

# Performance Modeling of Classification Techniques on Movie Sentiments



Pooja Rani, Jaswinder Singh

**Abstract:** The sentiment-based social media represents a gold-mine approach for analyzing the performance of the products, hotels, movies, politics, etc. Large opinions of the people are found over movie comments that are honest, informative, and casual as compared to the formal type of data-survey modeling using magazines or reports. The work proposed is based on the rating of movies. This paper analyzes the performance of classifiers for the prediction of sentiment class i.e., positive and negative by using artificial neural network, k-nearest neighbor and hybrid approach. The success of these classification techniques depends mainly on the appropriate extraction of the set of characteristics used to detect sentiments. Hybrid of two or more classifiers is mainly used to enhance the results. In the proposed experiment Hybrid of ANN and KNN shows improvement in precision and accuracy than other classifiers.

**Keywords:** Sentiment Analysis, Artificial Neural Network, K-Nearest Neighbor, Hybrid, Sentiments.

## I. INTRODUCTION

Sentiment analysis is an expansion of information mining which utilizes machine learning techniques to track the emotion of everyone about specific entity [7]. It gives the summated opinion of user. The incredible effect of web-based life has prompted the disclosure of estimation investigation [5]. The ongoing developments of keen innovations utilize portable based correspondence has involved enormous measure of information creation. People's sentiments and understanding are beneficial in decision making process, but to get profits from these comments and understanding, the gathered content should be mined and analyzed properly. The web-based life gives a capacity to share considerations and feelings with the establishment of famous networking sites like Facebook, Twitter, YouTube, Instagram etc. Information through these sites is easily accessible anytime, anywhere via desktop, laptop or mobile. The term sentiment classification is popularly recognized as opinion mining which is a process of emotion classification usually conveyed by a text that may be positive, negative or neutral given in fig. 1. It is profitable for users who need to find the sentiments of brands before any investment and also for the organizations who ought to know the opinions of public about any brand.

Examination of emotions can be useful in many areas such as decision making, policy making, product purchasing, quality improvement, market research, etc. [2].



Figure 1: Opinions

This paper considers different classification techniques of supervised machine learning [3][9] and includes following sections : Introduction, Types of Sentiment Analysis, Related Work, Proposed Methodology, Algorithm, Result, and Conclusion.

## II. TYPES OF SENTIMENT ANALYSIS

To see how to apply the examination of sentiments with regards to your business activity – one has to comprehend its various kinds. In this area, we will take a glimpse at the principle types of sentiment inspection.

**1. Fine-grained SA:** It is utilized to find the extremity of the assessment. It very well may be straightforward double positive or negative estimation dissimilarity. This sort can likewise go into higher portrayal (for instance positive, neutral, very positive, exceptionally negative), dependent upon the utilization case (for example as in five-star Amazon audits).

**2. Emotion acknowledgement:** It is utilized to distinguish indications of explicit emotional states exhibited in the content. Normally, there is a blend of learning algorithms and lexicons that figure out why and what kind of situations.

**3. Feature-based SA:** Its motivation is to distinguish an opinion with respect to a particular part of the item. For instance, sound of the speaker in the advanced smart phones. The evaluation i.e. aspect-based is generally utilized in item examination to watch out for how the item is seen and what are the solid and powerless focuses from the client perspectives.

**4. Intent Analysis:** Intent investigation is about the activity. Its motivation is to figure out what sort of expectation is communicated in the message. It is ordinarily utilized in client emotionally supportive networks to make an association or framework increasingly productive and compelling by utilizing quicker or more straightforward working strategies.

Revised Manuscript Received on December 30, 2019.

\* Correspondence Author

**Pooja Rani**, M.Tech Scholar, Department of Computer Science & Engineering, Guru Jambheshwar University of Science & Technology, Hisar, India. Email: poojabhatia46@gmail.com

**Jaswinder Singh\***, Assistant Professor Department of Computer Science & Engineering, Guru Jambheshwar University of Science & Technology, Hisar, India. Email: jaswinder\_singh\_2k@rediffmail.com

© The Authors. Published by Blue Eyes Intelligence Engineering and Sciences Publication (BEIESP). This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>)

### III. RELATED WORK

O. Kaynar, *et al.* [1] presented their work by applying three distinct classification methods i.e. Support Vector Machine, Artificial Neural Network and Centroid Based Algorithms (SVM, ANN, and CBA) have been applied together with feature selection methods i.e. GI, IG, and GA on the preprocessed tweet data of followers of GSM operators for sentiment exploration. Ensembling of feature selection method and classification algorithm are used to find the succession rate of modeling. It is specially used when classifier Support Vector Machines are used with the Genetic Algorithms as a hybrid. 96.8% favorable result has been gained for the analysis of the comments as +ve or -ve. M. M. Fouad, *et al.* [2] scheduled a miniature for analysis of tweeter-based sentiments describing weather the is or negative or positive form through the machine learning concept. The model proposed uses distinct type of methods for labeling the training phase input by using distinct forms of datasets. The method of classification was performed by the use of different types of classifiers in order to correlate and compare certain performances. The methodology of information gain and feature-based selection is mainly used in such type of work. G. Shidaganti, *et al.* [3] investigated over a method that presented a combined process of machine learning and data mining. The work proposed was done over tweeter data for the purpose of analyzing the tweet to collect the opinions of the user in regard to a specific type of topic or issue. The platform of tweeter is mainly used by individuals for expressing their sentiments (views) in a short type of message in response to several type of products, brands, celebrities, along with political criticisms. In this type of work, clustering and TF-IDF were mainly presented with their effective efficiency. D. Mumtaz, *et al.* [4] viewed and expected a methodology presenting a combined form of lexical-based and machine learning approach. Combination proposed provides high amount of accuracy than the classical-type of lexical method and it helps in providing the enhanced form of redundancy than the approach of machine learning. The approach i.e. proposed is mainly used for opinion/sentiment mining through NLP which helps in extracting the opinions from the content associated with an item. Mensikova, *et al.* [5] motivated the use of techniques of sentiment exploration for human trafficking analyzation. The conventional method did not focus over sentiment as human trafficking textual cue. Instead of the above, it focused over other cues or the textual cues. Here, the researchers applied two of the widely cited models of sentiment inspection i.e. Stanford and the Netflix. The binary and categorical models were also trained using escort data. The individual model exploratory and performance based has motivated construction of hybrid sentiment models which serves exactly as feature proxy in identifying human trafficking. Rajdev Tiwari, *et al.* [6] presented work on sentiment interpretation with respect to their distinct methods applied, architecture, different levels, tools used, applications, and research challenges. The research concluded that people usually express the feedback about any product, movie, book, etc. on shopping websites or social networking sites. Nowadays, if someone has to buy online product then he/she can view comments on the product site and take the right decision accordingly. B. Kaur, *et al.* [9] presented an approach which was used to identify the sentiments. The combination of SVM and KNN

was used, in which k-nearest neighbor classifier was working best for small length comments. The approach was experimented on SuperFetch comments. The approach concluded that combination of SVM and KNN produced better results on the basis of Recognition rate, Precision, TP rate, and F-Score. D. Virmani, *et al.* [10] explained the association of emotion with summarization, sentiment extraction, and further maintains the document of each student. In order to had collaborated and enhanced opinion (sentiment) about student helps in modifying the current algorithm. For analyzing the sentiments, a database has been used in the process. Firstly, it sets the score for any kind of sentiment or opinion word. In this process, when an opinion word is encountered in sentence, it performs the operation of matching with the set of database and further it sets the score as per the requirement. Bashir, *et al.* [11] focused on faced challenges of Twitter informational technology, focus over the ordering issues, and further it considers the streams for sentiment analysis and supposition mining. For the management of emitting unequal type of classes, the author has proposed a sliding or rolling window-based Kappa measurement for assessing informational streams that are of time-changing nature. The utilization of such type of measurements, the experts had investigated the information of Twitter that further utilizes the calculations for informational streams. Eduardo R. Hruschka, *et al.* [12] presented a new model that was useful for all the consumers who use sentiment anatomization for distinct products, for firms that target to monitor the sentiment of brands of public, and for many other kinds of applications. The classification of sentiments micro blogging services such as Twitter through the classifier gets ensemble and the lexicons were not explored so well explored in literature. The research on variety of sentiments of public tweet shows that the ensemble-based classifier got formed with the help of Logistic Regression, SVM, RF, and Multinomial NB that improved the accuracy. A. S. H. Basari, *et al.* [13] discussed the concept of opinions-based mining which referred to the function of computational linguistics, NLP and mining of the text in order to distribute the good or bad movie on the basis of message opinion. SVM mainly presents a supervised method of learning which helps in analyzing the data and recognizes the patterns used for the purpose of classification. Theresa Wilson, *et al.* [14] investigated the linguistic features-based utility for detection of sentiment of Twitter messages. The researchers further evaluated the lexical resources usefulness as well as the characteristics that were able to catch the creative and informal language that was recycled in the process of micro blogging. A supervised approach was taken to issue and it leveraged the hash tags that existed already in the Twitter data for designing of the trained type of data. R. Socher, *et al.* [15] discussed the concept of predicting the sentence-based level of label type distribution through a newly built approach established on the prediction of sentence level in recurrent type of auto-encoders. The work suggested was performed on sentiment shifting and lexica rules. Data used was the stories of personal user that was annotated with labels of multiple forms and it was further aggregated from multinomial kind of distribution capturing the sentimental reactions.

The related work concluded that combination (hybrid) of two or more classifiers produced better results on the basis of Accuracy, Precision, Recall, and F-Score. Next segment of the paper defines the technicalities related to the methodology followed.

#### IV. THE PROPOSED METHODOLOGY

This section of paper describes the detailed description of the number of steps followed for the research, programming language used, tools used and the proposed framework in fig. 2 to achieve the results.

##### Programming Language Used

- PYTHON 3.6.3
- PYCHARM (IDE)

##### Tools Used

- Scikit-Learn
- Tweet NLP
- NLTK
- Pandas

Number of steps followed and the techniques used for the research methodology to achieve results are given below in detail.

**Step1: Data Collection and Fetching:** Firstly, data will be collected for experiment from dataset i.e. IMDb. To conduct the experiment, the data which is given as input to the model has been taken from the Movie Sentiments. Dataset extracted from <https://www.kaggle.com/iarunava/imdb-movie-reviews-dataset>. After data collection, store it in database and define labels like positive, negative or neutral in order to reduce the subjective nature at sentiment annotation. Dataset contain 50K instances out of that 25K positive and 25K negative. It uses 30K for training and 20K for testing. Data is often incomplete, inconsistent, and contains errors. Pre-processing will resolve these issues in the next step.

**Step 2: Data Pre-processing:** Data preprocessing is the main step of the research that is used to clean the data. This has a major effect on the movie rating. During the data processing phase, links, usernames, punctuation marks, and the same sentences in related tweets were removed [4]. In addition, all words were converted to lower case. In this step reduce the noise from dataset by following steps [2]:

a) **Tokenization:** It is a process of text conversion as a string into elements process-able known as tokens. Tokens are separated by white space. In terms of tweets methodology, such elements can be emoticons, words, links, hashtags, etc [14].

b) **Stemming:** It is a process of removing the grammatical error. It includes verb, adverb, noun, pro-noun, and adjective for correction.

c) **Stop word removal:** It is a process of removing irrelevant data like commas, full stop, question mark, etc.

##### Proposed methodology: Flowchart

The figure below represents the framework of the proposed methodology. Comment of movie is given as the input and model provides classified data as the output. Next section of the paper describes the algorithms followed for the research.

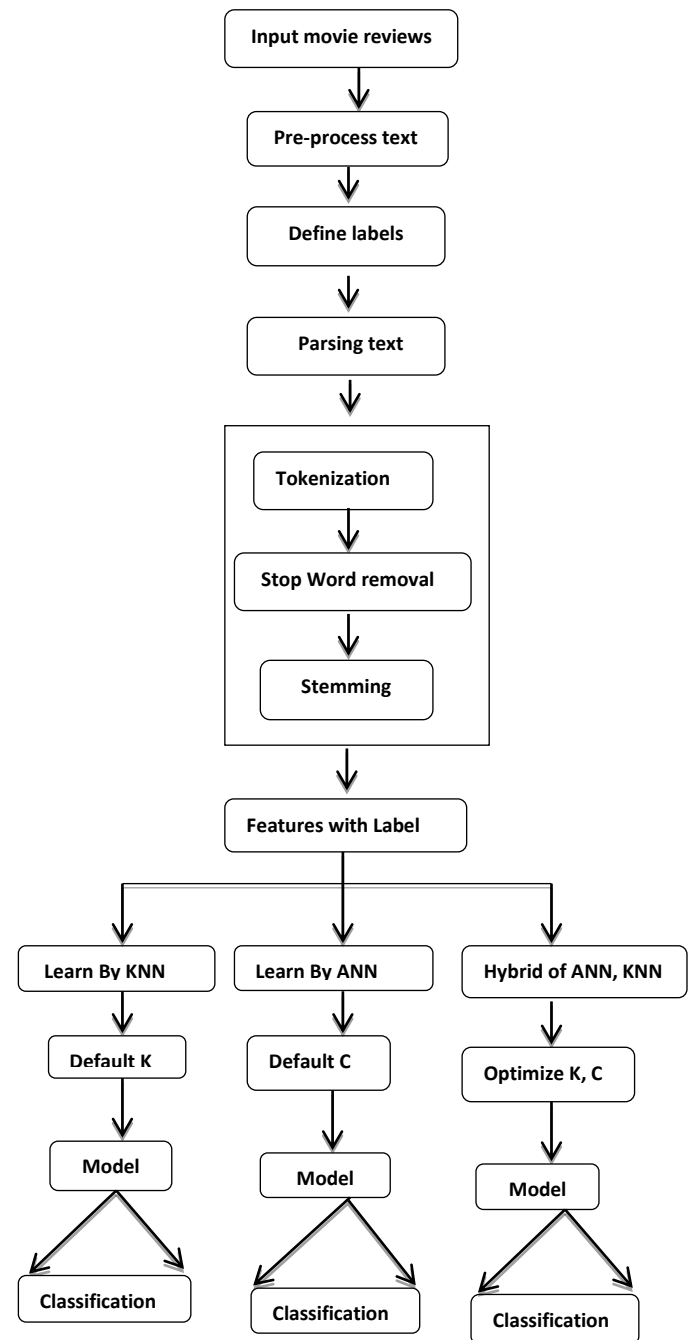


Figure 2: Proposed Framework

**Step 3: Feature extraction:** Feature extraction is a process of designing a feature-vector from known sentiments. In this process useful data will be taken from the dataset. Features are extracted from text with label. Here label will define the polarity like +1 or -1, it should be classified as positive and negative [9].

Extracted feature are very large in number. It increased the storage time and computation time so extracted features are presented in the form of matrix of finite dimensions.

**Step 4: Learning:** Artificial Neural Network, K-Nearest-Neighbor and Hybrid has the capability to classify the sentiments. Training samples are used for learning while test samples are used for estimating the accuracy of classifiers. Classifier modeling is the next process of learning that will distribute the comments into positive and negative. After feature extraction learning process is done by following approaches:

- **ANN**

Artificial Neural Network is three layers structure in which features are given by input layer then mapping on hidden layer and give the output at output layers [16].

- **KNN**

In K-Nearest Neighbor, firstly define the K value. In this research it takes default five then check testing movie opinions [17]. If comments are mostly near to positive then positive sentiment other-wise negative.

- **ANN-KNN (Hybrid)**

Grouping of two or more classifiers is known as hybrid process [13][11]. Artificial Neural Network and K-Nearest-Neighbor are example of supervised learning. In this classifier first make ANN model then define K in KNN and in last classify the movie sentiments.

**Step 5: Result optimization:** Result optimization is the last step that is used to check the model that no error is configured. As per the informational collection cross approval is utilized by splitting the dataset to get the exact outcome. The data gathered is split into two sets. One is learning or training set and another is testing set. Technique of K-fold is used to split the dataset into k mutually exclusive subset, almost equal size. Generally, K = 5 or 10 is preferred, but nothing's fixed and it can take any value. Flowchart of the proposed methodology is given in next section of the paper.

## V. ALGORITHM USED

The algorithm takes movie comments as input and interpretation of those emotions will be done in output process. Algorithm of ANN, KNN, and Hybrid has been mentioned below.

### Algorithm for ANN:

Input: Pre-process movie comment and initialize  $C=0.2$

Output: Analysis of classifier model

1.  $N \leftarrow$  no. of pre-process text

2. While ( $N > 0$ )

Start

$F \leftarrow$  Extract features

$M \leftarrow F[N]$

3.  $S_c \leftarrow$  Apply feed forward neural network  
(C, F, M)

For each  $n_i \in N$

do Back Propagation [ $B \in i$ ]

$B[i] \leftarrow N$

if  $N < C$

make model  $X[i]$

else

run step 3

4. return  $X[i]$

5. Analysis model

### Algorithm for KNN:

Input: Pre-process movie comment and initialize  $K=5$

Output: Analysis of classifier model

1.  $N \leftarrow$  no. of pre-process Text

2. while ( $N > 0$ )

Start

$F \leftarrow$  Extract features

$M \leftarrow F[N]$

3.  $S_K \leftarrow$  Compute nearest neighbour (F, M)

for each  $n_i \in N$

do  $V_j \leftarrow |M \cap G_i|$

4. return arg max  $V_j$

5. Analysis model

### Algorithm for Hybrid (ANN-KNN)

Input: Pre-process movie sentiment and initialize ( $C=0.2$  and  $K=5$ )

Output: Analysis of classification model

1.  $N \leftarrow$  no. of pre-process text

2. while ( $N > 0$ )

Start

$F \leftarrow$  Extract Features

$M \leftarrow F[N]$

3.  $S_K \leftarrow$  Compute nearest neighbour (F, M)

if ( $S_K < C$ )

then

for each  $n_i \in N$

do  $V_s \leftarrow |M \cap G_i|$

else

$S_c \leftarrow$  Apply feed forward (C,  $S_K$ , F, M)

if ( $S_c < C$ )

Meta Model  $X[i]$

Else

Run step 3

4. Return  $X[i]$

5. Analysis Model

## VI. RESULTS

This research presents the experiment on movie sentiments. In dataset, tweets are collected which have different classes. Proposed approach represents the comparative modeling of classifiers and their hybridization. Experimental inspection shows the results in form of Accuracy, Precision, Recall and F-score [8]. This section deals with the results obtained. Result will be analyzed with the help of confusion matrix [4].

**Confusion Matrix:** It is a popular way to represent the summarized findings. It defines how efficiently a classifier can observe dissimilar categories. It is a table of size n by n. Matrix for two classes is given below:



**Table 1 Confusion Matrix**

		Actual		
Predicted		C1	C2	
	C1	TP	FP	P'
	C2	FN	TN	N'
		Positive(P)	Negative(N)	ALL

Table 1 represents the confusion matrix of two classes i.e. Actual and Predicted.

TP: Positive tuple that are accurately classified.  
TN: Negative tuple that are accurately classified.  
FP: Negative tuple that are incorrectly classified.  
FN: Positive tuple that are incorrectly classified.  
P: Value after sum of TP and FN.  
N: Value after sum of FP and TN.  
P': Value after sum of TP and FP.  
N': Value after sum of FN and TN.

**Table 2 Different Parameters**

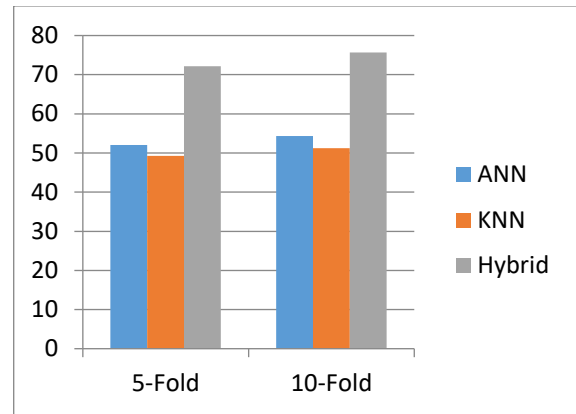
Accuracy	TP+TN/ALL
Recall	TP/TP+FN
Precision	TP/TP+FP
F-Measure	$2 \cdot \frac{\text{Precision} \cdot \text{Recall}}{\text{Precision} + \text{Recall}}$

In the given matrix sum of P' and N' is always equal to the sum of P and N. Table 2 represents the distinct parameters that are used in the research.

**Table 3 Accuracy of Classifiers**

No. of Validations	ANN	KNN	Hybrid
5-Fold	52	49.23	72.13
10-Fold	54.34	51.23	75.67

Table 3 shows the accuracy of used classifier i.e. ANN, KNN, and hybrid of both classifiers in different forms of validations. Accuracy of ANN in 5-Fold is 52 and in 10-Fold is 54.34. Accuracy of KNN in 5-Fold is 49.23 and in 10-Fold is 51.23. Hybrid of artificial neural network and k-nearest neighbor shows accuracy in 5-Fold is 72.13 and in 10-Fold is 75.67. Accuracy increases as we increase the number of validations. In the given above table maximum accuracy is shown by hybrid approach.



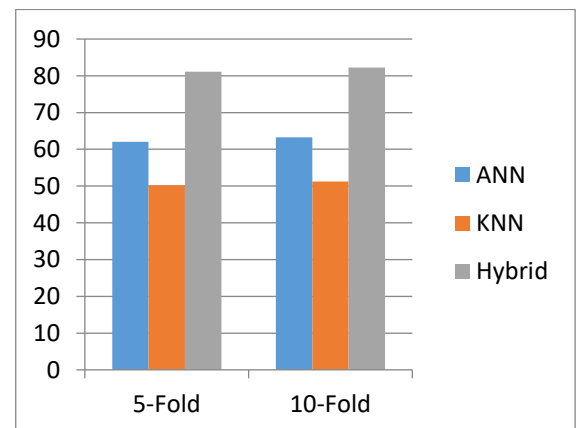
**Figure 3: Accuracy of Classifiers**

In fig. 3 it depicts the correctness of distinct classifiers that are ANN, KNN and Hybrid. X-axis on graph represents the validation fold and Y-axis represents the values of accuracy. The hybrid algorithm presents the highest accuracy in both folds. The lowest accuracy is 49.23 shown by KNN in 5-fold validation process and maximum accuracy is 75.67 in 10-fold process of Hybrid.

**Table 4 Precision of Classifiers**

No. of Validations	ANN	KNN	Hybrid
5-Fold	62	50.23	81.13
10-Fold	63.23	51.23	82.23

Table 4 represents the precision after using ANN, KNN, and Hybrid in distinct forms of cross validation folds i.e. 5-Fold and 10-Fold. Precision of ANN shown in 5-Fold validation is 62 and in 10-Fold is 63.23. Precision of KNN shown in 5-Fold validation is 50.23 and in 10-Fold is 51.23. Precision shown by Hybrid in 5-Fold is 81.13 and in 10-Fold is 82.23. In the given table hybrid algorithm shows the maximum precision.



**Figure 4: Precision of Classifiers**

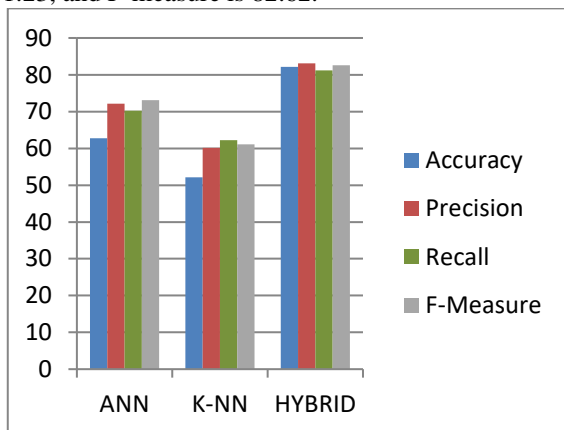
In fig. 4 it depicts the Precision of distinct classifiers that are ANN, KNN and Hybrid. X-axis on graph represents the validation fold and Y-axis represents the values of Precision. The minimum precision is 50.23 shown by KNN in 5-fold

validation process and maximum is 82.23 shown by hybrid of ANN and KNN in 10-fold cross validation. Fig. 3 and 4 in the paper shows the results of the classifiers in three distinct colors. Artificial Neural Network displayed by blue color, K- Nearest Neighbor displayed by orange color and Hybrid displayed by grey color.

**Table 5 Results of different Classifiers**

Classifiers	Accuracy	Precision	Recall	F-Score
ANN	62.73	72.13	70.23	73.13
K-NN	52.13	60.13	62.23	61.13
HYBRID	82.12	83.13	81.23	82.62

After experiment evaluation Table 5 represents the different parametric values. Value of overall accuracy of Artificial Neural Network is 62.73, precision is 72.13, recall is 70.23, and F-measure is 73.13. Result of overall accuracy of K-Nearest Neighbor is 52.13, precision is 60.13, recall is 62.23, and F-measure is 61.13. Value of overall accuracy of Hybrid of ANN and KNN is 82.12, precision is 83.13, recall is 81.23, and F-measure is 82.62.



**Figure 5: Analysis of Classifiers**

In fig. 5 depicts the comparison of the parametric values of ANN, KNN and Hybrid. X-axis on graph represents the classifiers and their hybridization. Y-axis represents the values of Accuracy, Precision, Recall and F-measure. Fig. 5 depicts the results of four parameters in distinct colors. Accuracy is displayed by blue color, Precision is displayed by red color, Recall is displayed by green color, and F-measure is displayed by grey color. The hybrid algorithm shows the maximum results on all parameters [13] and KNN shows the minimum result among all parameters.

## VII. CONCLUSION

The purpose of movie ranking is the supervision for people about films. They can suggest others what they should see and what they should not see. However, it does not justify that the movie is good or bad. It gives the main direction about languages, drugs, violence, concept of movie, etc. In this paper, the concept of artificial neural network, K-nearest neighbor, and hybrid of both classifiers are used for classification process. Such type of method

helps in analyzing different feature with an assigned class in order to identify the relation dependency between a sentiment and each of the features. These classifiers identify the features which divides the class into two types. One is the considered as positive and other is negative. Result concluded that Hybrid of ANN and KNN shows better results than other. More investigation is needed to increase the accuracy. Performance of classifiers can be further improved by combining it with other classifier and deep learning approaches can be used to make the sentimental analysis more effective.

## REFERENCES

1. Yelmen, M. Zontul, O. Kaynar, and F. Sonmez, "A Novel Hybrid Approach for Sentiment Classification of Turkish Tweets for GSM Operators," *International Journal of Circuits, Systems and Signal Processing*, vol. 12, pp. 637-645, 2018.
2. M. M. Fouad, T. F. Gharib, and A. S. Mashat, "Efficient Twitter Sentiment Analysis System with Feature Selection and Classifier Ensemble," in *The International Conference on Advanced Machine Learning Technologies and Applications (AMLTA2018)*, vol. 723, pp. 516-527, 2018.
3. G. Shidaganti, R. G. Hulkund, and S. Prakash, "Analysis and Exploitation of Twitter Data Using Machine Learning Techniques," in *International Proceedings on Advances in Soft Computing, Intelligent Systems and Applications*, vol. 628, pp. 135-146, Springer Singapore, 2018.
4. D. Mumtaz and B. Ahuja, "A Lexical and Machine Learning-Based Hybrid System for Sentiment Analysis," in *Innovations in Computational Intelligence*, vol. 713, pp. 165-175, Springer Singapore, 2018.
5. Mensikova Anastasija, and Chris A. Mattmann, "Ensemble Sentiment Analysis to Identify Human Trafficking in Web Data," *GTA3 2018*, Marina Del Rey, CA USA, 2018.
6. Surya Prakash Sharma, Dr Rajdev Tiwari, and Dr Rajesh Prasad, "Opinion Mining and Sentiment Analysis on Customer Review Documents- A Survey," in *International Journal of Advanced Research in Computer and Communication Engineering*, vol. 6, no. 2, pp. 156-159, 2017.
7. Suman, Jaswinder Singh, "Sentiment Analysis: A Survey," *International Journal for Research in Applied Science & Engineering Technology*, vol. 5, issue 8, pp. 1957-1963, 2017.
8. Suman, Jaswinder Singh, "Sentiment Analysis of Tweets using Support Vector Machine," *International Journal of Computer Science and Mobile Applications*, vol.5, issue 10, pp. 83-91, 2017.
9. B. Kaur and N. Kumari, "SVM and KNN based Hybrid Approach to Sentiment Analysis," in *International Journal of Technical Research & Science*, vol. 1, no. 5, pp. 67-74, 2016.
10. D. Virmani, V. Malhotra, and R. Tyagi, "Sentiment Analysis Using Collaborated Opinion Mining," vol. 4, p. 4, January, 2014.
11. F. H. Khan, S. Bashir, and U. Qamar, "TOM: Twitter opinion mining framework using hybrid classification scheme," *Decision Support System*, vol. 57, pp. 245-257, 2014.
12. Da Silva Nadia FF, Eduardo R. Hruschka, and Estevam R. Hruschka Jr. "Tweet sentiment analysis with classifier ensembles." In *Decision Support Systems* (2014): pp. 170-179, 2014.
13. A. S. H. Basari, B. Hussin, I. G. P. Ananta, and J. Zeniarja, "Opinion Mining of Movie Review using Hybrid Method of Support Vector Machine and Particle Swarm Optimization," *Procedia Engineering*, vol. 53, pp. 453-462, 2013.
14. Kouloumpis Efthymios, Theresa Wilson, and Johanna Moore, "Twitter Sentiment Analysis: The good the bad and the omg!" in *Proceedings of the Fifth International AAAI Conference on Weblogs and Social Media, (ICWSM-2011)*, pp. 538-541, 2011.
15. R. Socher, J. Pennington, E. H. Huang, A. Y. Ng, and C. D. Manning, "Semi-Supervised Recursive Autoencoders for Predicting Sentiment Distributions," in *Proceedings of the 2011 Conference on Empirical Methods in Natural Language Processing*, no. 2, pp. 151-161, 2011.
16. Pooja Rani and Jaswinder Singh, "Classification of Reviews using Artificial Neural Network," *International Journal of Electronics Engineering*, vol. 11, issue 1, pp. 882-888, 2019.

17. Pooja Rani and Jaswinder Singh, “Analysis of Sentiments using K-Nearest Neighbor,” *International Journal of Electronics Engineering*, vol. 11, issue 1, pp. 889-894, 2019.

#### AUTHORS PROFILE



**Pooja Rani**, M.Tech Scholar, Department of Computer Science & Engineering, Guru Jambheshwar University of Science & Technology, Hisar, Haryana, India.



**Jaswinder Singh**, Assistant Professor Department of Computer Science & Engineering, Guru Jambheshwar University of Science & Technology, Hisar, Haryana, India. He also has worked as Assistant Professor in the department of Computer Science & Engineering at Laxmi Devi Institute of Engineering & Technology, Chikani,

Alwar, Rajasthan and at Seth Jai Parkash Mukand Lal Institute of Engineering and Technology, Radaur, Yamunanagar, Haryana. He received his PhD in Computer Science & Engineering from department of Computer Science & Engineering, Deenbandhu Chhotu Ram University of Science & Technology, Murthal, Sonapat, Haryana. He has done M.Tech in Computer Science & Engineering from Kurukshetra University, Kurukshetra. His research interests include Web Information Retrieval, Web Mining, Search Engine Optimization.