A Movie Review Sentiment Analysis using Machine Learning Techniques

[1]P. Lakshmi Surekha [2] A. Jayanthi

[1](M.Tech)Student, Department of CSE, G.Pulla Reddy Engineering College, Autonomous, Kurnool, Andhra Pradesh, India
[2]M.Tech, CSE, Kurnool, Andhra Pradesh, India

Abstract

A sentiment analysis is a process where the sentiment describes is a movement behind meaning and something like irritated, happy, optimistic, harmful etc. This paper explains about the sentiment behind movie reviews. One of the most recent trend applications of Artificial Intelligence is also a Machine Learning (ML), in which software, computers and peripheral devices perform via cognition (very similar to human brain). AI part is important role in sentiment analysis talk about movie reviews like Negative, Positive, somewhat negative, somewhat positive, positive. The obstacles observe in like Language Ambiguity, Sarcasm, sentence Negation, Terseness etc., Sentiment analysis is categorization of sentiment division and its a part of written/printed text, the main job of the sentiment analysis is to determine the spoken words or opinion of person in document format or text format, like ex. Positive, negative. It is useful in social network monitoring and it allows us to get benefit an overview of maximum number of public opinion about certain topics. This sentiment analysis describes about opinion mining from client with Emotional feelings like happy or unhappy. So in this paper we represent Sentiment analysis using Machine Learning based approach for movie review. In this proposed system classification of sentiment analysis of present movie rating accuracy results. The existing movie ratings across in internet services it has a base to future research and implementation of domain. Then the execution of Feature Selection & Classification Algorithms are included in Movie review sentiment analysis applications using Machine Learning.

Keywords

Movie review, Feature Selection Alg., Classification Alg Sentiment study, Data Mining, Opinion, Classifier.

I. Introduction

In real life in sentiment analysis of ML based approach as shown in fig.1 as part we use some words like " I like to travel" is indicates positive sentiment and when I say " I like to travel/ I love nature, sometimes." Is indicates not a positive or nagative means it neutral[1]. So this sentiment analysis can be expressed as Sentiment Analysis = judgment Mining = Emotional AI (indicates happy or unhappy)

In present generation of internet utilization become a major database maintains a gigabytes of data which is utilized by the users[2]. This huge database is continuously growing stage at exponential rate. The users of social networks like Facebook, Youtube, Twitter etc. we review opinions of social networks maintained in document format. Different opinions of clients or users, collection of likes and dislikes observe in movie review. Present generation time valuable than money and dead line for a task completion also very important. To save time we can go through online shopping now a days. In this moment to purchase any type of product go through particular item and we can do observation and features. Then parallelly we can observe comments on that particular product and the

Manuscript received August 5, 2025
Manuscript revised August 20, 2025
https://doi.org/10.22937/IJCSNS.2025.25.8.3

same goods purchased by the customer and their opinions and comments.

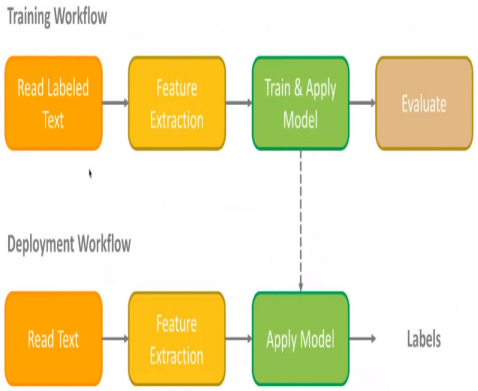


Fig:1. The block diagram of machine learning based approach

II. Sentiment Analysis of Movie Review

In this analysis how many peoples are like your product and how many peoples not like your product all these information getting from internet service[3]. There are so many use cases for Movies and Social media network monitoring, Political Issues, Business Issues, Home Applications and Customer Experience etc. some cases as follows:

1. In Movies or Pictures as shown in fig. 2 observe review either Positive or Negative.
2. Utilization of Social network applications in which we can tweet either it is positive or Negative
3. Discuss about the Political Issues what people thinking about a specific person or Political issues.
4. At last we got customer experience what peoples are think about the product.
5. We got their feedback either positive or negative

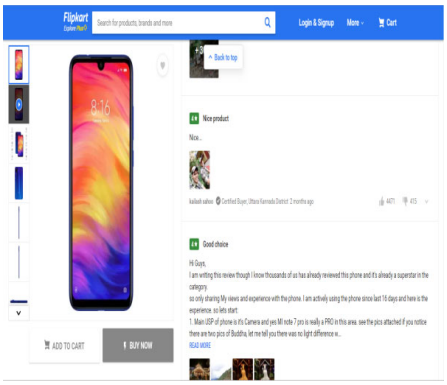


Fig:2. sentiment analysis of product

As shown In above figure feed back information about product like good choice. It means good experience on the product, some persons doesnot satisfy about the product they give negative feedback about the product means it idicates bad experience about the product.

2.1.Sentiment Analysis

In ML technique Sentiment analysis is a real-time application. It indicates opinion/information retreiving, and sentiment classification, etc. it is used to determining the opinion of the client/writer[4]. In other meaning of this context, this process of finding out the emotion from the text. The main goal of sentiment analysis is " what other people think?". Assume that some peoples watch and by their observation write 'the movie is bad or irritated.' To find out the actual opinion from the data (is it good/ bad) is the task of sentiment analysis. This SA(sentiment analysis) application to apply for such as in review based website, decision-making application.

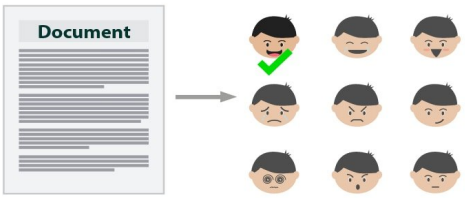


Fig:3. Opinion of client

Fig.3 indicates (ML) machine learning application is to constructs a system by extracting the knowledge from information. In machine learning techniques, learning algorithms are supervised and unsupervised approaches are used for SA.

Image identification system is also used in ML techniques. Here in this process many situations we classify the object as a digital image[5]. ML is used for face recognition system .ML is also used for Text recognition to determine handwritten or printed letters. We can devide information writing into smaller images, each containing a single character.

III. Design and Implementation of sentiment analysis of work in Machine

Forward selection is a method of iterative process in which we started with no features in the model. In step of iteration, adding the feature for best enhancements and improves our model and best performance of the model. Backward Elimination, technique start with all the best features and eliminate the least significant feature and it repeat this until no improvement is observed.

Fig:5. subset of features and train a model

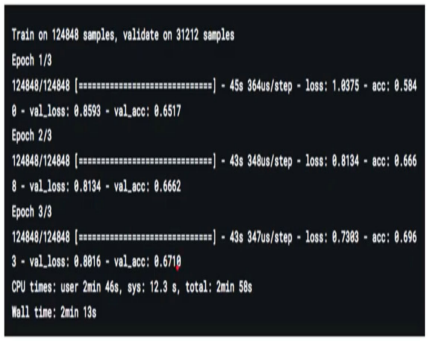


Fig:6. Python program for sentiment analysis with validation samples

As shown in fig.6 above code it explains about vectorizer range between 1 and 2, the full text in phrase values and it tokenizer then transform in test phrase and it train in sentiment. Python program explains about Logistic Regression, TFIDF, ngrams for sentiment analysis[7].

The information is always retrieves from search engines which are available in internet service. Information extraction means question and answering. As shown in below figure the text of mining workflow in fig. 7 using Machine learning. First part we take documentation of our opinion and add some stickers, part of speeches included in enrichment, in

preprocessing we remove odd words, in last end of block indicates classification of sentiment analysis. The sentiment score when it is greater than zero is positive sentiment when it less than zero it indicates negative sentiment. Therefore sentiment score = positive words – negative words total words

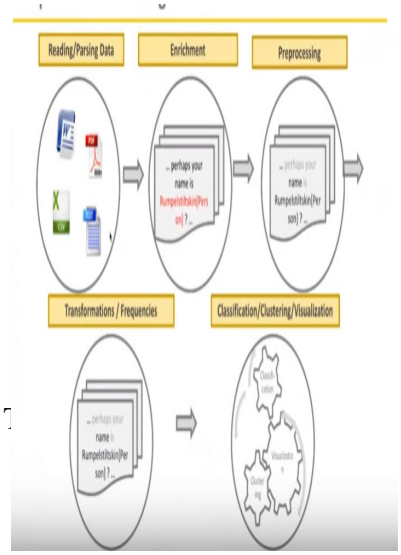
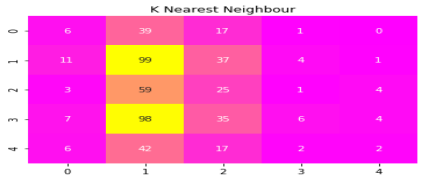
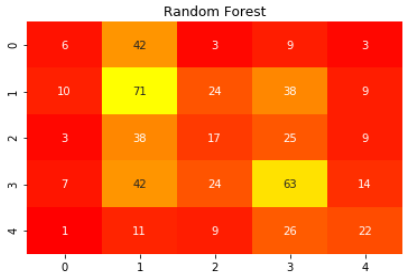


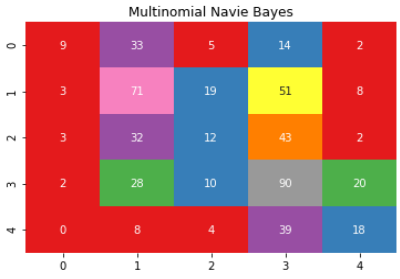
Fig:7. text of mining workflow using Machine learning

therefore sentiment score in text is indicates in text how many helpful words and negative words in document[8]. By preprocessing & transform we delete all unnecessary which I didn't take.

V. Results and Discussion







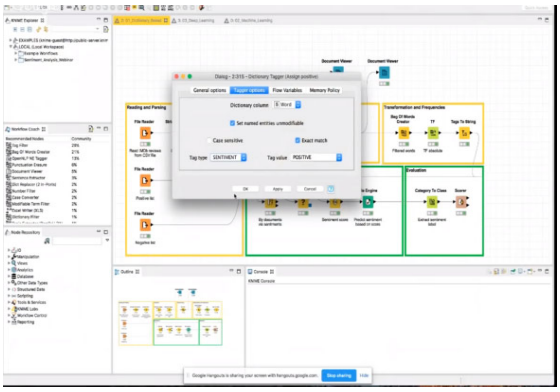

```
--------------------Support Vector Machine--------------------
              precision    recall  f1-score   support

           0       0.00      0.00      0.00        63
           1       0.00      0.00      0.00       152
           2       0.00      0.00      0.00        92
           3       0.29      1.00      0.44       150
           4       0.00      0.00      0.00        69

avg / total       0.08      0.29      0.13       526

--------------------Descion Tree--------------------
              precision    recall  f1-score   support

           0       0.23      0.13      0.16        63
           1       0.34      0.43      0.38       152
           2       0.11      0.08      0.09        92
           3       0.32      0.33      0.32       150
           4       0.21      0.23      0.22        69

avg / total       0.26      0.28      0.27       526
```

==========Random Forest==========

	precision	recall	f1-score	support
0	0.22	0.10	0.13	63
1	0.35	0.47	0.40	152
2	0.22	0.18	0.20	92
3	0.39	0.42	0.41	150
4	0.39	0.32	0.35	69
avg / total	0.33	0.34	0.33	526

==========Multinomial NB==========

	precision	recall	f1-score	support
0	0.53	0.14	0.22	63
1	0.41	0.47	0.44	152
2	0.24	0.13	0.17	92
3	0.38	0.60	0.47	150
4	0.36	0.26	0.30	69
avg / total	0.38	0.38	0.36	526



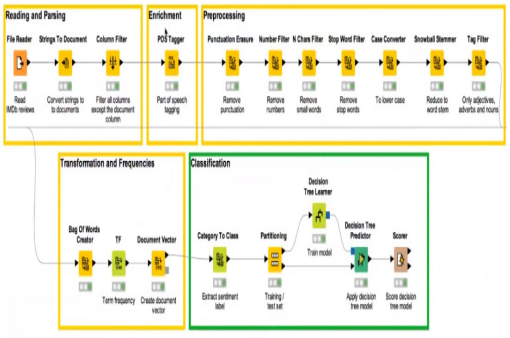


Fig:8. Machine Learning based approach process

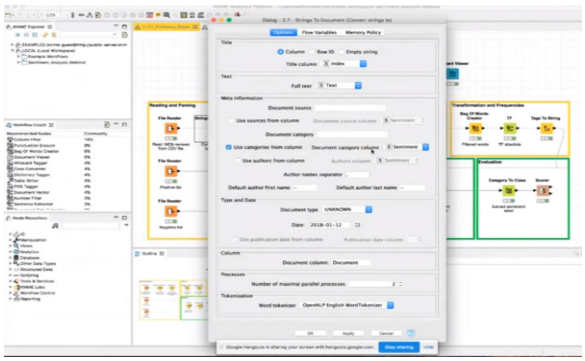


Fig:9. Positive text

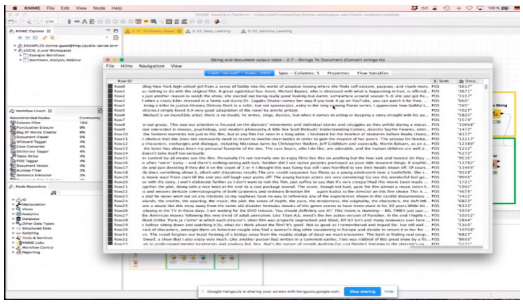
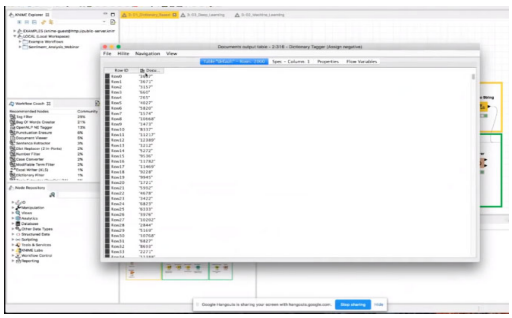


Fig:10. Output table



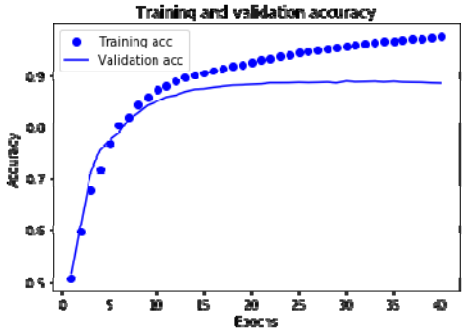


Fig:11. Text classification with movie reviews

The typical tasks which are included in it i.e. text categorization means document classification, text summarization, abstract of text, text clustering, Opinion mining with entity or concept extraction. For example before we purchase any products mobile, AC or any electronic goods from social network websites like Flipkart, Amazon, Alibaba from these observe reviews and comments or feedback about the product[10]. In this communication process we observe maximum number of persons positive feedback and negative number of persons feedback. We can take decision on which we got maximum number of positive feedbacks. Fig.8,9,10,11,12 all indicates the result analysis of movie review using ML process. Then we move on that particular item. In this way we got confident level on usage of the product by different persons. In this way we got sentiment analysis of different product which is available in society[9]. As shown in fig.12 sentiment analysis shows about movie reviews like Negative, Positive, somewhat negative, somewhat positive.

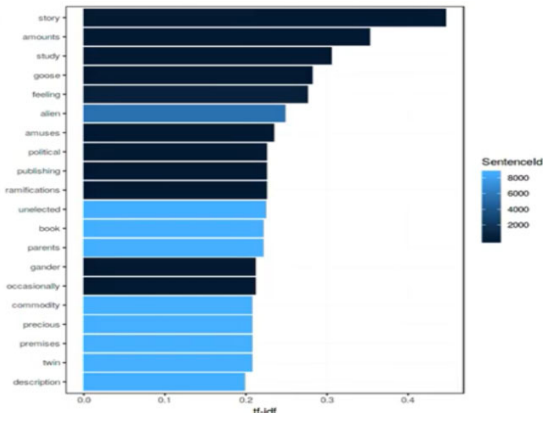


Fig:12. Graph about Sentiment distribution Movie review

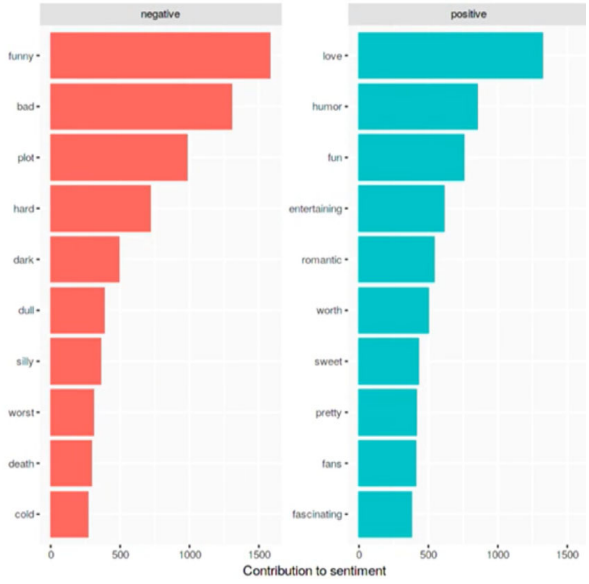


Fig:13. Comment about movie

As shown in fig.13 it indicates positive and negative words about the movie. Negative words are used for movie comments like worst, bad, dark, dull etc. positive words are used for movie comments like love, fun, worth, pretty etc. there is top 20 most common words used in movie review.

VI. CONCLUSION & FUTURE WORK

This paper represents proposed strategies it can develop the performance of selected sets substantially. By using the algorithms of Forward selection is a method of iterative process in which we started with no features in the model. In step of iteration, adding the feature for best enhancements and improves our model and best performance of the model. Backward Elimination, technique start with all the best features and eliminate the

least significant feature and it repeat this until no improvement is observed in movie reviews and determining the polarity of the movie reviews by applying the computational methods. Using this project we can easily analyze the tweets in twitter and we can easily predict the correct review. This project can be directly used for a better movie recommendation experience for the users in platforms like Netflix, Amazon Prime.

VII. REFERENCES

[1] sentiment classifications using machine learning techniques." Proceedings of the conference on Empirical methods in natural language processing by Pang, Bo, and Lillian Lee, Shivakumar-Volume 10, 2002.

[2] Pang, Bo, and Lillian Lee. "A sentimental education: Sentiment analysisusing subjectivity summarization based on minimum cuts." Association for Computational Linguistics, 2004.

[3] Nigel Collier, Mullen and Tony "Sentiment Analysis using Support Vector Machines with Diverse Information Sources." EMNLP. Vol. 4. 2004.

[4] Alina, Andreevskaia & Sabine Bergler. "Mining WordNet for a Fuzzy Sentiment: Sentiment Tag Extraction from WordNet Glosses." EACL. Vol. 6. 2006.

[5] Baccianella, Stefano, Andrea Esuli, and FabrizioSebastiani. "SentiWordNet : An advanced Lexical Resource for Sentiment Analysis and Opinion Mining." 2010.

[6] Patrick Paroubek, Pak & Alexander,. "Twitter as a Corpus for Sentiment Analysis and Opinion Mining." LREC. Vol. 10. 2010.

[7] Annett, Michelle, and GrzegorzKondrak. "A comparison of sentiment analysis techniques: Polarizing movie blogs." Advances in artificial intelligence, 2008.

[8]"Sentiment analysis of combined approach." Journal of Informetrics 2009.

[9] Stefano, Baccianella, Andrea Esuli, and Fabrizio. "SentiWordNet 3.0: An Enhanced Lexical Resource for Sentiment Analysis and Opinion Mining, 2010.

[10] Singh, V. K., et al. "Sentiment analysis of movie reviews: A new feature based heuristic for aspect-level sentiment classification." Automation, Computing, Communication, Control and Compressed Sensing, IEEE, 2013

P.Lakshmi Surekha, pursuing M.Tech in Computer Science and Engineering form work, G.Pulla reddy engineeing college, Kurnool, her research interest includes big data analytics, Machine learning, IoT, Computer Networks, Embedded Systems.

A. Jayanthi, M.Tech in Computer Science and Engineering. Her research interest includes big data analytics, Machine learning, IoT, Computer Networks, Embedded Systems.