IJISET - International Journal of Innovative Science, Engineering & Technology, Vol. 7 Issue 6, June 2020
ISSN (Online) 2348 – 7968 | Impact Factor (2020) – 6.72
www.ijiset.com

Comparative Analysis of Machine Learning Based Classification Algorithms for Sentiment Analysis

Tekendra Nath Yogi[1], Nawaraj Paudel[2]

[1, 2]Central Department of Computer Science and Information Technology, Tribhuvan University

Kathmandu, Nepal

*email: nawarajpaudel@cdcsit.edu.np

Abstract

Sentiment analysis is the process of predicting the sentiment polarity of review data based on a given data set. Huge amount of review data is generated in each day on the Web. This rapidly increasing day to day data need to be processed to detect the sentiment polarity of large review datasets as early as possible. The main aim of this research is to evaluate and compare the performance of three classification algorithms, Multinomial naïve Bayes (MNB), K-Nearest-Neighbors (KNN) and Support Vector Machines (SVM), for sentiment labeled sentences with three different datasets having different sizes. When comparing the performance of all three classification algorithms for sentiment analysis, SVM is found to be a better algorithm to detect sentiment polarity in all three sentiment labeled sentence datasets in every aspect, whereas MNB and KNN have less performance in every aspect as compared to SVM.

Keywords: *Machine learning, Review Data, Sentiment polarity, Sentiment analysis, KNN, SVM, MNB*

1. Introduction

An opinion is a viewpoint or judgment about a specific thing that acts as a key influence on an individual process of decision making. Collectively Opinions reflects the “wisdom of crowds” and can be good indicator of the future [1]. In addition People’s belief and the choices they make are always dependent on how others see and evaluate the world so, opinion holds high values in many aspect of life. Sentiment analysis is the process of determining opinions or sentiments as positive or negative from review which are expressed by people over a particular subject, area, product, or services offered by companies, governments and organizations [2]. In recent years, this field is widely appreciated by researchers due to its dynamic range of application in various numbers of fields. There are several areas such as marketing; politics; news analytics etc. which are benefited from the result of sentiment analysis [3].

Due to the vast range of product and services these days, it has become difficult for the users to select their preferred product. Product reviews turn out to be very useful reference. Despite of the willingness of people to share their thoughts and views about the product, a problem persists due to the huge amount of total reviews [3]. This develops a need for technology of data mining to uncover information automatically and assist in decision making. Such data mining technology is sentiment analysis for classifying opinion based on the review polarity [4].

There are two main approaches for sentiment analysis: machine learning approach and lexicon-based approach to solve the problem of sentiment classification [5]. The former approach is applied to classify the sentiments based on training as well as test data sets. The second category doesn’t require any prior training data set; it performs the task by identifying a list of words, phrases that consists of a semantic value. It mainly concentrates on patterns of unseen data [3].

This field becomes more challenging due to the fact that many demanding and interesting research problems still exist in this field to solve. Sentiment based analysis of a document is quite tough to perform in comparison with topic based text classification. The opinion words and sentiments are always varied with situations. Therefore, an opinion word can be considered as positive in one circumstance but may be that becomes negative in some other circumstance. The opinionated word ‘unpredictable’ is used in different senses in a different domain. For example, “an unpredictable plot in the movie” gives a positive opinion about the movie, while “an unpredictable steering wheel” is a negative expression considering the product, car [3].

Sentiment classification process has been classified into three levels: document level, sentence level, and feature level. In Document level the whole document is classify either into positive or negative class. Sentiment classification at the sentence level, considers the individual sentence to identify whether the sentence is positive or negative. Feature level sentiment classification concerns with identifying and extracting product features from the source data [5].

During this research study, the focus has been made on sentiment polarity classification based on three sentiment labeled sentence datasets, namely Yelp restaurant review dataset, Amazon cell phones and accessories review dataset and IMDB movie review dataset, by using three text classification algorithms, namely MNB, KNN and SVM. To improve the processing efficiency and sentiment polarity classification performance the data has been preprocessed, such as case folding, stop word removing, stemming and lemmatization, and chi-squared method has been used for feature selection and TFIDF method has been used for feature weighting before fed as input to the sentiment polarity classification algorithms. The sentiment polarity classification algorithms have been evaluated based four performance evaluation parameters accuracy, precision, recall and F-measure.

2. Related Works

Problems similar to the topic have been examined in the past in various contexts and the field is rapidly growing. There are various research aspects that need to be considered in order to get the better result of analysis to help in better decision making. A novel approach based on Support Vector Machines has been proposed to compare lexical-based and Machine learning based approaches. This approach shows that machine learning based approach for sentiment classification is quite successful and outperforms the lexical based approach [6].

Both supervised and unsupervised algorithms for automatic classification of sentiments from 2000 social network users have been employed. The researchers found that supervised machine learning technique outperformed the unsupervised machine learning techniques with low classification error [7].

An experiment on automatic classification of Sentiments in text documents using classification algorithms has been carried out. This experiment classified the text documents by topic, and overall sentiment of documents according to negative and positive sentiments. From their experiment the researchers found that classification algorithm perform poorly on the sentiment classification by topic [8] [9].

A comparative study of machine learning based, lexical based and rule based approaches for sentiment analysis has been provided. This study found that machine learning based approach outperforms both lexical based and rule based approaches and additionally shows that more the cleaner the data more correct the knowledge [10].

Five classification algorithms on movie review data set in order to identify fake review data set have been compared. In this comparative study data are used without preprocessing and it is found that SVM outperforms all other four classification algorithms namely Naïve Bayes, KNN, K*, and Decision Tree-J48 [11].

A method which combines machine learning based methods with preprocessing techniques to determine the sentiment of twitter data has been proposed and compared with usual machine learning methods. It is found that the proposed method outperforms the usual machine learning methods [12].

A method has been proposed to compare four feature selection methods chi-squared, Information gain, Mutual information and symmetrical uncertainty in combination with two machines learning methods SVM and NB. The result showed that SVM with Information gain outperforms all others. But the chi- Squared method has better noise tolerance [13].

The researchers analyzed twitter data and proposed a new machine leaning based method in combination with preprocessing and feature selection methods for sentiment analysis. The proposed Machine learning base method has four steps that followed for the sentiment analysis in the first step, the first step is applied in which data pre-processed. In the second step feature of the data will be extracted which is given as input to the third step in which data is classified for the sentiment analysis. They found that the proposed method outperformed the usual machine learning methods [14].

A framework for text classification base on the KNN algorithm and the TF-IDF method has been proposed. This framework proves a good result and provides the ability to upgrade and improve the present embedded classification algorithm [15].

A method in combination with preprocessing and TFIDF weighting has been proposed and performed a comparative analysis with the usual machine learning algorithms by using twitter data. The researchers found that the proposed method outperforms the usual machine learning methods for sentiment analysis [16].

A method by combining Multinomial Naive Bayes as a selected machine learning technique for classification, and TF-IDF as a vector space model for text extraction, and chi square technique for feature selection has been proposed [17]. This proposed method outperforms the framework proposed in [15].

A method by combining Support Vector Machine as a selected machine learning technique for classification, and TF-IDF as a vector space model for text extraction, and chi2 technique for feature selection has been proposed. This proposed method outperforms the normal SVM [18].

3. Methodology

There are different techniques to find the polarity of a review data. The most popular and efficient one is Machine learning based sentiment analysis technique. The machine learning based sentiment analysis technique decides the polarity of a data point as either positive class or negative class. To decide the polarity of a review data and to find the most efficient algorithm following steps were used in this research as shown in the figure below.

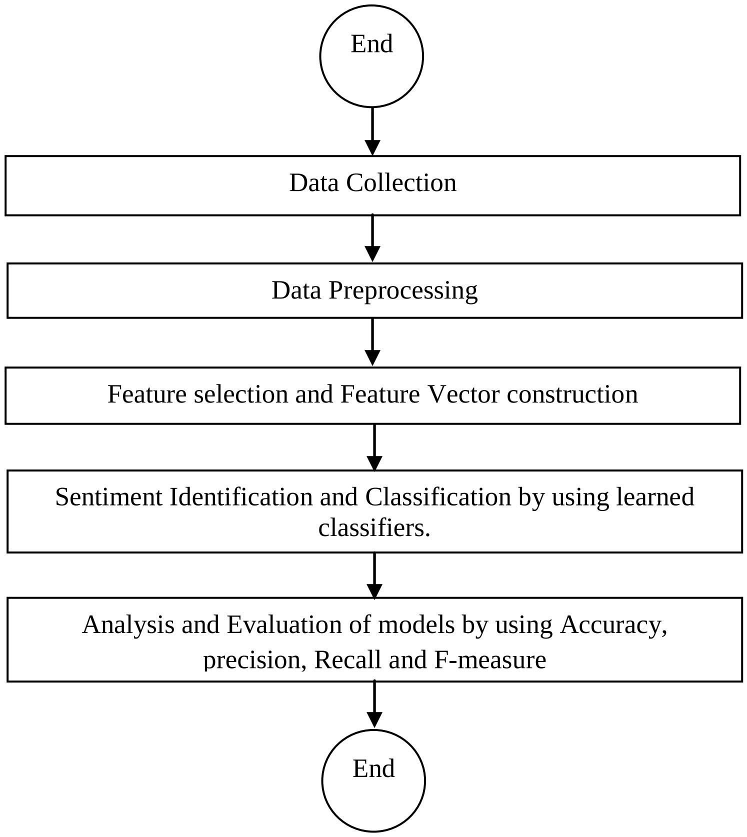

Fig 1. Flow Chart for Entire Process of Research.

3.1. Data collection

During the research three types, sentiment labeled sentence, data sets have been used which were collected from Kaggle (kaggle.com) machine learning repository. The data sets are different in size in terms of number of sentences. These data sets can be used for any sorts of text classification task but in this research all three data sets were used for sentiment analysis after minor preprocessing such as case folding, stop word removing, etc.

3.1.1. Dataset 1

The first data set that has been used in this research is Yelp restaurant review data set, which was collected from the Kaggle machine learning repository. The data set contains two attributes namely review and sentiment, the value of first one attribute is the review text and the value of next one is the sentiment category, 0 for negative sentiment category and 1 for positive sentiment category, for the corresponding review text. This data set has 992 reviews with their corresponding sentiment categories.

3.1.2. Dataset 2

The second data set that has been used in this research is Amazon cell phones and accessories review data set, which was collected from the Kaggle machine learning repository. The data set contains two attributes namely review and sentiment, the value of first one attribute is the review text and the value of next one is the sentiment category of the corresponding review text. This data set has 1000 reviews with their corresponding sentiment categories.

3.1.3. Dataset 3

The third data set that has been used in this research is movie review data set, which was collected from the Kaggle machine learning repository. The data set contains two attributes namely review and sentiment, the value of first one attribute is the review text and the value of next one is the sentiment category of the corresponding review text. This data set has 25000 reviews with their corresponding sentiment categories.

3.2. Tools Used

All three machine learning based algorithms for sentiment analysis have been implemented in python programming language by using Pycharm IDE.

3.3. Data Preprocessing

The preprocessing phase aims to prepare unstructured opinions text data (reviews) ready for further processing. Preprocessing step that were conducted in this research includes:
Case folding: DO NOT waste your time on this 'film -> do not waste your time on this film.
Stop word removing: do not waste your time on this film->do not waste your time this film.
Stemming: do not waste your time this film ->do not wast yo tim thi film. And
Lemmatization: do not waste your time this film-> do not waste your time this film.

3.4. Feature Selection and Feature Vector Construction

One inherent problem of a computer is that it cannot process text data directly. So need to represent text data in numeric form. Generally, terms are used as features to represent the text. This leads to high dimension in the text representation. To improve the classification performance and processing efficiency, features need to be filtered to reduce dimension and remove noise [17].

3.4.1. TFIDF

TFIDF is a method to calculate the numeric weight for each (t_i) in each document (d_j) as: $TF - IDF_{ij}(t_i, t_j, D) = TF_{ij} * \log\left(\frac{N}{1+DF_i}\right)$ where TF_{ij} is the occurrence frequency of the term t_i in the document dj. N, is the total number of document in training set. DF_i, total number of documents containing term t_i.
TF-IDF represents the importance of terms in the training set (D). But the problem with this method is that it is unable to represent the association between features and categories. One possible solution to this problem is to use feature selection method called the chi-square method. The chi-square method helps to measure the association between term t_i and class c_k. In addition chi-square method holds the strongest noise tolerance ability [17] [18].

3.4.2. Chi-Square Method

Chi-square is a method to find the top k features as follows:

First, feature selection value is calculated as $\chi^2(t_i,c_k) = \dfrac{N*(ad - bc)^2}{(a+c)*(b+d)*(a+b)*(c+d)}$ where, N, Total number of documents in training set. a, Number of documents with term ti and belong to a category c_k. b, Number of documents with term ti and do not belong to category ck. c, Number of documents without term ti and belong to a category ck. d, Number of documents without term t_i and do not belong to a category ck. The Higher value of $\chi^2(t_i,c_k)$ indicates the closer relationship between term t_i and class c_k and $\chi^2(t_i,c_k)=0$ indicates independent relationship between term t_i and class c_k. Second, the score of the term t_i in the entire training set (D) is calculated as: Max $\chi^2(t_i,c_k) = \underset{k=1}{\overset{|c|}{Max}}\{(t_i,c_k)\}$. Third features are ranked in descending order in terms of the feature selection values. Finally, choose the top K features [17] [18].

3.4.3. Combined Chi-Square and TFIDF Method

In this research, the document feature matrix was formed by multiplying feature selection value and TF-IDF value and normalized the product.

3.5. Classification Algorithms for Sentiment Analysis

There are many popular and widely used classification algorithms for sentiment polarity identification of opinions of users based on the given opinion data. Among them the most commonly used and popular classification algorithms that were compared in this research are discussed below.

3.5.1. Multinomial Naïve Bayes Algorithm

Naive Bayes is a family of algorithms based on applying Bayes theorem with a strong (naive) assumption, that every feature is independent of the others, in order to predict the category of a given sample. They are probabilistic classifiers, calculate the probability of each category using Bayes theorem, and the category with the highest probability is output [19].
MNB is a probabilistic classifier, meaning that for a document d, out of all classes $c_k \in C$ the classifier returns the class c_k which has the maximum posterior probability. MNB is always a preferred method for any sort of text classification as taking the frequency of the word into consideration, and get back better accuracy than just checking for word occurrence [20].

3.5.2. K-Nearest Neighbor Algorithm

KNN is a non-parametric, lazy learning algorithm [21]. To identify the sentiment of new test document KNN classifier computes the similarity between a new test document and every training document. Then KNN classifier sort the training documents in descending order of their similarity to the test document in order to pick the top K most similar training documents with a test document. Finally the KNN classifier assigns this new test document to a sentiment category that has the highest score of similarity [22] [23].

3.5.3. Support Vector Machines Algorithm

A Support Vector Machine (SVM) performs classification by finding the hyper plane (classifier) that maximizes the margin between the two classes subject to the constraint that all the training tuples should be correctly classified. Hyper plane is defined by using the data points that are closest to the boundary. These points are called support vectors and the decision boundary itself is called support vector machine. The main advantage of SVM classifier is that it minimizes the training set error and the test set error [24].
To obtain a SVM classifier with the best generalization performance, appropriate training is required. The most commonly used and popular algorithm for training SVM is the SMO algorithm [24]. The main advantage of SMO algorithm is that it works analytically on a fixed size working set by decomposing the large training data set. So, that it can works fine even for large data sets and it also gives superb performances in almost all kinds of training data sets [25].

3.6. Evaluation Metrics

The comparative analysis of MLBCAs for sentiment analysis was made by measuring the performance of each algorithm with the help of following parameters.

3.6.1. Confusion Matrix

A confusion matrix is a table for analyzing the result of sentiment analysis by using classification algorithms. It deals with how classification algorithm can recognize documents of different sentiment class (Either positive or negative). In order to develop the confusion matrix, the following terms should be considered [17]. True Positive (TP) is the number of Positive sentiment documents that are correctly labeled by the MLBCAs for sentiment analysis. True Negative (TN) is the number of Negative sentiment documents that are correctly labeled by MLBCAs for sentiment analysis. False Positive (FP) is the number of Negative sentiment documents that are incorrectly labeled as positive. False Negative (FN) is the number of Positive sentiment documents that are mislabeled as negative.

	Predicted positive sentiment	Predicted negative sentiment
Actual positive sentiment	TP	FN
Actual negative sentiment	FP	TN

Fig 2. Confusion Matrix

Accuracy

Accuracy of classification algorithm for sentiment analysis on given data dataset is the percentage of documents in a data set that are correctly classified as positive sentiment or negative sentiment. It also refers to the polarity detection rate of the classification algorithm for sentiment analysis.

$$\text{Accuracy} = \frac{TP + TN}{TP + FN + FP + TN}$$

Precision

Precision refers to the measure of exactness that means what percentage of documents labeled as positive sentiment category are actually such.

$$\text{Precision} = \frac{TP}{TP + FP}$$

Recall

Recall refers to the true positive or positive polarity that means the proportion of positive polarity documents that are correctly identified.

$$\text{Recall} = \frac{TP}{TP + FN}$$

F-Measure

The F-measure combines both measures precision and recall as the harmonic mean.

$$\text{F-measure} = \frac{2 \times precision \times recall}{pression + recall}$$

4. Experiments and Results

In this research study, the analysis of all three MLBCAs for sentiment analysis mentioned in above section has been compared for all three sentiment labeled sentence datasets mentioned in above section, which are compared based on four performance criteria's namely accuracy, precision, recall and F-measure. The results have been achieved by using whole test dataset for all three algorithms after training phase. In this research work the value of K- parameter for KNN algorithm has been taken k=9, because for the datasets taken in this research KNN algorithm gave better result on this value of K than other possible value of K.

4.1. Performance result of MLBCAs for sentiment analysis on dataset1 and their comparison

The dataset1 (i.e. restaurant review dataset) has 992 tuples in total, but after preprocessing only 702 tuples among them 348 belongs to 1 (i.e., positive sentiment) category and 354 belongs to 0 (i.e., negative sentiment) category. For training only 561 tuples has been taken and remaining 141 (65 belongs to 0 and 76 belongs to 1) for testing purpose. Table 1 shows the classification report that has been obtained after three MLBCAs applied on test dataset that is obtained from dataset1.

Table 1: Confusion matrix on dataset1

	KNN		MNB		SVM	
	Predicted Positive	Predicted Negative	Predicted Positive	Predicted Negative	Predicted Positive	Predicted Negative
Actual Positive	54	22	58	18	60	16
Actual Negative	16	49	12	53	13	52

Based on the classification report shown in Table 1 the calculated summary performance result for the comparison of all three algorithms applied on dataset1 is shown in Table 2. The precision, recall and F-measure value shown in Table 2 is the average of precision, recall and F-measure for both sentiment categories.

Table 2: Performance result on dataset1

Algorithms	Accuracy	Precision	Recall	F-Measure
KNN	73%	73%	73%	73%
MNB	78.7%	79%	79%	78.5%
SVM	79.5%	79.5%	79%	79.5%

Based on the Figure 3, it is clearly seen that the accuracy value of SVM is got high level of 79.5% and KNN got less accuracy of level 73%. In case of precision and the Recall value of implemented SVM for sentiment analysis had got high precision and recall level of 79.5% and 79% respectively. Whereas, KNN got less precision and recall level of 73% and 73 % respectively. Figure 3 also shows the F-measure of table 2 observed by implemented Machine learning based classification algorithms for sentiment analysis. Again SVM had outperformed two other compared algorithms with value of 79.5% and KNN had got minimum value of 73%.

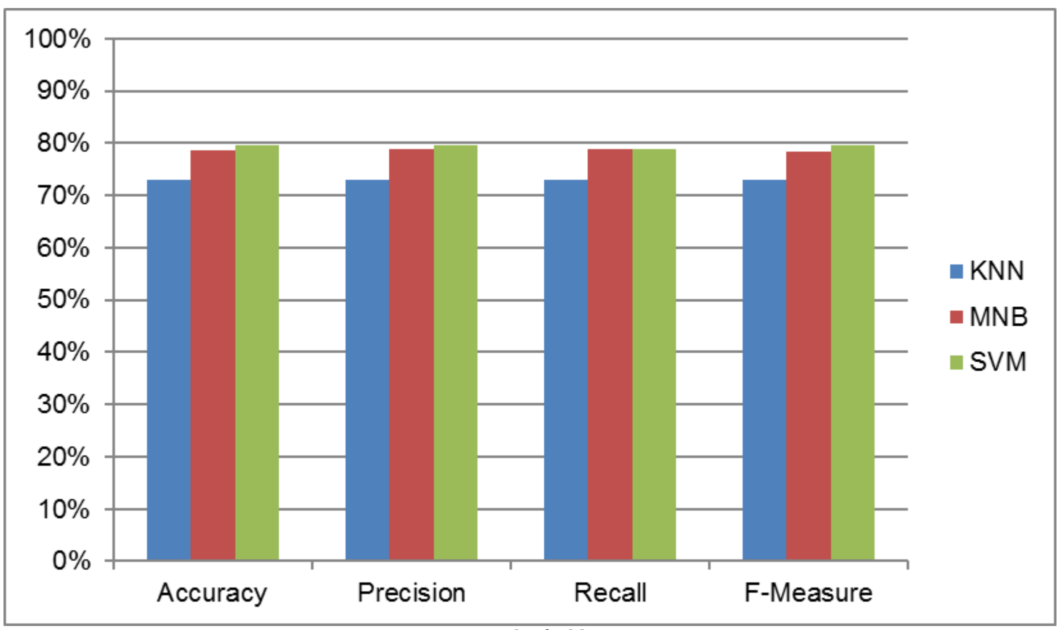

Fig 3. Graph of table 2

4.2. Performance result of MLBCAs for sentiment analysis on dataset2 and their comparison

The dataset2 (i.e. Amazon cell phones and accessories review dataset) has 1000 tuples in total, but after preprocessing only 770 tuples among them 387 belongs to 1 (i.e., positive sentiment) category and 383 belongs to 0 (i.e., negative sentiment) category. For training only 616 tuples has been taken and remaining 154 (77 belongs to 0 and 77 belongs to 1) for testing purpose.

Table 3 shows the classification report that has been obtained after three MLBCAs applied on test dataset that is obtained from dataset2.

Table 3: confusion matrix on dataset2

	KNN		MNB		SVM	
	Predicted Positive	Predicted Negative	Predicted Positive	Predicted Negative	Predicted Positive	Predicted Negative
Actual Positive	66	11	65	12	60	17
Actual Negative	25	53	25	52	16	61

Based on the classification report shown in Table 3 the calculated summary performance result for the comparison of all three algorithms applied on dataset2 is shown in Table 4. The precision, recall and F-measure value shown in Table 4 is also the average of precision, recall and F-measure for both sentiment categories.

Table 4: Performance result on dataset2

Algorithms	Accuracy	Precision	Recall	F-Measure
KNN	77.3%	77.5%	78%	77%
MNB	75.9%	76%	76.5%	76%
SVM	78.6%	78.5%	78.5%	78.5%

Based on the Figure 4, it is clearly seen that the accuracy value of SVM is got high level of 78.6% and MNB got less accuracy of level 75.9%. In case of precision and the Recall value of implemented SVM for sentiment analysis had got high precision and recall level of 78.5% and 78.5% respectively. Whereas, MNB got less precision and recall level of 76% and 76.5 % respectively. Figure 4 also show the F-measure of table 4.4 observed by implemented MLBCAs for sentiment analysis. Again SVM had outperformed other two compared algorithms with value of 78.5% and MNB had got minimum value of 76%.

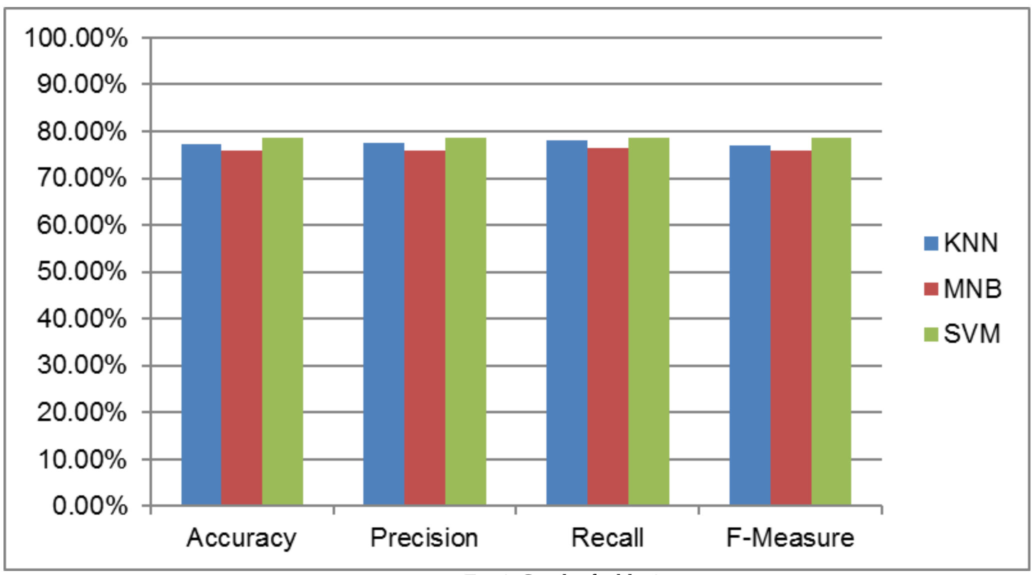

Fig 4. Graph of table 4

4.3. Performance result of MLBCAs for sentiment analysis on dataset3 and their comparison

The dataset3 (i.e. IMDB movie review dataset) has 25000 tuples in total, but after preprocessing all remains as it is among them 12500 belongs to 1 (i.e., positive sentiment) category and 12500 belongs to 0 (i.e., negative sentiment) category. For training only 20000 tuples has been taken and remaining 5000 (2548 belongs to 0 and 2452 belongs to 1) for testing purpose.
Table 5 shows the classification report that has been obtained after three MLBCAs applied on test dataset that is obtained from dataset3.

Table 5: confusion matrix on dataset3

	KNN		MNB		SVM	
	Predicted Positive	Predicted Negative	Predicted Positive	Predicted Negative	Predicted Positive	Predicted Negative
Actual Positive	1911	541	2163	289	2213	239
Actual Negative	497	2051	283	2265	288	2260

Based on the classification report shown in Table 5 the calculated summary performance result for the comparison of all three algorithms applied on dataset3 is shown in Table 6. The precision, recall and F-measure value shown in Table 6 is also the average of precision, recall and F-measure for both sentiment categories.

Table 6: performance result on dataset3

Algorithms	Accuracy	Precision	Recall	F-Measure
KNN	79.24%	79%	79%	79.5%
MNB	88.56%	88.5%	88.5%	88.5%
SVM	89.46%	89.5%	89%	89.5%

Based on the Figure 5, it is clearly seen that the accuracy value of SVM is got high level of 89.46% and KNN got less accuracy of level 79.24%. In case of precision and the Recall value of implemented SVM for sentiment analysis had got high precision and recall level of 89.5% and 89% respectively. Whereas, KNN got less precision and recall level of 79% and 79% respectively. Figure 5 also show the F-measure of table 6 observed by implemented MLBCAs for sentiment analysis. Again SVM had outperformed other two compared algorithms with value of 89.5% and KNN had got minimum value of 79.5%.

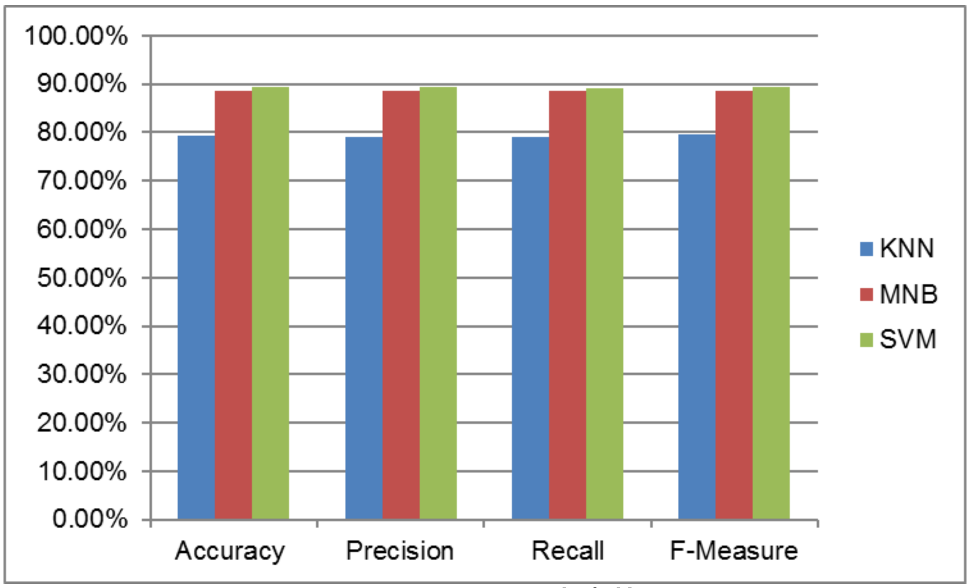

Fig 5. Graph of table 6

4. Conclusion

In this research, the comparative analysis of machine learning based classification algorithms for sentiment analysis using various performance measure parameters like accuracy, precision, recall and F-measures over the three different dataset with different size are evaluated. From this comparative study it has been found that, in Yelp Restaurant review dataset SVM had got higher performance in every aspect, whereas MNB and KNN had got less performance in every aspect as compared to SVM. The SVM algorithm has accuracy, precision, recall and F-measure with level of 79.5%, 79.5%, 79% and 79.5% respectively. In Amazon cell phones and accessories review dataset again SVM had got higher performance in every aspect, whereas KNN and MNB had got less performance in every aspect as compared to SVM. The SVM algorithm has accuracy, precision, recall and F-measure with level of 78.6%, 78.5%, 78.5% and 78.5% respectively. Also in the IMDB movie review dataset SVM had got higher performance in every aspect, whereas MNB and KNN had got less performance in every aspect as compared to SVM. The SVM algorithm has accuracy, precision, recall and F-measure with level of 89.46%, 89.5%, 89% and 89.5% respectively. Therefore, it has been concluded that, on balance datasets, SVM algorithm has predicted better Sentiment category result than other machine learning based classification algorithms for sentiment analysis studied for all three datasets.

5. Future Enhancement

In this research study only three traditional machine learning based classification algorithm has been study for sentiment polarity detection in three sentiment labeled sentence datasets. In the future more algorithms from the classification, clustering, and deep learning approach can be incorporated for further study to the studied datasets or other datasets which have text as well as image, audio or video type. Moreover some algorithms can be customized for the specific domain so that sentiment analysis could have more accurate and reliable results.

References

[1] A. Z. H. Khan, A. Mohammad and V. M. Thakare, "Sentiment Analysis Using Support Vector Machine," *International Journal of Advanced Research in Computer Science and Software Engineering,* vol. 5, no. 4, pp. 105-108, 2015.

[2] R. CL and G. S, "Machine Learning based Analysis of Twitter Data to Determine a Person's Mental Health Intuitive Wellbeing," *International Journal of Applied Engineering Research,* vol. 13, no. 21, pp. 14956-4963, 2018.

[3] M. Ghosh and G. Sanyal, "An ensemble approach to stabilize the features for multi-domain sentiment analysis using supervised machine learning," *Journal of Big Data,* vol. 5, 2018.

[4] G. Isabelle, W. Maharani and I. Asror, "Analysis on Opinion Mining Using Combining Lexicon-Based Method and Multinomial Naive Bayes," in *International Conference on Industrial Enterprise and System Engineering*, 2018.

[5] S. M. Vohra and J. B. Teraiya, "A Comparative Study of Sentiment Analysis Techniques," *Journal of Information, Knowledge and Research in Computer Engineering,* vol. 2, no. 2, pp. 313-317.

[6] M. Annett and G. Kondrak, "A Comparison of Sentiment Analysis Techniques: Polarizing Movie Blogs," in *21st Conference of the Canadian Society for Computational Studies of Intelligence:Advances in Artificial Intelligence*, 2008.

[7] L. Lopes, V. Machado and R. Rabelo, "Automatic Cluster Labeling through Artificial Neural Networks," in *International Joint Conferenceo on Neural Networks*, 2014.

[8] B. Pang, L. Lee and S. Vaithyanathan, "Thumbs up? Sentiment Classifcation using Machine Learning Techniques," in *Conference on Empirical Methods in Natural Language Processing (EMNLP)*, 2002.

[9] H. P. Rahmath and T. Ahmad, "Sentiment Analysis Techniques - A Comparative Study," *International Journal of Computational Engineering & Management,* vol. 17, no. 4, pp. 25-29, 2014.

[10] A. Kathuria and S. Upadhyay, "A Novel Review of Various Sentimental Analysis Techniques," *International Journal of Computer Science and Mobile Computing,* vol. 6, no. 4, pp. 17-22, 2017.

[11] E. Elmurngi and A. Gherbi, "Fake Reviews Detection on Movie Reviews through Sentiment Analysis Using Supervised Learning Techniques," *International Journal on Advances in Systems and Measurements,* vol. 11, no. 1 & 2, pp. 196-207, 2018.

[12] Z. Jianqiang and G. Xiaolin, "Comparison Research on Text Pre-processingMethods on Twitter Sentiment Analysis," *IEEE Access,* vol. 5, pp. 2870-2879, 2017.

[13] S. D. Sarkar and S. Goswami, "Empirical Study on Filter based Feature Selection Methods for Text Classification," *International Journal of Computer Applications,* vol. 81, no. 6, pp. 38-43, 2013.

[14] K. Huda, M. T. Nafis and N. K. Shaukat, "Classification Technique for Sentiment Analysis of Twitter Data," *International Journal of Advanced Research in Computer Science,* vol. 8, no. 5, pp. 2551-2555, 2017.

[15] B. Trstenjaka , S. Mikacb and D. Donkoc, "KNN with TF-IDF Based Framework for Text Categorization," in *Procedia Engineering*, 2014.

[16] S. Ahmed, S. Hina and R. Asif, "Detection of Sentiment Polarity of Unstructured Multi-Language Text from Social Media," *International Journal of Advanced Computer Science and Applications,* vol. 9, no. 7, pp. 199-203, 2018.

[17] M. Mowafy, A. Rezk and H. M. El-Bakry, "An Efficient Classification Model for Unstructured Text Document," *American Journal of Computer Science and Information Technology,* vol. 6, no. 1:16, 2018.

[18] U. I. Larasati, M. A. Muslim, R. Arifudin and A. Alamsyah, "Improve the Accuracy of Support Vector Machine Using Chi Square Statistic and Term Frequency Inverse Document Frequency on Movie Review Sentiment Analysis," *Scientific Joural of Informatics,* vol. 6, no. 1, pp. 138-149, 2019.

[19] M. Abbas, K. A. Memon, A. A. Jamali, S. Memon and A. Ahmed, "Multinomial Naive Bayes Classification Model for Sentiment Analysis," *International Journal of Computer Science and Network Security,* vol. 19, no. 3, pp. 62-67, 2019.

[20] D. Jurafsky and J. H. Martin, Speech and Language Processing: An Introduction to Natural Language Processing,Computational Linguistics, and Speech Recognition, 2019, pp. 56-74.

[21] R. Alhutaish and N. Omar, "Arabic Text Classification using K-Nearest Neighbour Algorithm," *The International Arab Journal of Information Technology,* vol. 12, no. 2, pp. 190-195, 2015.

[22] H. Deka and P. Sarma, "A Machine Learning Approach for Text and Document Mining," *International Journal of Computer Science Engineering,* vol. 6, no. 5, pp. 115-123, 2017.

[23] S. Tan, "An effective refinement strategy for KNN text classifier," *Expert Systems with Applications,* vol. 30, pp. 290-298, 2006.

[24] J. Platt, "Sequential Minimal Optimization: A Fast Algorithm for Training Support Vector Machines," *Technical Report MSR-TR-98-14, Microsoft Research,* 1998.

[25] K. P. Bennett and C. Campbell, "Support vector machines: hype or hallelujah?," *SIGKDD Explorations,* vol. 2, no. 2, pp. 1-13, 2000.

Tekendra Nath Yogi, was a student of Master of Science in Computer Science and Information Technology, Tribhuvan University, Nepal. He is actively involved in researches related data mining and machine learning. He is also involved in teaching different subjects related with Computer Science and Information Technology.

Nawaraj Paudel, is a Head of Department of Central Department of Computer Science and Information Technology, Tribhuvan University. He is actively involved in teaching and researches related with databases, data mining and machine learning.