**Sentiment Analysis and Opinion Feature Extraction of Movie Reviews**

**Lubaina Kamran Siddiqui**

**F2019114006@umt.edu.pk**

**University of Management and Technology, Lahore.**

**Abstract:** Sentiment analysis or opinion mining can be described as study of public opinion on politics, sports, products, movies or any other area of interest. Because of social media access, it has become rather easy to give out opinion on anything and it has made easier for the researchers to work because of immense amount of data. Sentiment analysis is a subset of Natural Language Processing, a subfield of Artificial Intelligence. In this paper, we have performed opinion mining of Movie reviews on Rotten Tomatoes Dataset using three algorithms; Logistic regression, Naïve Bayes and Decision Tree. This research serves as a comparison between mentioned three algorithms, we have focus on the training set as the main objective to present the retrieval of optimal probability thresholds using the ROC Curve and results are further compared using confusion matrix.

**Keywords**- Artificial Intelligence, Sentiment Analysis, NLP, Opinion Mining, Naïve Bayes, Decision Tree, Logistic Regression.

# **INTRODUCTION:**

Natural Language Processing, NLP is a subdomain of Artificial Intelligence. NLP mines meaning from natural languages and makes decision in accordance to that information. Sentiment Analysis or Opinion mining makes use of Natural Language Process to identify, extract, measure and study sentiments or meanings (Mtetwa et al., 2018). NLP research work has been around since many years, but its usage has greatly increased ever since digitalization and internet has sky rocketed. Basics of natural language processing lie in many different disciplines such as Math, Engineering, Robotics, Electronics and Artificial Intelligence (Chowdhury, 2019). Natural language processing is considered as one of the difficult areas because of the nature of human language. Predictive text in emails or in google search, advertisement on social media speech to text and text to speech transformation, language translation and much more. Earlier language based NLP techniques were the focus in limited fields but now a day’s researchers are more down towards using machine learning and Statistical methods for applying natural language processing techniques. This trend has given AI a new outlook to explore and predict analysis on sentiments of people; therefore, Sentiment Analysis came into being.

With the advent of Web 2.0 and 3.0, usage of social media sites, people from all over the world express their opinions about various topics now more openly (Kaur & Gupta, 2013). Since the advent of blogging on Facebook, Instagram, Twitter, etc. people share their views, opinion, and judgement by reviewing various products, movies, restaurants, etc.

Modern boom in technology such as big data, mobile technologies, and social media has allowed the creation of many new research opportunities and challenges. Sentiment Analysis, Opinion or Emotion mining provides organization an edge over their competitors by extracting insights about particular subjects. Monitoring social media gives a wider public point of view, predicts the ongoing trends and helps to make certain decisions. Organizations can improve their quality of products and services based on reviews by people. This gives a positive impression on their organization’s reputation. Sentiment Analysis assists people with the help of recommendation systems. Such systems plays a vital role in decision-making process of the user. Recommendation systems evaluate user’s opinion and sentiments and help to predict what kind of items are be recommended to users or not. The outcome of sentiment analysis can help recognize the stance of people towards certain policies and products.(M & Mehla, 2019)

Because of so much advancement and possibilities, multiple organizations have started mining opinion of users on various topics. One of the best examples is movie reviews. Movie reviews can help determine whether the reviewer has given a positive, negative or neutral opinion on a movie. This analysis of movie reviews can help us understand if the movie has created a good or bad reputation among other movies. People rate different movies by giving reviews and based on their opinions we can perform various machine learning and rule based techniques to determine whether the review is positive or negative.(Agarwal & Mittal, 2013).

In this paper, we are going to perform opinion mining based on movie reviews. Various algorithms help us determine the sentiment of movie for example Naïve Bayes, Logistic Regression and Decision Tree.

## **Levels of Opinion Mining:**

Opinion mining is performed on three granularity levels: (Chinsha & Joseph, 2015)

1. Document Level
2. Sentence Level
3. Aspect\Feature Level

**1. Document Level opinion mining:**

Document level sentiment analysis is the most simplest and easiest to achieve. The entire document is considered as a single entity. The objective of opinion mining in document is to determine overall polarity. Sentiment document analysis is achieved by supervised machine-learning methods. Various pre-processing steps like removing stop words, POS tagging, etc. are carried out on the document. In Supervised machine-learning methods model is trained on training data using algorithms like Naïve Bayes, SVM, KNN, etc.

**2. Sentence Level opinion mining:**

Sentence level sentiment analysis works in two main steps, First, whether the sentence is conveying an opinion or not. Second, to classify the sentiment into positive, negative and neutral class. Therefore, Sentence Level opinion mining includes two main steps:

* Subjectivity Classification
* Sentiment Classification

**2.1 Subjectivity Classification**

The objective of Subjectivity Classification is to differentiate between subjective sentences and objective sentences. Subjective sentences includes opinions, feelings and point of view. Objective sentence express facts that are universal. For example, “Pakistan is an agricultural country. It is a very beautiful country.” The first sentence that ‘Pakistan is an agriculture country’ is a fact and does not state any opinion whereas, the sentence ‘it is a beautiful county’ is an opinion, that is positive.

**2.2 Sentiment Classification**

Sentiment Classification of a Sentence determines the polarity of the sentence that is positive, negative or neutral. Some sentences are difficult to interpret their opinion for instance, sentences that contain sarcasm, interrogative, conditional and comparative words.

**3. Aspect/Feature Level opinion mining**

Aspect or Feature Level opinion mining refers to understanding of those sentences that give mix reviews for example, “The food is tasty but the ambience of restaurant is not good”. Now these kind of sentences are difficult to classify as they hold multiple opinions. Aspect or Feature Level Opinion Mining is therefore, divided into three steps:

* Identify the aspects
* Find the opinion related to the aspect
* Polarity of the word( positive, negative and neutral)

In our proposed study, we will use sentence-based opinion mining on movie review dataset, while exploring various classification algorithms like Linear SVC and Naïve Bayes using vectorization methods such as Count Vectorization and tf-idf.

In section 2, we present Literature Review. Section 3 includes the methodology that we have implemented furthermore; section 4 discusses the results and its analysis with related work.

# **Literature Review:**

## **Criteria for paper:**

In this section, we present the literature review. Following are the steps taken for gathering relevant literature for our research paper:

* **Searching Criteria:**

Our search criteria includes all papers from the year 2013 onwards. Online repositories such as IEEE, ACM, Google Scholar, Springer and Science Direct are part of literature review. Key words like opinion mining, feature extraction, NLP, sentiment analysis, etc. were taken into consideration for advance search to find relevant papers.

* **Inclusion Criteria:**

Inclusion Criteria of our paper includes those literature that is published in well-known journals that is ACM, IEEE, and Springer etc. are included as part of study. Thorough investigation of minimum 20 different research papers, are included based on search criteria that contains minimum 10 citations per paper.

* **Exclusion Criteria:**

Omission of Grey literature, such as presentations, reports, abstracts, unpublished papers, etc. from this study, while conducting literature review. Papers that are not in English language are excluded as well.

## **Related Work**

Some of the previous work by authors is mentioned below:

Authors (Moraes et al., 2013) performed document level sentiment analysis using SVM and ANN classifiers. The dataset chosen for training was the standard movie dataset (Pang & Lee, 2004) along with some reviews extracted from amazon. Results indicated that ANN outperformed SVM in terms of accuracy, precision and recall. However, in terms of noisy data SVM was less affected compared to ANN.

(V.K. Singh, R. Piryani, A. Uddin, 2013) implemented sentiment analysis using document and aspect level classification. A movie review dataset was used for evaluation. They used SentiWordNet and document level analysis presents a new scheme of features as Adverb Adjective combination (AAC) and Adverb Adjective and Adverb Verb combination (AAAVC). This new scheme was set in comparison to Alchemy API method. AAAVC outperformed both methods with an accuracy of 78.7%. The aspect-level of 10 movies gives score on different aspects and it is very easy to understand and incorporate.

Authors (Singh & Shahid Husain, 2014) used two datasets for sentiment analysis. One is product dataset from amazon.com, other is the (Pang & Lee, 2004) IMDB movie data set. Three classification algorithms SVM, Naïve Bayes and multi-level perceptron are implemented on N-gram feature. On both data sets, SVM performed the best with an accuracy of 81.15% for movie review dataset and 79.40% on product data set.

Authors (Sharma et al., 2014) performed opinion mining on document level. They used the IMDB movie dataset. The authors evaluated by comparing document based sentiment analysis with AIRC system. Document based sentiment orientation accomplished accuracy of 63%, which was better in contrast to AIRC system.

(Indhuja & Reghu Raj, 2014) proposed a methodology for sentiment analysis which uses fuzzy logic. The steps involved were pre-processing, feature extraction and fuzzy opinion classification. The authors used product review data set. They achieved a high accuracy of 85.58%.

(Chinsha & Joseph, 2015) did some research work on aspect-based opinion mining. They used restaurant reviews collected from TripAdvisor. In this proposed method, polarity of words that was determined by using adjective, SentiWordNet and combination of adverb verb and adverb adjective, had achieved an accuracy of 78.04%.

In this research, authors (Tang et al., 2015) authors propose a new model namely, User Product Neural Network (UPNN) for document level classification. The authors used the IMDB and Yelp Dataset for their study. Empirical results of this model concluded, that user and product representations incorporated together could give better results than state-of-art sentiment analysis classifiers.

In this document level sentiment analysis, the authors(Xia et al., 2016) propose a three stage model called Polarity Shift detection, Elimination and Ensemble to detect polarity shift in reviews. Polarity shift refers to those words that change the orientation of sentences and depict different meaning. The authors have used the (Blitzer et al., 2007) dataset. In the first step, the authors suggested a hybrid approach that includes rule based and machine learning methods to identify polarity shift words. The second step involves antonym revision to remove polarity shifts. Third step comprises of weighted ensemble on the dataset. The results showed significant improvement compared to other classifiers.

(Sahu & Ahuja, 2016) authors have performed a study on sentiment analysis of movie review data set and evaluated using classification algorithms like Naïve Bayes, KNN, Bagging, COCR, Random Forrest and Decision Tree. Random Forrest scored the highest accuracy of 88.95% among other classifiers.

(Saberi & Saad, 2017) These authors conducted a non-experimental review of sentiment analysis describing its main use, levels of sentiment analysis, and resources of sentiment analysis.

(Kai et al., 2017) The authors propose a hybrid model, by combining SVM and GBDT. They conclude that SVM works well under the condition that sentences have a simple structure and strong opinion words whereas GBDT works well when sentence is too long and has more polarity words.

(Zvarevashe & Olugbara, 2018) Authors used Sentence polarity Based Model (SPBM). They used OpinRank dataset containing 259000 reviews of hotel. Four main classification algorithms were used namely Naïve Bayes, Sequential Minimal Optimization (SMO), Compliment Naïve Bayes (CNB) and Composite hyper cubes on Iterated Random Projections (CHIRP). The results obtained after experimentation was that Naïve Bayes got highest accuracy of 80.9% while CNB was 80.5% and CHIRP 75.6%.

(Penubaka Balaji, 2018) says there are three main categories of sentiment analysis based on corpus. Document level sentiment analysis, Sentence level analysis and lastly aspect-based which classifies features and objects.

There can be three common components of any opinion.

* Opinion holder: One who gives the opinion, for example, I do not like this movie. Me being the opinion holder.
* Opinion Object: Object on which opinion has been passed on, for example, a movie is an object in above-mentioned sentence.
* Opinion Orientation: Degree of the comment, i.e. positive, negative or neutral. For example, I **do not like** this movie.

Furthermore, opinions can be direct i.e. this movie had great story or they can be indirect like Brad Pit became famous after Troy.

(Gamal et al., 2018) did a comparative analysis of machine learning algorithms for sentiment analysis, using three different feature extraction algorithms and after that using multiple machine learning algorithms like Logistic Regression, Bernoulli NB, Multinomial NB, Naïve Bayes, Maximum Entropy, Passive Aggressive, Stochastic Gradient Descent, Support Vector Machines, Ridge Regression and Adaptive Boosting. It was concluded that among all above mention algorithms the one with best performance was Passive Aggressive for all utilized datasets like IMDB, Cornell Movies, Twitter and Amazon with unigrams. It had values from 87% to 99.96% for evaluation matrices.

(Vo et al., 2018) developed a method of automatically extracting a knowledge base system, which further can be utilized in extraction of aspects for any product and resulting in opinion. They have tried to solve the problem of feature extraction precisely from a bigger and larger corpus. Natural Language Processing gears like named entity recognition, Dependencies Parser and Co-reference Resolution were used in two steps: Knowledge introduction and broad syntactic knowledge extraction.

(Mtetwa et al., 2018) worked on web based API with JSON doing sentiment analysis of movie reviews to show results on any operating system and furthermore, system can be trained for different other domains as well.

(Mäntylä et al., 2018) did a review paper that was computer-assisted literature review. They used mining and qualitative coding to analyze around 6996 papers from Scopus. It was found that sentiment analysis was performed in the start of 20th century but it really started booming after 2004 with availability of subjective text on web. Out of all the papers reviewed, 99% were after 2004. It was also observed in the review that sentiment analysis has gone past just product reviews, many other domains like disasters, politics, stocks, business medicine and software engineering are also being touched.

(Alsaeedi & Khan, 2019 Sentiment Analysis is actually a way of checking the degree of written and spoken languages, it also shows extent i.e. it is positive, negative or neutral. Great number of customer feedback is being gathered through different social media platforms in the form of surveys and reviews on different products. Sentiment analysis than helps in assessing these and currently number of very good tools are available in market for such purpose.

(Badaro et al., 2019) To work better with computerized feelings of any content, notion examination is also an advanced area of research. Sentiment analysis can be widely divided into two categories namely; lexicon based approach and supervised learning approach. First is where sentiment lexica is used in algorithms for prediction of sentiment whereas, supervised methods focuses on more complex relationship between text and sentiment, algorithms from this approach work on labeled examples.

(Ruseti et al., 2020) Proposed a tested solution of an accessible and easily extendable model for creating a customized sentiment analysis for a precise domain using Multinomial Naïve Bayes, Deep Neural Network and SVM. Testing was done 200,000 game reviews and had more adequate results than previous works for game reviews.

# **Methodology:**

The methodology that we are following is summarized below:

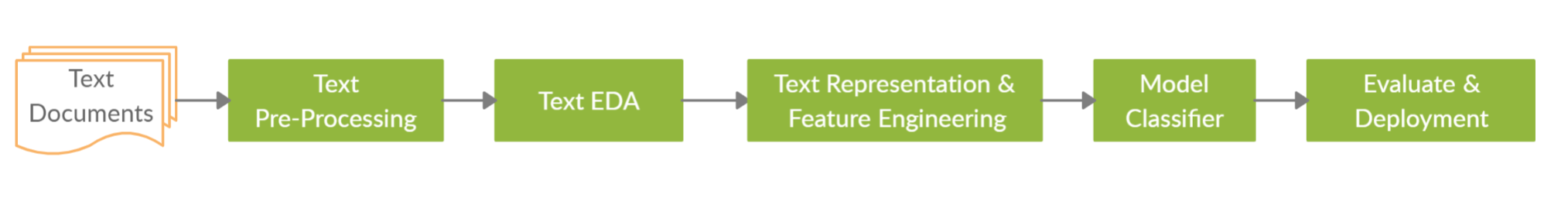


Figure : Process Pipeline

## **Experimental Work:**

The platform for experimental work was Jupyter Notebook, a python IDE. Python version 3.7 is used along with major libraries like numpy, pandas, seaborn, matplotlib and nltk.

### **Data Set:**

First step is to gather data, for our proposed methodology, we are going to use Rotten Tomatoes large movie review dataset of 80,000 reviews. This dataset is mainly, used for binary sentiment classification. The dataset consists of two columns. First column contains movie reviews. The second column is our target column that we have to predict, it comprises sentiments of relevant movie reviews. The sentiment is further divided into 5 categories i.e. positive to negative reviews. Neutral reviews are also included in this data set.

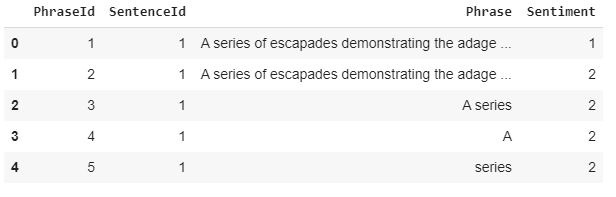


Figure 2: Rotten Tomatoes Dataset

Below figure shows total reviews as per categories created.

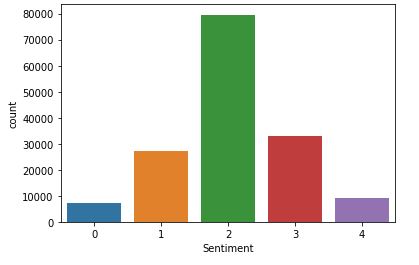


Figure 3: Sentiments Bar chart

### **Data Preprocessing:**

After acquiring large Rotten Tomatoes movie review dataset, Data Pre-processing helps to clean the text and improves accuracy as well as helps in better prediction. The steps that are involved in our data preprocessing are:

* Eliminating HTML tags in reviews like ‘<b>’
* Eliminating special characters example: ‘@’ and punctuation marks example ‘!’ using regular expressions.
* Converting all the text from Upper case into lower case.
* Eliminating the stop words like ‘and’, ‘or’ ‘the’ ‘but’ because they do not provide any opinion in a sentence.
* In order to bring words in their root form we have to perform Stemming/ Lemmatization for example, the word ‘running’ or ‘runs’ after stemming and lemmatization its root form is ‘run’.
* Perform Feature Extraction or Vectorization to convert text into numerical form.

### **Training the Model:**

For our proposed methodology, we have used three main machine-learning algorithms to train our data i.e. Logistic Regression, Decision Tree and Naïve Bayes classifier. After cleaning the data, we split our movie dataset into 25% testing and 75% training data. Since we cannot use text directly, we need to convert it into numeric form for our training model. This can be achieved by converting our text into vectors. This process is called vectorization or feature extraction. One of the most effective method is Bag of Word Model. We are going to perform two kinds of feature extraction or vectorization. First is Count Vectorization and second TF-IDF.

**A. Naïve Bayes Classifier:**

This is the most used and common classification classifier. A probabilistic classifier that works on the principle of Naïve Bayes theorem. It is a conditional probability model, which works of the assumption that predictors are independent, meaning that one specific feature does not affect the other.

**B. Logistic Regression:**

Logistic regression is based on statistical model and it uses logistic function to model binary variable. Logistic regression can be described as linear algorithm but prediction are converted using logistic function.

**C. Decision Tree:**

Decision tree algorithm is based on supervised learning model where predictions are done using tree representation where every node is an attribute and every leaf node represents class. It can be simply described as logical graphical representation of solutions to a problem.

**D. TF-IDF:**

Term Frequency – Inverse Document Frequency. Term Frequency refers to how frequently the word appears in a document whereas; Inverse Document Frequency tells us if the word is common or infrequent across the document.

**E. Performance Metrics:**

* *Precision* of the classifier is defined as the fraction of correctly predicted positive values to total predicted positive values. It tells how precise our model in terms of those predicted positive.
* *Recall* of the classifier is defined as the percentage of errors correctly predicted out of all the errors that actually occurred. Recall is mainly used when there is high cost asscosiated with False Negative.
* *Accuracy* of a classifier was defined as the fraction of correctly predicted values to the total observations. Accuracy is one of the best measure to evaluate a classifier. The higher th accuracy the better the model.
* *F1-score* is optimal if we need both precision and recall. It takes both measures and develops a balance between.
* *ROC Curve* to show cost of binary classifiers by plotting true positive rate against false positive rate

# **Results and Discussion:**

Before data preprocessing, let us analyze the distribution of review length in data set.

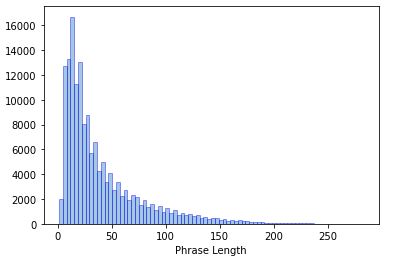


Figure 4: Reviews Length

Figure 4: shows that most reviews are under length 0 to 50 meaning most of the opinions can be extracted from shorter reviews.

After that, we compared sentiments based on phrase length to see the trend of reviews and figure 5 shows that most people who reviewed somewhat positive or somewhat negative are of greater length than other categories. It shows that absolute positive and negative reviews are short and to the point.

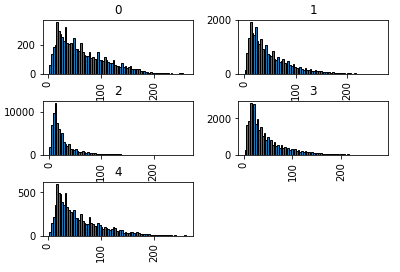


Figure 5: Phrase lengths as per Label

After phrase length analysis boxplot is generated, boxplot is a standard way of displaying the data distribution. It also tells you how your data is grouped furthermore it shows if your data is symmetrical or not. We generated boxplots of sentiments and phraseId.

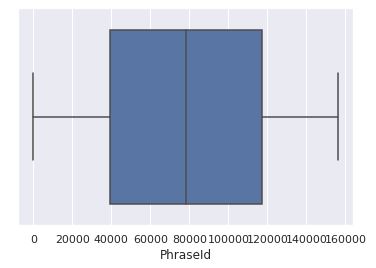
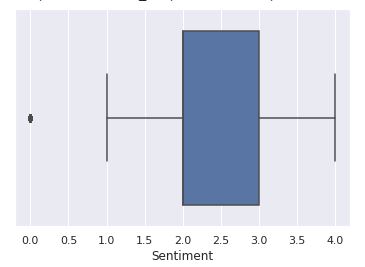


Figure 6: Phrase lengths as per Label

Figure 6 of boxplots shows that for this dataset, we shall let sentiment values above 2 represent positive ones. As a result, positive movie reviews make up less than 50% of the dataset.

After above analysis, preprocessing is done and data sit is split in training and testing, all features are extracted using tf-idf and model is trained using all three algorithms.

ROC curves are generated to measure the performance of our models, below figures show results of ROC Curves.

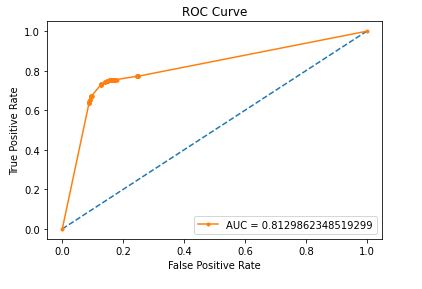


Figure 7-: ROC Curve for Decision Tree

As you can see in the above figure ROC curve of Decision Tree, the AUC curve value is 0.812 that is not very impressive compared to our other two models.

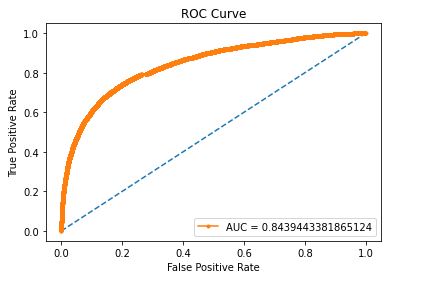


Figure 8: ROC Curve for Naïve Bayes

According to our results, Naïve Bayes performs better than Decision Tree and its AUC curve is 0.8543.

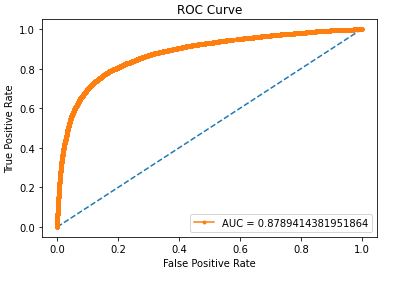


Figure 9: ROC Curve for Logistic Regression

Above figures show that AUC (Area Under the Curve). Most outstanding results are shown by logistic regression as AUC of Logistic Regression Model is greater than other two, meaning that logistic regression has shown best performance so far, we’ll move to confusion matrix and accuracy score before and after threshold to see if our results are according to the curve or not.

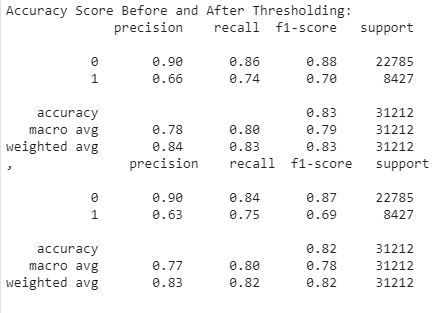
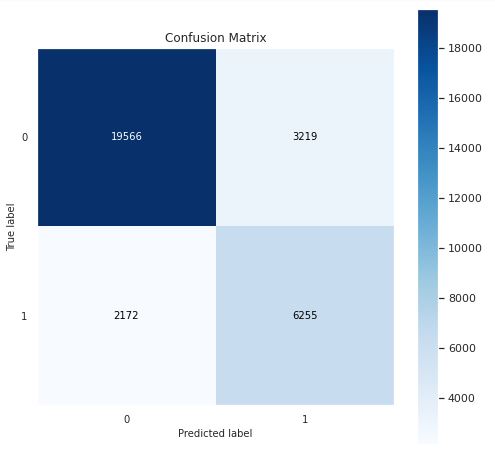
**Results from Decision Tree:**

Figure 10: Results from Decsion Tree

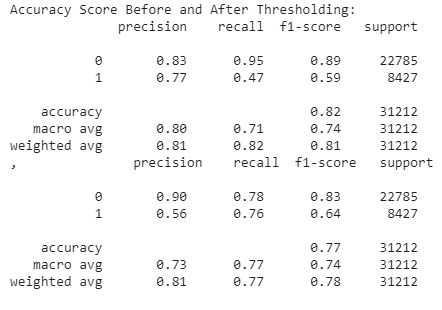
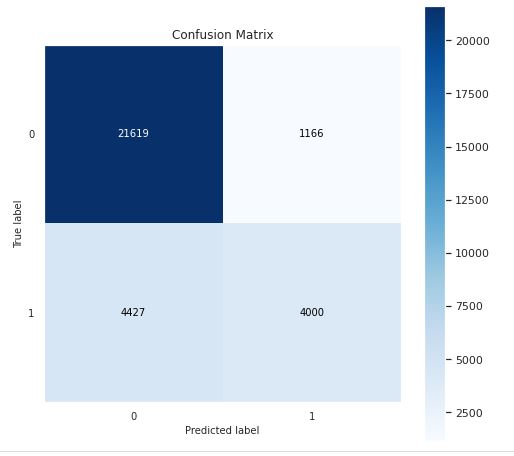
**Results from Naïve Bayes:**

Figure 11: Results from Naïve Bayes

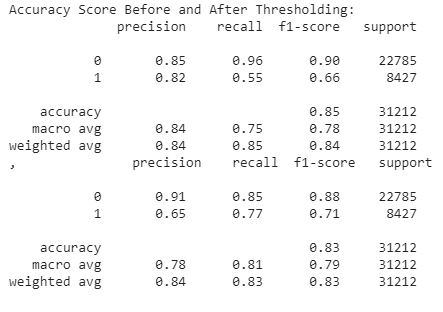
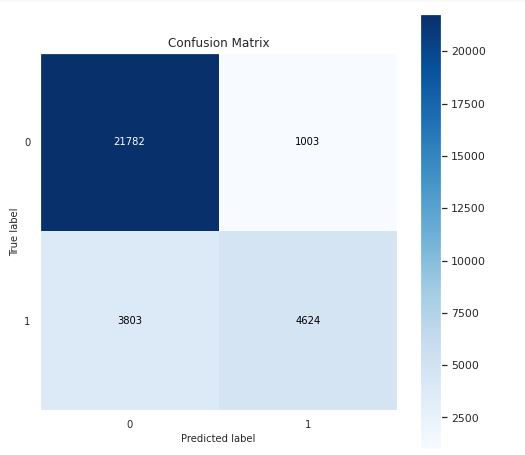
**Results from Logistic Regression:**

Figure 12: Results from Logistic Regression

We can summarize our results by saying that accuracy, confusion matrix as well as the ROC curve for all three models illustrate that Naive Bayes has an accuracy of 77%, Decision Tree has an accuracy of 82% and Logistic Regression has outperformed the other two algorithms before and after threshold with an accuracy of 83%.

**Comparison with Related Work:**

|  |  |  |  |
| --- | --- | --- | --- |
| **SR#** | **Table. 1-Related Work** | | |
| **Study** | **Methodology** | **Accuracy** |
| 1 | Our proposed Methodology | Naïve Bayes, Logistic Regression and Decision Tree | NB=0.77, LR=0.83, DT=0.82 |
| 2 | (V.K. Singh, R. Piryani, A. Uddin, 2013) | (AAC) and Adverb Adjective and Adverb Verb combination (AAAVC) | 78.7 |
| 3 | (Singh & Shahid Husain, 2014) | SVM | Accuracy of 81.15% for movie review dataset  79.40% on product data set |
| 4 | (Indhuja & Reghu Raj, 2014) | Fuzzy Logic | 87.58% |
| 5 | (Chinsha & Joseph, 2015) | Combination of adverb verb and adverb adjective | 78.04% |
| 6 | (Sahu & Ahuja, 2016) | Naïve Bayes, KNN, Bagging, COCR, Random Forrest and Decision Tree. | Random Forrest scored the highest accuracy of 88.95% |
| 7 | (Xia et al., 2016) | ensemble classifier with SVM, NB and LR as base learners | Majority Voting Ensemble with IG = 93.94 |
| 8 | (Kai et al., 2017) | Hybrid Model (SVM + GBDT) | P=0.81, R=0.92, f1=o.88 |
| 9 | (Zvarevashe & Olugbara, 2018) | Naïve Bayes multinomial, sequential minimal optimization, compliment Naïve Bayes and Composite hypercubes | Precision of NBM=80.9 |
| 10 | (Gamal et al., 2018) | Naïve Bayes (NB), Stochastic Gradient Descent (SGD), Support Vector Machines (SVM), Passive Aggressive (PA), Maximum Entropy (ME), Adaptive Boosting (AdaBoost), Multinomial NB (MNB), Bernoulli NB (BNB), Ridge Regression (RR) and Logistic Regression (LR) with TFIDF | NB=65.56, BNB=84.86, MNB=85.46, AdaBoost=80.46, ME=64.36, SGD=85.56, SVM=84.56, LR=86.56, PA=99.96, RR=98.6 |
| 11 | (Ruseti et al., 2020) | SVM, MNB, DNN | 66, 65, 67 |

According to our methodology our result of Naïve Bayes model is not good compared to (Zvarevashe & Olugbara, 2018) work but our accuracy of Naïve Bayes model is better compared to the authors (Gamal et al., 2018) who did a comparative analysis using various machine learning models. Our highest accuracy is achieved using Logistic Regression which is 0.83 and compared to the works of (Gamal et al., 2018) and (Xia et al., 2016) it is less. Therefore, we can say that our models still have some room for improvement.

# **Conclusion:**

In the end, we would conclude our research paper as, that opinion mining or sentiment analysis is getting vital part of many industries day by day, it is helping to understand current trend and allowing companies to fulfil customer’s demands in a better way along with flourishing their own business. Furthermore, we can safely say that Logistic Regression has shown good performance in opinion mining and in future, we can simply take Logistic Regression to compare with any other algorithm’s performance for further studies to make sentiment analysis more accurate. All other analysis before preprocessing show the credibility of data and of course, ROC curve tells us how credible is our algorithm that can ultimately save lots of time and effort.

# **References:**

Agarwal, B., & Mittal, N. (2013). *LNCS 7817 - Optimal Feature Selection for Sentiment Analysis*. 13–24.

Alsaeedi, A., & Khan, M. Z. (2019). A study on sentiment analysis techniques of Twitter data. *International Journal of Advanced Computer Science and Applications*, *10*(2), 361–374. https://doi.org/10.14569/ijacsa.2019.0100248

Badaro, G., Baly, R., Hajj, H., El-Hajj, W., Shaban, K. B., Habash, N., Al-Sallab, A., & Hamdi, A. (2019). A survey of opinion mining in Arabic: A comprehensive system perspective covering challenges and advances in tools, resources, models, applications, and visualizations. *ACM Transactions on Asian and Low-Resource Language Information Processing*, *18*(3). https://doi.org/10.1145/3295662

Blitzer, J., Dredze, M., & Pereira, F. (2007). ACL07 Biographies, Bollywood, Boom-boxes and Blenders. *Association for Computational Linguistics - ACL 2007*, *June*, 440–447.

Chinsha, T. C., & Joseph, S. (2015). A syntactic approach for aspect based opinion mining. *Proceedings of the 2015 IEEE 9th International Conference on Semantic Computing, IEEE ICSC 2015*, 24–31. https://doi.org/10.1109/ICOSC.2015.7050774

Gamal, D., Alfonse, M., M. El-Horbaty, E.-S., & M. Salem, A.-B. (2018). Analysis of Machine Learning Algorithms for Opinion Mining in Different Domains. *Machine Learning and Knowledge Extraction*, *1*(1), 224–234. https://doi.org/10.3390/make1010014

Indhuja, K., & Reghu Raj, P. C. (2014). Fuzzy logic based sentiment analysis of product review documents. *2014 1st International Conference on Computational Systems and Communications, ICCSC 2014*, *December*, 18–22. https://doi.org/10.1109/COMPSC.2014.7032613

Kai, Y., Cai, Y., Dongping, H., Li, J., Zhou, Z., & Lei, X. (2017). An effective hybrid model for opinion mining and sentiment analysis. *2017 IEEE International Conference on Big Data and Smart Computing, BigComp 2017*, 465–466. https://doi.org/10.1109/BIGCOMP.2017.7881759

Kaur, A., & Gupta, V. (2013). A survey on sentiment analysis and opinion mining techniques. *Journal of Emerging Technologies in Web Intelligence*, *5*(4), 367–371. https://doi.org/10.4304/jetwi.5.4.367-371

M, M., & Mehla, S. (2019). Sentiment Analysis of Movie Reviews using Machine Learning Classifiers. *International Journal of Computer Applications*, *182*(50), 25–28. https://doi.org/10.5120/ijca2019918756

Maas, A. L., Daly, R. E., Pham, P. T., Huang, D., Ng, A. Y., & Potts, C. (2011). Learning Word Vectors for Sentiment Analysis. *Proceedings of the 49th Annual Meeting of the Association for Computational Linguistics: Human Language Technologies*, 142–150. http://www.aclweb.org/anthology/P11-1015

Mäntylä, M. V, Graziotin, D., & Kuutila, M. (2018). The Evolution of Sentiment Analysis. *Computer Review*, *27*(February), 16–32. https://doi.org/10.1016/j.cosrev.2017.10.002

Moraes, R., Valiati, J. F., & Gavião Neto, W. P. (2013). Document-level sentiment classification: An empirical comparison between SVM and ANN. *Expert Systems with Applications*, *40*(2), 621–633. https://doi.org/10.1016/j.eswa.2012.07.059

Mtetwa, N., Awukam, A. O., & Yousefi, M. (2018). Feature Extraction and Classification of Movie Reviews. *5th International Conference on Soft Computing and Machine Intelligence, ISCMI 2018*, 67–71. https://doi.org/10.1109/ISCMI.2018.8703235

Pang, B., & Lee, L. (2004). *A Sentimental Education: Sentiment Analysis Using Subjectivity Summarization Based on Minimum Cuts*. http://arxiv.org/abs/cs/0409058

Penubaka Balaji, D. H. and O. N. (2018). 6. An Overview on Opinion Mining Techniques and Sentiment Analysis. *International Journal of Pure and Applied Maths*, *India*(19), 61–69.

Ruseti, S., Sirbu, M. D., Calin, M. A., Dascalu, M., Trausan-Matu, S., & Militaru, G. (2020). Comprehensive exploration of game reviews extraction and opinion mining using nlp techniques. *Advances in Intelligent Systems and Computing*, *1041*(January), 323–331. https://doi.org/10.1007/978-981-15-0637-6\_27

Saberi, B., & Saad, S. (2017). Sentiment analysis or opinion mining: A review. *International Journal on Advanced Science, Engineering and Information Technology*, *7*(5), 1660–1666. https://doi.org/10.18517/ijaseit.7.5.2137

Sahu, T. P., & Ahuja, S. (2016). Sentiment analysis of movie reviews: A study on feature selection and classification algorithms. *International Conference on Microelectronics, Computing and Communication, MicroCom 2016*. https://doi.org/10.1109/MicroCom.2016.7522583

Sharma, R., Nigam, S., & Jain, R. (2014). Opinion Mining of Movie Reviews At Document Level. *International Journal on Information Theory*, *3*(3), 13–21. https://doi.org/10.5121/ijit.2014.3302

Singh, P. K., & Shahid Husain, M. (2014). Methodological Study Of Opinion Mining And Sentiment Analysis Techniques. *International Journal on Soft Computing*, *5*(1), 11–21. https://doi.org/10.5121/ijsc.2014.5102

Tang, D., Qin, B., & Liu, T. (2015). Learning semantic representations of users and products for document level sentiment classification. *ACL-IJCNLP 2015 - 53rd Annual Meeting of the Association for Computational Linguistics and the 7th International Joint Conference on Natural Language Processing of the Asian Federation of Natural Language Processing, Proceedings of the Conference*, *1*, 1014–1023. https://doi.org/10.3115/v1/p15-1098

V.K. Singh, R. Piryani, A. Uddin, P. W. (2013). Sentiment Analysis of Movie Reviews A new Feature-based Heuristic for Aspect-level Sentiment Classification. *2013 International Mutli-Conference on Automation, Computing, Communication, Control and Compressed Sensing (IMac4s)*. https://doi.org/https://doi.org/10.1109/iMac4s.2013.6526500

Vo, A. D., Nguyen, Q. P., & Ock, C. Y. (2018). Automatic knowledge extraction for aspect-based sentiment analysis of customer reviews. *ACM International Conference Proceeding Series*, 110–113. https://doi.org/10.1145/3177457.3177462

Xia, R., Xu, F., Yu, J., Qi, Y., & Cambria, E. (2016). Polarity shift detection, elimination and ensemble: A three-stage model for document-level sentiment analysis. *Information Processing and Management*, *52*(1), 36–45. https://doi.org/10.1016/j.ipm.2015.04.003

Zvarevashe, K., & Olugbara, O. O. (2018). A framework for sentiment analysis with opinion mining of hotel reviews. *2018 Conference on Information Communications Technology and Society, ICTAS 2018 - Proceedings*, 1–4. https://doi.org/10.1109/ICTAS.2018.8368746