

Sentiment Analysis Based on Movie Reviews using Various Classification Techniques : A Review

Karishma Kaushik, Mahesh Parmar

Department of Computer Science and Engineering and Information Technology, Madhav Institute of Technology and Science, Gwalior India

ABSTRACT

Article Info

Volume 7, Issue 3

Page Number: 197-208

Publication Issue :

May-June-2021

Article History

Accepted : 01 May 2021

Published : 05 May 2021

Sentimental analysis is also called "opinion mining" analyses attitudes and classifies text views. It relates to the use of natural language processing, text, and linguistic processing. A huge amount of data is created with the rapid growth of web technologies. Social networking sites are now popular and normal places where feelings can be shared by short messages. These sentiments involve happiness, sadness, anxiety, fear, etc. The analysis of short texts tends to recognize the crowd's sentiment. Sentiment Analysis on IMDb moviereviews describes a reviewer's general feeling or impression of a movie. Since the perceptions of humans improve the effectiveness of products & since a movie's success or failure depending on its review, costs are rising, and a good sentiment analysis model needs to be developed, that classifies moviereviews. Machine learning methods use ML algorithms to carry out sentiment analysis as a standard classification problem using syntactic and language characteristics. There are some methods of machine learning used for sentiment analysis in this paper. Most of the sentiment analysis is performed using SVM, RF, ANN, and NB, Algorithms of DT, BN, & KNN.

Keywords : Sentiment Analysis, Movie Review, Text Processing, Machine Learning, Feature Extraction, Classification Techniques.

I. INTRODUCTION

Today, the internet era has transformed people's way of expressing thoughts. This is achieved more by blog posts, forums, blogs for product testing, social media, and so forth. Today, millions of people have been using social networking platforms like Facebook, Twitter, Google Plus, etc. to communicate their emotions. We get immersive media through online communities that inform and influence other people through forums. Social media generates a vast

number of sentimentally rich information in the form of tweets, updates, blog posts, remarks, feedback, etc. Also, social media offers companies the ability to communicate with their advertisement clients. Most people rely mainly on the content created by users online. For example, if someone wishes to purchase or use a product, they first lookup their review online, then address it on social media before making a decision. The amount of user information is too large to be analyzed by a regular user. Therefore, this must be automated, different methods for feeling

processing are commonly employed[1]. Social media sites were recently used to exchange views and perspectives on everyday life, business, learning, celebrities, etc.

The sentimental analysis is often regarded as "opinion mining" and also applies to the application of natural language processing (NLP), text mining, electronic linguistics, & bio-metallization, & tools for systematic assessment, review, and study. SA referred to as sentiment classification, sentiment mining, examination mining, subjectivity analysis, opinion mining, or evaluation extraction and polarity classification deals with the computational processing and subjectivity of the document. The textual analysis aims to define a speaker's decision on a trustworthy issue. The role might represent assessment, evaluation or decision, feeling, or the expected disturbing communication. It is also an indication that "subjective" does not mean something is not real in this situation [2]. Analysis of sentiments may be characterized by text, NLP & language analysis as a method for sensing or identified sentiments. Only a training data set and a testing dataset are applied for developing a classifier in the machine-learning methodology. It is easier than the basic method of knowledge.

ML is a method used to show computers how to work with data more easily and to produce a better performance. We can't grasp the model or derive details from the data in certain situations after presenting the dataset. We use machine learning techniques in such cases to predict the results. A large number of datasets can be obtained from numerous sources, and machine learning is required. Machine learning is used in numerous fields, from medical to military, to derive useful knowledge from accessible data sets. Machine learning is specifically meant to learn from current records. A wide variety of algorithms is programmed to learn from computers.

Often mathematicians and programmers use a variety of methods to solve this problem.. [3].

Classification is a DM feature that assigns categories or groups in a list. The classification aims to forecast the target class correctly for each data record. Classification between objects is a simple task for humans, but for machines, it has proven difficult. The increased availability of high-capability processors, high quality and low-cost video cameras as well as the increased need for automated video analyses have created interest in algorithms for object classification. A basic scheme of classification consists of a camera that is far above the area concerned, where photographs are collected and processed. The classification covers image sensors, pre-processing of images, object recognition, object segmentation, extraction of objects. Classification scheme comprises a database of predefined patterns which are classified into the intended form compared to the observed object. Image classification in diverse applications, like biomedical imaging, biometry, video monitoring, vehicle sailing, industrial visual control, robot navigation, and remote sensing, is a critical and challenging activity [4].

The remaining part of the paper is structured according to Section 2 gives a brief review of Sentiment analysis and its Levels, Applications and challenges, etc. Section 3 provides movie reviews using the IMDb dataset. And Section 4 gives a small introduction of Machine learning and its various algorithms and Section 5 describes various classification techniques, and after this, in Section 6 we present the related study in Sentiment analysis using various classification techniques and IMDb dataset, and finally, section 7 concludes the review paper.

II. LITERATURE REVIEW

M. Butler and S. Robila(2016)This study outlines the creation and deployment of the IMDb dataset file extraction and import to database method. The prior study used relational databases, which vary from most published methods or analyses. This method uses document-driven data structures and enables anyone to add structures to modify based on their needs. The project creation involved the use of presently needed technology for site developers & software engineers, helping other developers to fork variants of the work & use that for their specific research. [20].

A. Yenter & A. Verma (2017) are experimenting with a wide variety of regularization methods, network architectures, and kernel sizes for 5 high-performance models. These models can forecast the polarity in sentiments of IMDb dataset feedback with precision above 89%. Initially, the precision of the model of best performance goes beyond the previous ones, and secondly, it significantly enhances the baseline CNN+LSTM model. Simultaneous kernel functionality from many divisions of LSTM-based CNN architecture can also prove beneficial to other sentimental analytics or text classification datasets. The suggested framework also has the potential in visual / voice machine learning [21].

Dipak R. Kawade et al. (2017)Current research retrieves Uri attack tweets & finds tweets emotions & polarity. Text mining methods are also used to mine emotions including polarity in tweets. Approximately 2 million tweets are pre-coded to construct a dataset of words that are most viewed. R is used for polarity and mining sentiments. Test results reveal that the Uri attack disgusted 94.3 percent of the population [22].

A. U. Hassan et al. (2017)In this report they will explain the deep depression level of an individual by

analyzing and gathering emotional data on a range of social media channels using concepts of sentiment, machine learning & NLP techniques..[23].

S. Mathapati et al. (2018)In this study they identify feedback of deep learning technology to boost predictions using these functions. They work together to achieve higher accuracy and a shorter computation time. CNN's and LSTM- RNN. [24].

M. R. Haque et al. (2019)This study comparing CNN, LSTM, and LSTM-CNN sentiment classification frameworks in the IMDb Film Reviews to identify the most appropriate dataset architecture. Exploratory findings showed that the 91% F-score for CNN outperforms the existing LSTM, LSTM-CNN, and other cutting-edge approaches to the classification of sentiments in IMDb film reviews. [25].

A. Gupta et al. (2019)This review is about the use of machine learning technologies to analyze the feelings of postings on the site for microblogging - Twitter. There have been the application of four main machine learning methods: DTC, LR, SVM, as well as NN. They evaluate the findings and address the benefits of various approaches compared to others [26].

P. Karthika et al. (2019)The dissertation being proposed is evaluated using the RF-referred ML algorithm & simulated using SPYDER. In their scheme, they quantify the accuracy, precision, F-measure, or recall for both the RF & SVM algorithms, & then compare the accuracy of the 2 algorithms. Where the RF outperforms the SVM by an average of 97 percent. [27].

L. Mandloi and R. Patel (2020)evaluate machine Learning Methods in this study, including the NB Classification method, the SVM Classification method, as well as the Maximum Entropy Classification method. They'll look at how this

classification algorithm analyses opinions and how accurate and precise it is in these contexts. [28].

Sandesh Tripathi et al. (2020) In the research schemes, They conducted SA on the Kaggle Bag of Words dataset of IMDb film reviews to show how useful insights can be gained from an array of textual data gathered across the Internet. These ideas are derived from the implementation of four standard ML algorithms, namely NB, LR, RF, and DT. Moreover, based on six test measures - confusions matrix, accuracy, precision, recall, F1 measure & AUC, the outcome of all four algorithms was evaluated.

III. SENTIMENT ANALYSIS

Sentiment analysis is an ML that is used to interpret polarity texts from positive to negative. The automatic machine learns how to analyze human feelings without intervention or human insertion. Social media was a part of people's lives; people use social media to study some areas of politics, film review, and commercialization. There are several websites such as Twitter, Facebook, and Instagram. They use this website to voice their opinion on many subjects. They are the conduit. Sentiment Analysis thus analyses the text that somebody from the various countries has inserted using the training data collection to analyze the sentiments of that specific language, understanding its emotions.

The terms opinion, sentiment, view, and beliefs are being used interchangeably, but they vary.

- **Opinion:**An open-ended inference (because various experts have various views).
- **View:**Opinion Subjective.
- **Belief:**Intellectual approval and conscious recognition.

- **Sentiment:**the opinion of one's emotional responses.

A. Levels of Sentiment Analysis

- **Document-level:** -Examination of the document stage used for the whole document. A single-subject paper is included with the level of classification. The customer has an attitude towards comparing two or two subjects in the study of the level of documents. The supervised and unchecked methodology for the characterization at the SA of document stage.
- **Sentence level:** -The classification of subjectivity is directly linked to the study of sentence level. Examination of sentence-level is to locate the word from the sentence as good, negative, or neutral. For sentence-level sentiment analysis, all classifiers in the text level analysis are used.
- **Aspect level:** -The sentiment analysis for the aspect level is used to detect sentiments about the aspect of these individuals. Let's take this instance. "My car has decent dealing but it is a little bit." There is an opinion in this case of a vehicle that the handling of a car is good but the car is negative. The competitive comment is part of the SA on the aspect stage.
- **Phrase level:** -The phrase with the words of opinion identifies the designation of their expression level. These have both benefits and drawbacks because the benefit is that they have an exact opinion. Yet relational polarity is disadvantaged so that the outcome cannot be precise.
- **Feature Level:** -The feature of the product is to define product attributes. In the text Analysis of these elements for the identification of feelings is called a feeling analysis at a functioning stage. The positive, negative or neutral view is based on the features extracted [5,6].

B. Applications of Sentiment Analysis

- **Online Commerce:**E-commerce practices are most often used for opinion analyses. Websites allow their users to share their shopping & product quality experience. The consumer and various characteristics of the product are described by giving assessments or grades. Customers can read reviews and recommendations on all products and individual product features quickly.
 - **Voice of the Market (VOM):**VOM determines what consumers know about rivals' goods or services. Precise and appropriate market intelligence contributes to the strategic edge and production of the new products.
 - **Voice of the Customer (VOC):**The Customer's voice is concerned with what each customer says about goods or services. It requires the analysis of consumer reviews and comments. VOC is a central factor in the management of user experience.
 - **Brand Reputation Management:** BRM is concerned with the business reputation. Customers' or some other party's opinions will harm or improve your credibility. BRM is not only a consumer but a product and organization. Now, a whole series of discussions are held online at a fast pace.
 - **Government:**By evaluating public perceptions, SA makes the government evaluate its strengths and disadvantages. For instance, "What do you expect from the truth if that is the state? The MP who examines the 2g racket is profoundly unethical." This post shows pessimistic government sentiments.
- corresponds to. The core resolution will increase opinion mining accuracy.
- **Temporal Relations:** For sentimental research, the time to examine may be significant. The reviser will think Windows Vista is nice in 2008 but now, because of the latest Windows 7, he might be pessimistic in 2009. To test those opinions which are modified in time, the efficiency of the sentiment analysis method may be improved.
 - **Sarcastic sentences:** Sarcastic & sarcastic phrases can be the text. "It stopped running on the second day, which is a wonderful car, for instance." Positive words may have a negative implication in such cases. It may be hard to recognize sarcastic or satirical phrases that can generate misconceptions.
 - **Requirement of World Knowledge:** Knowledge of the facts, events, and people of the world is also essential to classify the text correctly. See the example below. "Casablanca and rice and fish lunch: a great Sunday" The scheme of not having global awareness categorizes the above phrase as "nice," but it is a phrase as the name of the popular movie is Casablanca.
 - **Domain Considerations:**The precision of the classification of sentiment can be determined by the domain of the elements used. It's how certain words switch their meaning from domain to domain. "Go read the novel," for instance. This phrase has a positive impression on the book domain and a negative impression on the film domain.
 - **Review Spam Detection:** Numerous people write misleading reviews, referred to as spam reviews, to encourage their goods by offering undeserved favorable views or to defame the products of their rivals by providing false-negative views. The spam detection challenge affects industrial societies greatly. If several spams are presented

B. Challenges for Sentiment Analysis

- **Coreference Resolution:** The main resolution is the issue of what a pronoun or a noun sentence

in the opinions, the user experience would be affected. [7].

IV. MOVIE REVIEW USING IMDB

A movie review is a work that reflects the opinions of the authors on and criticizes a particular movie positively or negatively, allowing everyone to appreciate the general idea of the film and decide to choose whether or not to watch it. A film analysis will influence the entire crew involved in the film. An analysis shows that the popularity or failure of a film depends in certain ways on its ratings. A significant issue, however, is to be captured, retrieved, quantified, and analyzed more accurately by classifying movie reviews. The rating of movie reviews in favorable or negative opinions is linked to terms in the review text as well as the climate is shown positively or negatively previously. These considerations help to improve the process of assessment comprehension using SA where SA is now the portal to customer understanding [8].

Reviews are short texts which usually convey an opinion on films or goods. These reviews play an essential role in the popularity of the films or goods' sales. People usually look at blogs, review websites such as the IMDb to find out about films, teams, reviews, and scores. Therefore the public is not only brought to the theatres by the word of mouth; reviews still play a leading role here. This means that SA in film reviews simplifies the process of summing up opinions by removing the feeling the critic expresses. Text is the semantic material of video frames. It can be readily obtained compared to other semantic contents, & is useful for explaining the contents of video files, and allows applications including keyword video recovery, automated video recording, and text-based picture indexing. [9].

A. Movie Review Dataset

IMDb is the database that is the most used to obtain information about a film worldwide. IMDb Movie Review Dataset is included in the proposal because of its success and the availability of a wide range of reviews about a single film. It is one of the standard benchmarks used in film criticism for SA. The dataset includes 25,000 optimistic and negative assessments each. We have, indeed, randomly selected 5000 test feedback due to constraints of computing tools. It is a dataset of 25 000 sentimentally (positively/negative) film ratings from IMDB. Reviews are pre-processed and all reviews are represented as word indexes (integers). Words in a dataset are conveniently sorted by the total frequency to encode the "3," for example, the 3rd most common term in the report. This allows rapid filtering, like "Just take the top 10,000 words into account but delete the top 20 most widely used words" [10,11].

V. MACHINE LEARNING (ML)

The techniques to machine learning focus on machine learning algorithms to evaluate sentiments as a typical problem in the classification of syntaxes and language functions. These algorithms use a series of approaches in which the data package provides the structure to identify and detect various patterns. The trained models are then used to recognize and distinguish unexplored patterns based on their emotions. Two methods in machine learning, i.e. unsupervised learning and supervised learning, are introduced. Supervised learning usually occurs where the outcome can be estimated by the given input and if the result is unpredictable unsupervised learning is carried out. Different algorithms in machine learning, which fall under supervised learning algorithms such as naive Bayes, decision tree learning, etc, or

unsupervised learning techniques, such as neural networks and cluster approaches, etc[12].

- **Supervised Learning:**In supervised learning, we study an objective feature that can be used as accepted or not authorized to estimate the values of a distinct class attribute. The ML algorithm predicts a certain sample collection and supervised learning algorithms check for trends inside data-point labels. The algorithm is a consequence variable from which a specific set of predictors, i.e. different variables, is to be expected. We create a function using these variables, which maps the input to the desired results. The phase of testing continues until the model is correct with the training results. This whole technique purposes to decrease manual checks and software outlays. Examples: NN, Regression, DC, KNN, LR, SVM, NB, and so on.
- **Unsupervised Learning:**unsupervised learning is mentioned as studying valuable framework of out-of-label classes, design parameters, input signal, or some other information outside raw data. We may not have a goal variable in this algorithm to predict forms in which we have no data point mark or may claim the class labeling of training data is unclear. This method is used to assemble data into the cluster category & explain its organization. The clustering of data shows valuable partitions as well as hierarchies. Examples: clustering fuzzy, clustering hierarchy, K-means.

A. ALGORITHMS

ML is now one of the most important IT remains. With the growing availability of data, there is cause to think that intelligent data processing is becoming more prevalent as a critical component for technical purposes. All of ML's work is about addressing these issues and finding a response. An algorithm can

model an issue based on its experiences with the context or environment or input data in various ways.

- **Classification:**Classification assigns new records to (predefined classes) advance the establishment of classes to predict the future outcome that is true or false describes range high or low. Classification is the most popular data mining technique used to create a model that can classify the population of data in general pre-classified instances.
- **Clustering:**Clustering is the collection of a similar group of objects in data. Another form of the cluster is dissimilar of objects. The detection of related types of objects may be considered clustering. By using clustering methods we can detect dense and distributed areas in object space and can detect overall patterns and associations between data attributes.
- **Prediction:**For prediction, the regression method can be modified. Regression analysis is used to map out the connection between one or more separate variables& the dependent variable.
- **Association rule:**Association technique is used to find frequent patterns, correlates from datasets. The Association rule is mainly used to express sales. Association rule mining is used to analyzing customer behavior and predict customer behavior in a supermarket. Mining of association rules discovers the common patterns of itemsets. It is designed to extract important associations, frequent patterns & similarities between communication framework sets.
- **Neural networks:**NN is a series of connected I/O units with a weighted current on each attachment. Neural networks have an outstanding potential to infer significance from complex or incorrect information. They can be applied to identify correlations and predict trends that can be too complex for people or other computational technology to notice.

- **Regression:** It is another predictive data modeling type known as the supervised learning technique. This analysis depends on another attribute present in the same item regression that knows the target value. Regression models are computed by many statistical measures and the difference between predicted and expected values—E.g., prediction of house value and some rooms, size, etc [13].

B. CLASSIFICATION TECHNIQUES

Data mining classification techniques can process enormous data. It can forecast categorical class labels and classify data training/testing sets, so new data available can be classified. It can thus be defined as part of data mining and is becoming increasingly common [14].

Classification is a method used by DM/ML for feature vectors in the prevision of group membership[15]. However several methods are possible for machine learning, classification is commonly used[16]. Classification is an admirable activity in machine learning, specifically in future designs and exploration of information. [17].

In this section, important classification algorithms were described.

1. Support Vector Machine

The classification aim in SVM is to distinguish the two groups by preparing the relevant data with a feature and produced a classifier that works well on additional data that are not visible. The simplest type of SVM is the maximum range classification. It is common for binary classification of linear separable training data to resolve the main classification problem.

2. Artificial neural network (ANN)

ANN is a biological NN computational model. NN is also known as ANN. The idea of ANN

comes primarily from the biology subject, in which the NN plays a central and significant role in the human body. Practice in the NN is performed in the human body. A NN may be a set of associated input/output units, where each connection has a corresponding weight.

3. Decision Tree (DT)

In classification, DT algorithms are most frequently used. It also helps to simplify the classification task. DT offers an easily defined modeling technique. A decision tree is a straightforward tool that allows users to quickly follow a tree structure to see how the decision is taken.

4. K- Nearest Neighbor (KNN)

NN is a method used to determine the unknown data point focused on the nearest neighbor whose value is already identified. Closest Point Search. 2 strategies can be divided into the NN mechanism. Structure and function are fewer NN techniques of classification. K-NN is classified as less method under the scheme. KNN method exploits the NN for the value of k, which defines how numerous NN to a sample data point must be examined in the description of the class. The NN strategy is classified into two groups, i.e. KNN dependent structure and KNN less structure.

5. Naive Bayes

The NB classifiers are a class of simple probabilistic classifiers that help the application of bays theorems with durable assumptions of independence between decisions. Bayes' classifier depends on the contingent occurrences and is thus likely to occur in the future as a result of the preceding occurrence of a continuous event.

6. Bayesian network

BN is also known as the belief networks. A BN is a visual-spatial distribution of probabilities. There are two elements of this BN. The first element is primarily a directed acyclic graph

(DAG), that calls the graph nodes the random variables & which depicts probabilistic addition problems on the subsequent random variables on the edges seen between nodes or random variables [18,19].

Table I - Summary & comparison of numerousSAMethods

S. No.	Methods	Learning methodology	Advantages	Disadvantages
1.	SVM	Supervised	<ul style="list-style-type: none"> • Very high accuracy • Lesser overfitting • Robust to noise • Few irrelevant features • Document vectors are sparse • High dimensional input space 	<ul style="list-style-type: none"> • Incapable of multiclass classification • Computationally expensive • Slow
2.	Naïve Bayes (NB)	Supervised	<ul style="list-style-type: none"> • A simple and intuitive approach • It blends reliability with very good precision • Faster preparation and classification • Not sensible of irregular features • Treated streaming data well. 	<ul style="list-style-type: none"> • Assumes independence of feature • Less accurate than SVM • Mostly used while the training set is smaller • The language characteristics are conditionally independent
3.	Centroid classifier	Supervised	<ul style="list-style-type: none"> • High dimensional data set • Can combine several features. • Low computation costs. 	<ul style="list-style-type: none"> • Class term dependence • Training data vulnerable too • Many features in the vector
4.	KNN	Supervised	<ul style="list-style-type: none"> • Easy and simple to comprehend • Strong to noisy training data • Manages huge data sets • Capable of handling a large number of data. • Premised on the reality that the designation of an instance is very similar to others next to it in the space of the vector 	<ul style="list-style-type: none"> • Biased by K • High machine performance • Makes irrelevant attributes easy to mislead • Computationally intensive recall • Large storage required
5.	Winnow classifier	Supervised	<ul style="list-style-type: none"> • Failure to implement oriented • More vulnerable to an interaction between functionalities 	<ul style="list-style-type: none"> • Less accurate than SVM • Not robust tuning on various training sets

			<ul style="list-style-type: none"> The only active characteristics are weights changed 	
6.	K-means clustering	Unsupervised	<ul style="list-style-type: none"> Due to a failure to execute focused Multi-functionality interactions more fragile Weight changes are the only successful features 	<ul style="list-style-type: none"> Less precise than a controlled study Hard to estimate K Does not perform well in various sized and density cluster
7.	ME Classifier	Supervised	<ul style="list-style-type: none"> Individual features like the NB model are not assumed by this approach. Can manage high data volumes. 	<ul style="list-style-type: none"> Simplicity is hard
8.	N gram SA	Supervised	<ul style="list-style-type: none"> Using 1 and 2 grams to forecast sentiment Will enhance the model accuracy as compared to the term characteristic. 	<ul style="list-style-type: none"> Based on the fact that a corpus of data is to be trained from
9.	Multilingual SA	Supervised and Unsupervised	<ul style="list-style-type: none"> messages in various languages without a translation will be interpreted Fifteen languages operated with 	<ul style="list-style-type: none"> Corpus of various languages training is essential
10.	Feature Driven SA	Unsupervised	<ul style="list-style-type: none"> It is a succinct method that can be adapted to software development 	<ul style="list-style-type: none"> Not involved in smaller initiatives

The reliability of different sentiment analysis techniques and methods was used to evaluate their performance. That is, how much of the sentence was correctly classified by the sentiment analysis technique? Table 1 provides a short comparison of various methods used in sentiment classification. Using different methods for sentiment analysis can different outcomes. Each method has advantages and disadvantages. The machine learning approach produces the best results when key variables such as performance, reliability, and accuracy are considered, and the majority of the efforts have been made in this method. Several methods for accomplishing this mission have advanced, as seen in Table 1.

VI. CONCLUSION

Sentiment analysis can be characterized by text analysis, natural language processing, and linguistic processes as emotion detection mechanisms or emotion identification. Sentiment Analysis of casual short texts is a difficult process. Owing to the small number of characters and the enormous dimensional characteristics, the complication grows. This paper covers different classification methods and a study of some of them used in sentiment analysis. Sentiment analysis is a broad range of methods, including machine learning, artificial intelligence, data processing, statistics, and design identification, for large-scale analysis of data..Sentiment analysis uses machine learning techniques and without any human interruption, the machine will give an accurate sentiment of the people. In this, we are seeing some papers which help the new researcher to found a proper path for their new research.

VII. REFERENCES

- [1]. Vishal A. Kharde & S.S. Sonawane," Sentiment Analysis of Twitter Data: A Survey of Techniques", International Journal of Computer Applications (0975 – 8887), Volume 139 – No.11, April 2016
- [2]. N.Mahendran," A Survey: Sentiment Analysis Using Machine Learning Techniques for Social Media Analytics" International Journal of Pure and Applied Mathematics, Volume 118 No. 8 2018, 419-423
- [3]. Suraj D. M., Rohan A. R. and Dr. Vimuktha Evangeleen Salis (2019)," Survey on Sentiment Analysis", International Journal of Engineering Research & Technology (IJERT), ISSN: 2278-0181, Vol. 8 Issue 04, April-2019
- [4]. Pooja Kamavisdar," A Survey on Image Classification Approaches and Techniques", International Journal of Advanced Research in Computer and Communication Engineering Vol. 2, Issue 1, January 2013
- [5]. Suraj D. M., Rohan A. R et al. Survey on Sentiment Analysis," International Journal of Engineering Research & Technology (IJERT)", ISSN: 2278-0181, Vol. 8 Issue 04, April-2019
- [6]. Nirag T. Bhatt," Sentiment Analysis using Machine Learning Technique: A Literature Survey", International Research Journal of Engineering and Technology (IRJET), Volume: 07 Issue: 12 | Dec 2020
- [7]. Vishal A. Kharde," Sentiment Analysis of Twitter Data: A Survey of Techniques", International Journal of Computer Applications (0975 – 8887) Volume 139 – No.11, April 2016
- [8]. Mais Yasen, Sara Tedmori," Movies Reviews Sentiment Analysis and Classification", https://www.researchgate.net/publication/332321070_Movies_Reviews_Sentiment_Analysis_and_Classification/link/5cadd038458515cd2b0d62e5/download
- [9]. H. M. Keerthi Kumar¹, B. S. Harish², H. K. Darshan³," Sentiment Analysis on IMDb Movie Reviews Using Hybrid Feature Extraction Method", Regular Issue, International Journal of Interactive Multimedia and Artificial Intelligence, Vol. 5, N^o 5
- [10]. https://www.researchgate.net/publication/330014159_Sentiment_Analysis_on_IMDb_Movie_Reviews_Using_Hybrid_Feature_Extraction_Method/link/5ce44f2299bf109da773a72/download
- [11]. Shi, S., Cheng, T., Xiao, S., & Lv, X. (2009). Text Processing in Video Frames with Complex Background. 2009 International Forum on Information Technology and Applications. doi:10.1109/ifita.2009.128
- [12]. Francesco Musumeci, Member, IEEE, Cristina Rottondi, Member, IEEE, Avishek Nag, Member, IEEE, Irene Macaluso, Darko Zibar, Member, IEEE, Marco Ruffini, Senior Member, IEEE, and Massimo Tornatore, Senior Member, IEEE," An Overview on Application of Machine Learning Techniques in Optical Networks", arXiv:1803.07976v4 [cs.NI] 1 Dec 2018
- [13]. V.A. Kanimozhi¹, Dr. T. Karthikeyan²," A Survey on Machine Learning Algorithms in Data Mining for Prediction of Heart Disease", International Journal of Advanced Research in Computer and Communication Engineering, Vol. 5, Issue 4, April 2016
- [14]. RAJ, M. A. 2012. Mrs. Bincy G, Mrs. T. Mathu. Survey on common data mining classification Technique. International Journal of Wisdom Based Computing, 2.
- [15]. Kesavaraj G, Sukumaran S. A study on classification techniques in data mining. in Computing, Communications and Networking Technologies (ICCCNT), 2013 Fourth International Conference on, 2013; pp. 1-7.
- [16]. Singh M, Sharma S, Kaur A. Performance Analysis of Decision Trees. International Journal of Computer Applications 2013; 71.
- [17]. N.Satyanarayana," Survey of Classification Techniques in Data Mining", IJSET - International Journal of Innovative Science, Engineering & Technology, Vol. 1 Issue 9, November 2014.
- [18]. N. Chandra Sekhar Reddy, K. Sai Prasad, and A. Mounika," Classification Algorithms on Datamining: A Study", International Journal of Computational Intelligence Research, ISSN 0973-1873 Volume 13, Number 8 (2017), pp. 2135-2142

- [19]. Aized Amin Soof," Classification Techniques in Machine Learning: Applications and Issues", Journal of Basic & Applied Sciences, 2017, 13, 459-465
- [20]. M. Butler and S. Robila, "Interface for querying and data mining for the IMDb dataset," 2016 IEEE Long Island Systems, Applications and Technology Conference (LISAT), Farmingdale, NY, USA, 2016, pp. 1-6, doi: 10.1109/LISAT.2016.7494103.
- [21]. A. Yenter and A. Verma, "Deep CNN-LSTM with combined kernels from multiple branches for IMDb review sentiment analysis," 2017 IEEE 8th Annual Ubiquitous Computing, Electronics and Mobile Communication Conference (UEMCON), New York, NY, USA, 2017, pp. 540-546, doi: 10.1109/UEMCON.2017.8249013.
- [22]. Dipak R. Kawade," Sentiment Analysis: Machine Learning Approach", International Journal of Engineering and Technology (IJET), Vol 9 No 3 Jun-Jul 2017
- [23]. A. U. Hassan, J. Hussain, M. Hussain, M. Sadiq and S. Lee, "Sentiment analysis of social networking sites (SNS) data using machine learning approach for the measurement of depression," 2017 International Conference on Information and Communication Technology Convergence (ICTC), Jeju, Korea (South), 2017, pp. 138-140, doi: 10.1109/ICTC.2017.8190959.
- [24]. S. Mathapati, A. K. Adur, R. Tanuja, S. H. Manjula and K. R. Venugopal, "Collaborative Deep Learning Techniques for Sentiment Analysis on IMDb Dataset," 2018 Tenth International Conference on Advanced Computing (ICoAC), Chennai, India, 2018, pp. 361-366, doi: 10.1109/ICoAC44903.2018.8939068.
- [25]. M. R. Haque, S. Akter Lima and S. Z. Mishu, "Performance Analysis of Different Neural Networks for Sentiment Analysis on IMDb Movie Reviews," 2019 3rd International Conference on Electrical, Computer & Telecommunication Engineering (ICECTE), Rajshahi, Bangladesh, 2019, pp. 161-164, doi: 10.1109/ICECTE48615.2019.9303573.
- [26]. A. Gupta, A. Singh, I. Pandita and H. Parashar, "Sentiment Analysis of Twitter Posts using Machine Learning Algorithms," 2019 6th International Conference on Computing for Sustainable Global Development (INDIACom), New Delhi, India, 2019, pp. 980-983.
- [27]. P. Karthika, R. Murugeswari and R. Manoranjithem, "Sentiment Analysis of Social Media Network Using Random Forest Algorithm," 2019 IEEE International Conference on Intelligent Techniques in Control, Optimization and Signal Processing (INCOS), Tamilnadu, India, 2019, pp. 1-5, doi: 10.1109/INCOS45849.2019.8951367.
- [28]. L. Mandloi and R. Patel, "Twitter Sentiments Analysis Using Machine Learning Methods," 2020 International Conference for Emerging Technology (INCET), Belgaum, India, 2020, pp. 1-5, doi: 10.1109/INCET49848.2020.9154183.
- [29]. S. Tripathi, R. Mehrotra, V. Bansal and S. Upadhyay, "Analyzing Sentiment using IMDb Dataset," 2020 12th International Conference on Computational Intelligence and Communication Networks (CICN), Bhimtal, India, 2020, pp. 30-33, doi: 10.1109/CICN49253.2020.9242570.

Cite this article as :

Karishma Kaushik, Mahesh Parmar, "Sentiment Analysis Based on Movie Reviews using Various Classification Techniques : A Review", International Journal of Scientific Research in Computer Science, Engineering and Information Technology (IJSRCSEIT), ISSN : 2456-3307, Volume 7 Issue 3, pp. 197-208, May-June 2021. Available at doi : <https://doi.org/10.32628/CSEIT217329>
Journal URL : <https://ijsrcseit.com/CSEIT217329>