B because the Csf value is smaller than Cfs. Test results form McNemar's for the multi-domain reviews and movie reviews in Tables X to Table XI; show that GB has produced significantly improved results than GNN.

At 5% right-tailed test, H1 is accepted with a level of significance. For Book reviews dataset, DVD reviews dataset, Kitchen reviews dataset and Movie reviews dataset GB classifier performs better than GNN. For the Electronics reviews dataset, the GNN classifier performs better than GB.

[1]*www.cs.cornell.edu/people/pabo/movie-review-data*
\[2]*www.cs.jhu.edu/_mdredze/datasets/sentiment*

*Retrieval Number: B3677129219 /2020©BEIESP*
*DOI: 10.35940/ijeat.B3677.029320*

1887

*Published By:*
*Blue Eyes Intelligence Engineering*
*& Sciences Publication*

# Sentiment Classification using Neural Network and Ensemble Model based on Genetic Algorithm

**Table- I: Description of multi domain data set and movie review Dataset**

| Dataset | Number of stratified samples | Number of Positive Reviews dataset | Number of Negative Reviews dataset | Total No. of Attributes | Total No. of Attributes (Weight by IG) | Total No. of Attributes (Optimize Selection) |
|---|---|---|---|---|---|---|
| BOOK REVIEWS | 191 | 100 | 91 | 407 | 367 | 297 |
| DVD REVIEWS | 199 | 99 | 100 | 392 | 342 | 279 |
| ELECTRONICS REVIEWS | 200 | 100 | 100 | 303 | 236 | 196 |
| KITCHEN REVIEWS | 190 | 99 | 91 | 234 | 214 | 146 |
| MOVIE REVIEWS | 200 | 100 | 100 | 1719 | 1718 | 1069 |

**Table- II: Positive Precision of multi domain data set**

| Dataset | Positive Precision GNN | | | | | Positive Precision GB | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | Attribute Weight | | | | | Attribute Weight | | | | |
| | ≥ 0.500 | ≥ 0.400 | ≥ 0.300 | ≥ 0.200 | ≥ 0.100 | ≥ 0.500 | ≥ 0.400 | ≥ 0.300 | ≥ 0.200 | ≥ 0.100 |
| BOOK REVIEWS | 0.65 | 0.76 | 0.79 | 0.88 | 0.89 | 0.65 | 0.76 | 0.80 | 0.87 | **0.91** |
| DVD REVIEWS | 0.75 | 0.75 | 0.77 | 0.85 | 0.85 | 0.76 | 0.75 | 0.77 | 0.81 | **0.87** |
| ELECTRONICS REVIEWS | 0.63 | 0.70 | 0.79 | 0.88 | **0.91** | 0.65 | 0.75 | 0.78 | 0.88 | **0.91** |
| KITCHEN REVIEWS | 0.82 | 0.83 | 0.86 | 0.86 | 0.86 | 0.82 | 0.83 | 0.88 | 0.88 | **0.88** |
| MOVIE REVIEWS | 0.85 | 0.78 | 0.96 | 0.96 | **0.97** | 0.68 | 0.78 | 0.91 | 0.92 | **0.97** |

**Table- III Negative Precision of multi domain data set**

| Dataset | Negative Precision GNN | | | | | Negative Precision GB | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | Attribute Weight | | | | | Attribute Weight | | | | |
| | ≥ 0.500 | ≥ 0.400 | ≥ 0.300 | ≥ 0.200 | ≥ 0.100 | ≥ 0.500 | ≥ 0.400 | ≥ 0.300 | ≥ 0.200 | ≥ 0.100 |
| BOOK REVIEWS | 0.85 | 0.92 | 0.90 | 0.85 | 0.91 | 0.84 | **0.92** | 0.88 | 0.87 | 0.90 |
| DVD REVIEWS | 0.96 | 0.93 | 0.91 | 0.90 | 0.90 | 0.89 | 0.91 | 0.91 | 0.86 | **0.98** |
| ELECTRONICS REVIEWS | 0.98 | **1.00** | 0.91 | 0.91 | 0.93 | 0.90 | 0.93 | **0.97** | 0.92 | 0.93 |
| KITCHEN REVIEWS | 0.62 | 0.70 | 0.77 | 0.79 | 0.91 | 0.59 | 0.71 | 0.76 | 0.79 | **0.88** |
| MOVIE REVIEWS | 0.77 | 0.81 | 0.87 | 0.95 | 0.97 | 0.76 | 0.83 | 0.91 | 0.96 | **0.97** |

**Table- IV Positive Recall of multi domain data set**

| Dataset | Positive Recall GNN | | | | | Positive Recall GB | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | Attribute Weight | | | | | Attribute Weight | | | | |
| | ≥ 0.500 | ≥ 0.400 | ≥ 0.300 | ≥ 0.200 | ≥ 0.100 | ≥ 0.500 | ≥ 0.400 | ≥ 0.300 | ≥ 0.200 | ≥ 0.100 |
| BOOK REVIEWS | 0.93 | **0.95** | 0.93 | 0.86 | 0.92 | 0.92 | 0.95 | 0.91 | 0.88 | 0.91 |
| DVD REVIEWS | 0.91 | 0.91 | 0.93 | 0.91 | 0.91 | 0.91 | 0.93 | 0.93 | 0.93 | **0.95** |
| ELECTRONICS REVIEWS | 0.99 | **1.00** | 0.98 | 0.91 | 0.94 | 0.95 | 0.95 | 0.98 | 0.92 | 0.93 |
| KITCHEN REVIEWS | 0.52 | 0.66 | 0.76 | 0.80 | 0.93 | 0.55 | 0.69 | 0.75 | 0.79 | **0.90** |
| MOVIE REVIEWS | 0.82 | 0.82 | 0.86 | 0.95 | 0.97 | 0.80 | 0.84 | 0.91 | 0.96 | **0.97** |

**Table- V Negative Recall of multi domain data set**

| Dataset | Negative Recall GNN | | | | | Negative Recall GB | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | Attribute Weight | | | | | Attribute Weight | | | | |
| | ≥ 0.500 | ≥ 0.400 | ≥ 0.300 | ≥ 0.200 | ≥ 0.100 | ≥ 0.500 | ≥ 0.400 | ≥ 0.300 | ≥ 0.200 | ≥ 0.100 |
| BOOK REVIEWS | 0.45 | 0.67 | 0.73 | 0.87 | 0.88 | 0.47 | 0.67 | 0.75 | 0.86 | **0.90** |
| DVD REVIEWS | 0.70 | 0.68 | 0.73 | 0.86 | 0.85 | 0.71 | 0.70 | 0.72 | 0.80 | **0.87** |
| ELECTRONICS REVIEWS | 0.42 | 0.59 | 0.74 | 0.91 | 0.87 | 0.48 | 0.68 | 0.73 | 0.87 | **0.90** |
| KITCHEN REVIEWS | 0.88 | 0.88 | 0.87 | 0.86 | 0.89 | 0.87 | 0.85 | **0.89** | 0.88 | 0.79 |
| MOVIE REVIEWS | 0.59 | 0.77 | 0.97 | 0.96 | 0.96 | 0.62 | 0.76 | 0.91 | 0.94 | **0.97** |

**Table- VI Positive f score of multi domain data set**

| Dataset | Positive f score GNN | | | | | Positive f score GB | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | Attribute Weight | | | | | Attribute Weight | | | | |
| | ≥ 0.500 | ≥ 0.400 | ≥ 0.300 | ≥ 0.200 | ≥ 0.100 | ≥ 0.500 | ≥ 0.400 | ≥ 0.300 | ≥ 0.200 | ≥ 0.100 |
| BOOK REVIEWS | 0.77 | 0.84 | 0.85 | 0.87 | 0.90 | 0.76 | 0.84 | 0.85 | 0.88 | **0.91** |
| DVD REVIEWS | 0.82 | 0.84 | 0.84 | 0.88 | 0.88 | 0.88 | 0.83 | 0.84 | 0.84 | **0.89** |
| ELECTRONICS REVIEWS | 0.77 | 0.82 | 0.87 | 0.91 | 0.91 | 0.77 | 0.84 | 0.87 | 0.90 | **0.91** |
| KITCHEN REVIEWS | 0.64 | 0.75 | 0.81 | 0.83 | 0.86 | 0.66 | 0.75 | 0.81 | 0.83 | **0.88** |
| MOVIE REVIEWS | 0.83 | 0.80 | 0.91 | 0.96 | **0.97** | 0.74 | 0.81 | 0.91 | 0.94 | **0.97** |

**Table- VII Negative f score of multi domain data set**

| Dataset | Negative f score GNN | | | | | Negative f score GB | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | Attribute Weight | | | | | Attribute Weight | | | | |
| | ≥ 0.500 | ≥ 0.400 | ≥ 0.300 | ≥ 0.200 | ≥ 0.100 | ≥ 0.500 | ≥ 0.400 | ≥ 0.300 | ≥ 0.200 | ≥ 0.100 |
| BOOK REVIEWS | 0.45 | 0.64 | 0.71 | 0.89 | 0.88 | 0.47 | 0.64 | 0.74 | 0.87 | **0.91** |
| DVD REVIEWS | 0.63 | 0.63 | 0.69 | 0.84 | 0.83 | 0.77 | 0.65 | 0.68 | 0.78 | **0.86** |
| ELECTRONICS REVIEWS | 0.38 | 0.52 | 0.71 | 0.89 | 0.85 | 0.45 | 0.63 | 0.69 | 0.86 | **0.91** |
| KITCHEN REVIEWS | 0.96 | 0.95 | 0.91 | 0.90 | 0.77 | 0.98 | 0.90 | 0.95 | 0.93 | **0.99** |
| MOVIE REVIEWS | 0.74 | 0.76 | 1.02 | 0.97 | 0.96 | 0.61 | 0.75 | 0.91 | 0.91 | **0.97** |

**Table- VIII AUC of classifiers**

| Dataset | GNN | | | | | GB | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | Attribute Weight | | | | | Attribute Weight | | | | |
| | ≥ 0.500 | ≥ 0.400 | ≥ 0.300 | ≥ 0.200 | ≥ 0.100 | ≥ 0.500 | ≥ 0.400 | ≥ 0.300 | ≥ 0.200 | ≥ 0.100 |
| BOOK REVIEWS | 0.77 | 0.82 | 0.80 | 0.80 | 0.81 | **0.79** | 0.85 | 0.88 | 0.91 | 0.91 |
| DVD REVIEWS | 0.77 | 0.78 | 0.73 | 0.79 | 0.77 | **0.83** | 0.87 | 0.88 | 0.92 | 0.89 |
| ELECTRONICS REVIEWS | 0.84 | 0.82 | 0.82 | 0.87 | 0.86 | **0.81** | 0.86 | 0.90 | 0.94 | 0.93 |
| KITCHEN REVIEWS | 0.70 | 0.79 | 0.78 | 0.79 | 0.73 | **0.71** | 0.81 | 0.85 | 0.87 | 0.86 |
| MOVIE REVIEWS | 0.75 | 0.82 | 0.94 | 0.95 | 0.95 | **0.77** | 0.85 | 0.95 | 0.99 | 0.99 |
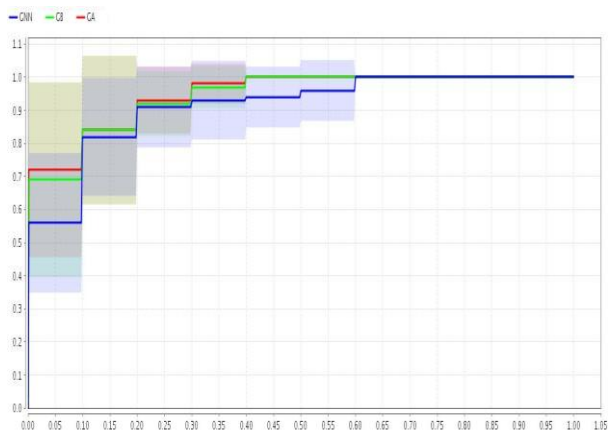


**Fig. 1. ROC Curves of GA, GNN and GB for book reviews**
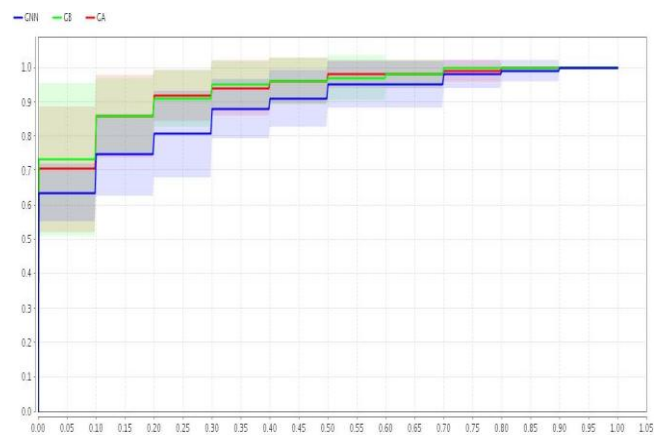


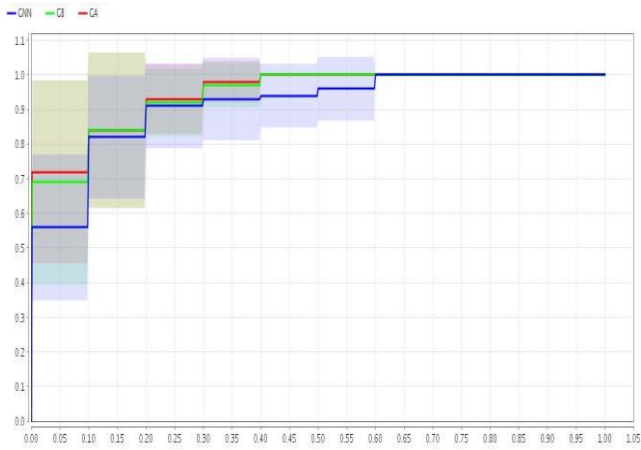**Fig. 2. ROC Curves of GA, GNN and GB for DVD reviews**

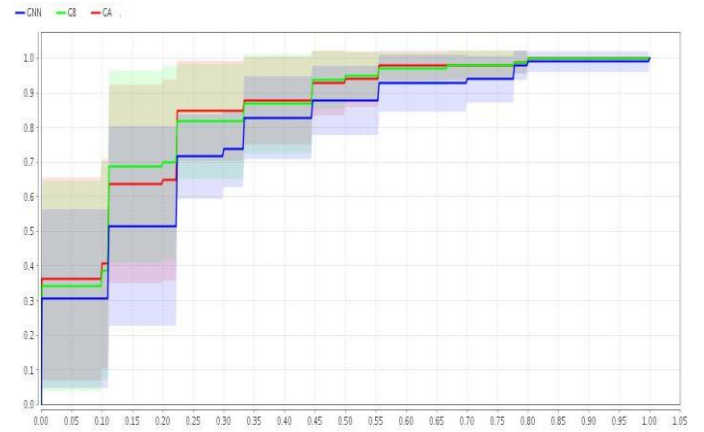**Fig. 3. ROC Curves of GA, GNN and GB for electronics reviews**



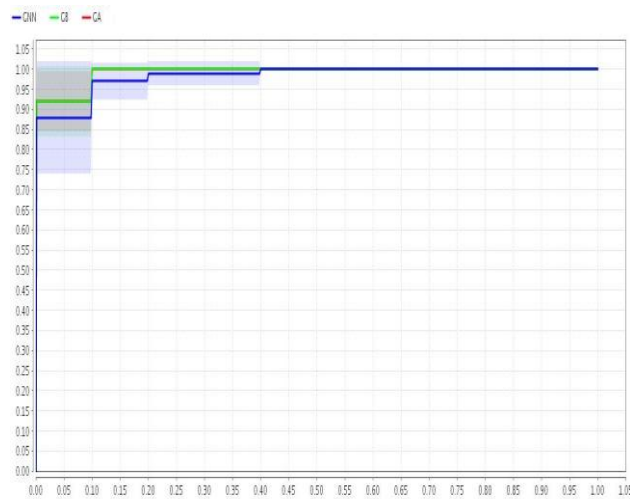**Fig. 4. ROC Curves of GA, GNN and GB for kitchen reviews**



**Fig. 5. ROC Curves of GA, GNN and GB for movie reviews.**

**Table- IX Effectiveness comparison of the GB and GNN for book and DVD reviews**

| Classifier B (GNN) Attribute Weight | Classifier A (GB) | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | Book Reviews | | | | | DVD Reviews | | | | |
| | ≥0.500 | ≥0.400 | ≥0.300 | ≥0.200 | ≥0.100 | ≥0.500 | ≥0.400 | ≥0.300 | ≥0.200 | ≥0.100 |
| ≥0.500 | 0.05(↑) | | | | | 0.05(↑) | | | | |
| ≥0.400 | - | 0.05(↑) | | | | - | 0.05(↑) | | | |
| ≥0.300 | - | - | 0.05(↑) | | | - | - | 0.05(↑) | | |
| ≥0.200 | - | - | - | 0.05(↑) | | - | - | - | 0.95(←) | |
| ≥0.100 | - | - | - | - | 0.00(↑) | - | - | - | - | 0.05(↑) |

**Table- X Effectiveness comparison of the GB and GNN for Electronics reviews**

| Classifier B (GNN) Attribute Weight | Classifier A (GB) | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | Electronics Reviews | | | | | Kitchen Reviews | | | | |
| | ≥0.500 | ≥0.400 | ≥0.300 | ≥0.200 | ≥0.100 | ≥0.500 | ≥0.400 | ≥0.300 | ≥0.200 | ≥0.100 |
| ≥0.500 | 0.15(↑) | | | | | 0.15(↑) | | | | |
| ≥0.400 | - | 5.25(←) | | | | - | 0.05(↑) | | | |
| ≥0.300 | - | - | 0.10(←) | | | - | - | 0.05(↑) | | |
| ≥0.200 | - | - | - | 0.25(←) | | - | - | - | 0.05(↑) | |
| ≥0.100 | - | - | - | - | 5.48(←) | - | - | - | - | 0.25(←) |

**Table- XI Effectiveness comparison of the GB and GNN for Movie reviews**

| Classifier B (GNN) Attribute Weight | Classifier A (GB) | | | | |
|---|---|---|---|---|---|
| | ≥0.500 | ≥0.400 | ≥0.300 | ≥0.200 | ≥0.100 |
| ≥0.500 | 0.05(↑) | | | | |
| ≥0.400 | - | 0.05(↑) | | | |
| ≥0.300 | - | - | 0.05(←) | | |
| ≥0.200 | - | - | - | 0.05 (←) | |
| ≥0.100 | - | - | - | - | 0.05(↑) |

## VI. CONCLUSION

Experiments carried for multi-domain reviews and movie reviews for positive and negative reviews with different attribute weight relation using genetic algorithm, hybrid genetic NN and hybrid genetic bagging algorithm resulted in certain significant results. The NN approach of sentiment classification incorporated with a genetic algorithm improved the average accuracy of nearly 90.25%. The accuracy produced by the GNN algorithm is increased by nearly 1% when compared to GA. The GB approach of sentiment classification incorporated with a genetic algorithm improved the average accuracy of nearly 90.28%. The accuracy produced by the GB algorithm is increased by nearly 1.03% when compared to GA. The GB, GNN methods are compared with Mc Nemar's statistical to the importance of classification by an empirical analysis.

## REFERENCES

1. A. Abbasi, H. Chen and A. Salem. "Sentiment Analysis in Multiple Language: Feature Selection for Opinion Classification in Web Forms", ACM Transaction on Information Systems, 2008, vol 26, No3.
2. O. Appel, F. Chiclana, J. Carter and H. Fujita," A hybrid approaches to the sentiment analysis problem at the sentence level. Knowledge-Based Systems", 2016, 110–124, Advance online publication. 2016. vol 05, pp 1-34.
3. L.S. Chen, C.H. Liu, H.J. Chiu, "A neural network based approach for sentiment classification in the blogosphere." Journal of Informatics, 2011, vol 5, pp. 313–322.
4. M. Gamon , "Sentiment classification on customer feedback data: noisy data, large feature vectors, and the role of linguistic analysis".2004, In: Proceeding of the 20th Intl. Conference on computational linguistics.
5. M. Ghiassi, M. Olschimke, B. Moon and P. Arnaudo, "Automated text classification using a dynamic artificial neural network model". Expert System Applicatins, 2012, pp. 10967–10976.
6. J. Wiebe, T.Wilson, R.Bruce, M.Bell and M.Martin, "Learning Subjective Language," The Association for Computational Linguistics, vol. 30, no. 3,2004, pp. 277-308
7. J.Demsar., "Statistical comparisons of classifiers over multiple data sets," Journal of Machine Learning Research, 2006, pp. 1–6.
8. R. Moraes, JOF. Valiati, "Document-level sentiment classification: An empirical comparison between SVM and ANN", 2013, Expert Systems with Applications, Vol. 40, No. 2, pp. 621-633.
9. R. Prabowo and M.Thelwall, "Opinion analysis: a combined approach. Journal of Infometrics, 2009, pp143–157.
10. R. Xia, C. Zong and S. Li., "Ensemble of feature sets and classification algorithms for sentiment classification", Information Sciences, 2011, pp 1138–1152.
11. F. Salvetti, S. Lewis and C. Reichenbach," Automatic opinion polarity classification of movie reviews". Colorado research in linguistics. University of Colorado, Boulder, 2004, Vol. 17, No. 1.
12. A. Sharma and S.Dey, "A comparative study of feature selection and machine learning techniques for sentiment analysis", In Proceedings of the ACM Research in Applied Computation Symposium, ACM, 2012, pp. 1-7.
13. S.B. Tan and J. Zhang, "An empirical study of sentiment analysis for Chinese documents." Expert System. Applications, 2008, pp. 2622–2629.
14. T. Mullen and N. Collier, "Sentiment analysis using support vector machines with diverse information sources". In Proceedings of the Conference on Empirical Methods in Natural Language Processing, 2004, pp. 412-418, Barcelona, Spain.
15. K. Tsutsumi, K. Shimada and T. Endo, "Movie review classification based on a multiple classifier", In: The 21th Pacific Asia Conference on Language, Information and Computation, 2007.
16. P.D. Turney, "Thumbs up or thumbs down? Semantic orientation applied to unsupervised classification of reviews". In Proceedings of the 40th Annual Meeting on Association for Computational Linguistics. Association for Computational Linguistics, Philadelphia, Pennsylvania, using contextual valence shifters". Computational Intelligence, 2002, pp.110–125.
17. T. Wilson, J.Wiebe and R.Hwa, "Recognizing strong and weak opinion clauses". Computational Intelligence, 2006, pp.73–99.
18. Q. Ye, Z. Zhang and R. Law, "Sentiment classification of online reviews to travel destinations by supervised machine learning approaches ", Expert Systems with Applications, 2009, pp. 6527–6535.
19. Z. Zhang, Q. Ye, Z. Zhang and Y. Li,"Sentiment classification of Internet restaurant reviews written in Cantonese". Expert System Applications, vol.38, no.6, 2011, pp.7674–7682.
20. Jain Zhu, Chen Xu, Han-shi Wang, "Sentiment classification using the theory of ANNs", Published by Elsevier B.V.,The Journal of China Universities of Posts and Telecommunications, vol. 17, July 2010, pp. 58-62.

*Retrieval Number: B3677129219 /2020©BEIESP*
*DOI: 10.35940/ijeat.B3677.029320*

1891

*Published By:*
*Blue Eyes Intelligence Engineering*
*& Sciences Publication*