Fine Grained Analysis of Intention for Social Media Reviews Using Distance Measure and Deep Learning Technique

R. Akila[1] and S. Revathi[2*]

[1]Assistant Professor (Sr. Gr), Department of Computer Science and Engineering, B.S. Abdur Rahman Crescent Institute of Science and Technology, Tamil Nadu. niceakila@gmail.com
Orcid: https://orcid.org/0000-0003-4000-4535

[2*]Professor, Department of Computer Science and Engineering, B.S. Abdur Rahman Crescent Institute of Science and Technology, Tamil Nadu. srevathi@crescent.education
Orcid: https://orcid.org/0000-0001-9584-5089

Received: February 14, 2023; Accepted: March 28, 2023; Published: May 30, 2023

Abstract

Intent analysis and classification are performed to identify the expressions of intent in the given text. In this paper, the dataset is classified into emotion classifications by utilizing machine learning model SVM, Bipolar classification, Fine Grained Analysis, and Sarcasm detection, with Naïve Bayes and Random Forest techniques of deep learning, including Long Short-Term Memory to perform intention analysis on social media data. Then Fine-grained or Multi-Class Sentiment analysis is used for further classification of the five classes, viz. negative, strong negative, neutral, positive, and strong positive, which detects the sarcastic reviews in the movie dataset. The emotional intention behind the review comments is classified as happiness, rage, sadness, joy, anger, and disgust by using SVM. The reviews are analyzed and calculated based on their subjectivity and context level similarity using Related Relaxed Word Mover Distance (RRWMD) semantic similarity measure. With the advantage of the RRWMD algorithm, the reviews from the context containing deviated or irrelevant contents were removed before being applied to the classification algorithms, thereby reducing the execution time, which obtains a 3% improvement in accuracy. The disadvantage of the RRWMD algorithm is only one deep learning algorithm is compared. From the observed accuracy scores and classification reports, the LSTM has provided higher accuracy, despite the long execution time. The Naïve Bayes model has produced lower accuracy than the neural network model but was efficient, taking less time to fit and classify. The results from various experiments have proven that the semantic similarity measure provides more accurate results than the state-of-the-art model.

Keywords: Fine Grained Analysis, Semantic Similarity, Deep Learning, Emotion Classification, Related Relaxed Word Mover Distance (RRWMD).

1 Introduction

Social media and micro-blog tools are progressively utilized by people to communicate their intentional state and opinions as short instant messages. Intent mining aims to extract the information from the textual reviews using machine learning processes, which are closely related tothe text mining process

Journal of Internet Services and Information Security (JISIS), volume: 13, number: 2 (May), pp. 48-64.
DOI: *10.58346/JISIS.2023.I2.003*

*Corresponding author: Professor, Department of Computer Science and Engineering, B.S. Abdur Rahman Crescent Institute of Science and Technology, Tamil Nadu.

and Natural Language Processing (NLP). Using this analysis, we can predict the reviewer's attitude, state of mind at the time of providing the movie reviews, and the emotions (Merriment, Disgust, Rage) which make them provide the negative, somewhat negative, neutral, somewhat positive, and positive reviews.

This paper is focused on the Intent Mining process for predicting the user intention about movies by considering the reviews through analyzing the positive and negative opinions or sentiments of the users for that particular type of movie, which is more important for other people to know about the movies. By predicting the sentiment of the movies, the user can gain more information about the movie they want to know with the help of reviews given by the person who has already viewed it. To perform this prediction the datasets obtained from Kaggle are used which consist of positive (1), and negative (0) sentiment and reviews of movies. Before using the dataset, pre-processing should be done to detach the noises such as HTMLtags, stop words (am, is, are, the, and, a), and non-letters which can be visualized.

Using the Bag of words model, the reviews are represented as a numerical feature vector. The total count of the words was calculated from the reviews. The words which are frequently occurring in the feature vectors were re-weighted using the transformation techniques. Then the model is trained for the movie review classification.

The Movie Review Rating Prediction (MR2P) was implemented byusing five different machine learning algorithms (Adetunji et al., 2020). The results of this study will help movie stakeholders (producers, directors, crew, and cast who already work in the industry or who aspire to) understand the type of film to make, invest in, which will benefit the stakeholders in terms of increased profit.

The algorithms like Naive Bayes, SVM, Random Forest, and LSTM are used out of which the LSTM is the most convenient one as the precision obtained is higher compared to other models. It helps in the detection of the intention of users with the reviews given to the movies, which is required for the person who wants to know more about the different emotions or sentiments of the other viewers of that movie. In this way, details of the movies can be gained without spending any time watching that particular movie.

To detect the intention of review furthermore, the Fine-grained analysis, also called Multi class analysis is done in this paper. For performing this challenging activity, the sentiment for the movie reviews has been classified into five classes such as positive, negative, neutral, somewhat positive, and somewhat negative. Each of these sentiments was denoted by the numerical values from 0-4 which is used for the fine-grained sentiment analysis. From this process, people can get more detailed information about movies compared to the previous process of analysis which consists of binary labels. Also, this fine-grained sentiment analysis helps to retrieve the data from the comparative expressions or reviews of the people.

Sarcasm is a comment or review that is hardly difficult to define in general. It may be the words denoting the intent of humor or made to hurt. While analyzing the movie reviews, there are lots of possibilities for the presence of sarcastic reviews, because those reviews are taken from directly collected words, and sentiments of people or viewers of the movies, which can be related to the quality of the person who acted in that story. The sarcastic words or statements generally mean the opposite of the meant words. Sarcasm detection is a challenging task in sentiment analysis which is difficult to detect because it may positively refer to negative emotions. So, it is highly important to overcome this challenge of sentiment analysis to get the correct meaning of the sentiment of the people.

By detecting the sarcasm in the reviews, the user can obtain better meaning and understand the accurate thought of the reviews provided by the other viewers. This sarcasm detection process of movie reviews also helps the persons who have been involved directly in the movies like directors, actors, and actresses to improve their skills in attracting the viewers, by updating the ideas in the required way as expected by their audience so that it reaches them quickly with the help of the reviews suggested to their movies. As a result of the analysis made from the reviews and sentiments, the intention of the users is predicted accurately by overcoming the challenges in analyzing the sentiment. The semantic similarity index of the movie comments or reviews can be calculated using RRWMD (Werner and Laber, 2020).

Word Mover's Distance (WMD) is a fundamental method for calculating how similar two documents are to one another. By using an optimum transport formulation, the core of WMD may take use of the word space's underlying geometry.

In Relaxed Word Moving Distance (RWMD), when p - the number of different words, the optimum average time to solve WMD is around O (p3 log p). There are two methods to improve computation speed because it is a little slow.

The first one summarizes the lowest bound distance between and is called Word Centroid Distance (WCD). The second method, known as Relaxed Word Moving Distance (RWMD), uses the closet distance without taking into account the possibility of many words becoming single words.

Related Relaxed Word Mover's Distance (RRWMD) is among the most effective due to its ease of use, efficiency, and speedy deployment. We suggest a method to accelerate WMD and RWMD based on presumptions that are validated by empirical characteristics of the distances between embeddings.

2 Related Work

(Matheus and Eduardo, 2020) have proposed the Distance Measure for computing the similarity among documents. They have also discovered that the properties of the application are mostly on the distribution of the distances. The distance between related and unrelated words is computed using RRWMD. For the larger dataset of operations, this similarity measure is more useful and occupies less memory space thus speeding the computation process.

The Researchers (Bouazizi and Ohtsuki, 2019) have discussed the Twitter Sentiment analysis using the Multi-class classification, by comparing the accuracy obtained by the classification with the accuracy of Binary classification, to know about the efficiency and to understand more about the relationship between different sentiments using some of the metrics.

(Molinera et al.,2020) have proposed a novel on group decision-making for dynamic context. Since the social network is a rapidly growing place that contains a large amount of information thus resulting in a dynamic changing environment. Sentimental analysis is used to order the text in each round and interval Type-2(IT-2), Hesitant Fuzzy (HF) ontology to store the best alternative result. Finally, they have combined both interval type-2 and hesitant fuzzy sets to represent the text within the ontology.

In this paper (Ruza et al., 2020), the researcher has addressed the issues of sentimental analysis while critical events like disasters, social events, etc., occur naturally. Bayesian network classifiers are used to perform sentiment analysis on Twitter data. In addition, Support Vector Machine (SVM) and random forest has been used. But when compared to Bayesian networks classifiers predict the better result. They also bring TAN and BF TAN present in it to attract quality information socially and historically grasp the important features of the dynamics event, even in the case of the reduced number of training examples.

(Araújo et al., 2020), the authors have discussed MT for MS-level SA (Sentiment Analysis) by translating the input texts in other languages (except English) to English and performing sentimental analysis using the existing method for English. They have also analyzed the existing method of sentimental analysis for English.

In this paper (Gan, 2019), the researcher talks about targeted sentiment analysis based on the concept of sparse attention separable dilated Convolution Neural Networks (CNN). The methods of LSTM and classical Convolution Neural Networks (CCNN) have been used for analyzing sentimental analysis which consists of four layers i) Multichannel Embedding Layer ii) module of separable dilated convolution iii) sparse attention layer and iv) output layer.

(Ranganath et al., 2020), the authors have proposed several methods to detect sarcasm in social networks like Twitter. They took the Twitter data and pre-processed using Text Blob and then performed polarity detection using Rapid Miner which is then validated using Weka. They have even used emojis to detect sarcasm in their tweets. This paper has also discussed various techniques to detect sarcasm like learning models and deep learning models.

In this paper (Kumar et al., 2020) Sarcasm Detection has been done using Multi-Head Attention on the concept of Bidirectional LSTM. Sarcasm is mainly handled to express a negative opinion in social media using positive or intensified positive words. They have proposed MHA-BiLSTM (Multi-Head Attention-based Bidirectional Long-Short Memory), a network to detect sarcastic comments inthe given dataset. The result obtained after performing the simulation of BiLSTM has achieved a good result compared to feature-rich SVM models.

(Chatterjee et al.,2018) have proposed a novel Deep Learning-based way to deal with distinguishing feelings of Happy, Sad, and Angry in printed discourses and proposed a Deep Learning based methodology called "Intent and Semantic-Based Emotion Detector (SS-BED)". The embodiment of the methodology lies in consolidating both semantic and intent-based portrayals by progressively discovering the exact feelings. The author utilizes semi-robotized techniques to accumulate huge scope of preparing information with differing methods of communicating feelings in the model. Assessment of the methodology on genuine discourse datasets reveals that it altogether beats conventional Machine Learning baselines just as the other off-the-rack Deep Learning models (Johnson, C., 2020).

(Shah et al., 2019) have proposed a directed machine classifier to recognize feelings utilizing WORDNET and EMOSENTICNET apparatuses and actualized it with the proposed model which performs Emo Sentic Net superior to WordNet, which practically classifies all the fundamental classifiers to give nearly a similar accuracy while distinguishing the feelings and the author has closed it with emotional recognition, perhaps the hardest issue to explain. Identifying feelings from text is a part of the testing work and the greater part of the examination works have some thoughtful confinements, in particular, language uncertainty, the different feeling-bearing content, text which doesn't contain any feeling words, and so forth.

A classification and summarizing strategy for movie reviews has been put out by (Atif Khan et al., 2020). The bag-of-words feature extraction technique is employed to extract unigrams, bigrams, and trigrams as a feature set from the review documents and represent the review documents as a vector for the purpose of categorizing movie reviews. The movie reviews (represented as a feature vector) are then divided into negative and positive reviews using the Naive Bayes method.

(Umer et al., 2019) developed a sentiment-based approach for predicting whether enhancement reports are likely to be granted or refused, so that developers can prioritise requests that are likely to be approved. This could help software programme compete in the industry by improving their functionality as needed by users.

(Nizamani et al., 2017) proposed a multinomial naive Bayes strategy for predicting whether a new enhancement report is likely to be approved or refused. They obtained open-source software programme enhancement reports from Bugzilla for examination. (Lin et al., 2018) shared their experience developing a software library recommender using Stack Overflow developer opinions.

(Umer et al.,2018) presented an emotion-based automatic technique for predicting report priority. They evaluate the suggested approach on Eclipse open-source projects, and the findings of the cross-project evaluation indicate that the proposed approach outperforms the state-of-the-art. It improves the F1 score by more than 6% on average.

(Mashal and Asnani, 2017) suggested a novel solution to the challenge of determining emotion intensity from social media data.

(Williams and Mahmoud, 2017), (Bhavsar and Manglani, 2019) and (Mars and Gouider, 2017) analyzed emotions expressed in tweets which are related to software and reviews of products.

(Sahithya et al., 2019) investigated convolution function extractors and long temporary memory as two types of low-level network structures for representing review phrases.

(Riaz et al., 2019) prepared a study demonstrates that, depending on the technique used and the number of features chosen, feature selection greatly enhances the classification accuracy.

(Bavakhani et al., 2019), multilayer neural networks and deep learning techniques were utilised to extract the polarity of consumer comments and opinions in two different product/service categories, spanning from laptops to restaurants.

(Ireland and Liu, 2018), proposed framework in order to help designers make better choices, the framework tries to transform massive amounts of qualitative data into quantitative insights on product characteristics.

(Shrivastava et al., 2019) work being presented provides a brand-new corpus that expresses various emotions gleaned from a TV show's transcript.

They used semantic k-means (SKM) clustering in their analysis because it uses the Euclidean distance between sentence embeddings (the semantic representation of sentences) of related sentences. The suggested approach uses a clustering technique to get a summary from the categorized reviews after the NB classifier has classified them as positive or negative.

3 Methodology

System Design

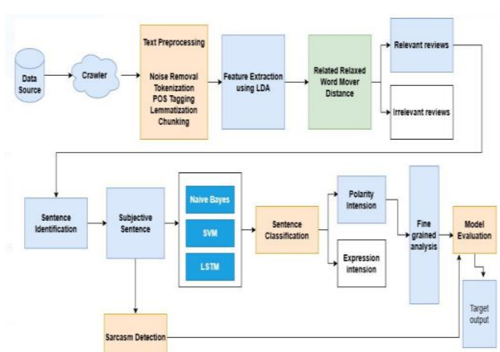

Figure 1: Architecture of Fine-Grained Analysis and Emotional Classification of Social Media Reviews

Figure 1shows the Architecture of Fine-Grained Analysis and Emotional Classification of Social Media Reviews. It shows the data source which contains review comments obtained from Micro Blogs such as social media. The crawler extracts the review comments from social media through API or from the repository and prepares it as Data Set.

The essential and primary thing in the data mining process is preprocessing. The Special symbols and erroneous messages are removed in the noise-removal process. After removing noise from the comments, the sentences are broken into words by using tokenization. The stop words such as articles, nouns, and pronouns are removed and only the adverb and adjectives are taken for analysis through POS tagging. The Lemmatization method is used for removing es and ed from the words. Chunking is a process of taking out phrases from unstructured text. POS tag output is given as input to the chunking process and gives chunks as output. Feature extraction is done to discover the number of occurrences of tokens in the document and the token appears in the number of documents which is determined by TF and IDF respectively.

Naive Bayes

A straightforward and effective supervised learning algorithm is naive Bayes. The Naive Bayes method can produce accurate predictions even with little training data. Every pair of features being classed is independent of one another, which is the essential premise on which it operates.

The simplistic presumption is that each pair of features is independent and contributes equally to the result. It is considered that the traits are profoundly independent of one another.

For instance, if a fruit is orange in colour, rounded, and around 4 cm in radius, it can be regarded as orange. The Nave Bayes method assumes that each of these characteristics independently increases the likelihood that the fruit is Orange (without any correlation between the color, shape and size)

- Naive Bayes classifiers are the most effective and straightforward supervised machine learning algorithms. They use the Bayes theorem to calculate probability.
- P(L) Represents the likelihood that a random characteristic produced the label. The prior probability that a given feature set is labeled as P(Feature/L). The last chance that a particular feature set is occurring is called P(Feature)

Random Forest Classifier

The random forest classifier is used for prediction as shown in Figure 2 and for providing the F1 scores. The output of both the vectorizer and classifier is taken in the pipeline and the features extracted are given a separate score which is an improved score from the process before the extraction of the index. A random forest is a group of decision trees formed together. The classifier extracts the features from the result of the TF-IDF model and predicts their accuracy for them. The classification report for both the vectorizer and the classifier is taken as one and the negative and positive labels are given the precision, recall, and support scores respectively.

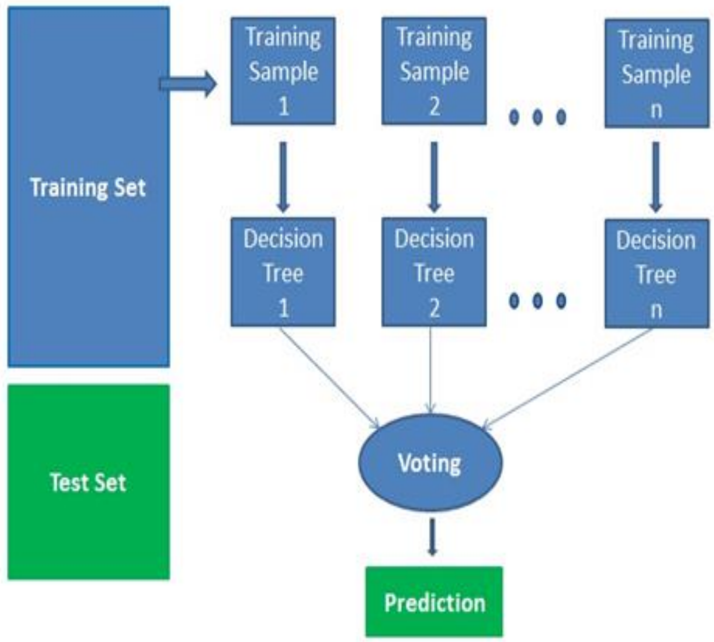

Figure 2: Random Forest Classifier

Long Short-Term Memory (LSTM)

A typical LSTM unit comprises of a cell i) input gate ii) output gate and iii) forget gate. The cell recalls esteems over subjective time stretches and the three gates control the progression of data into and out of the phone". LSTMs are a sort of deep learning model that is generally utilized for the examination of successive information (time-arrangement information prediction).Different application regions are utilized: Language model, Neural machine interpretation, Music age, Time arrangement expectation, Financial forecast, Robot control, Time arrangement expectation, Speech acknowledgment, Rhythm learning, Music piece, Grammar learning, Handwriting acknowledgment, Human activity acknowledgment, Sign Language Translation, Time arrangement peculiarity identification, Several expectation assignments in the territory of the business process the executives, Prediction in clinical consideration pathways, Semantic parsing, Object Co-division.

The feed-forward neural network has a limitation that it cannot pass feedback to prior nodes which are addressed in RNN (Recurrent Neural Network). However, the problem with RNNs is that they eventually start forgetting their initial inputs because data is lost at every stage of their processing. This calls for the use of long-term memory. Long short-term memory (LSTM) was developed to solve these RNN problems.

LSTM can remember information for a longer period. They have internal mechanisms called gates to regulate the information flow. These gates help indetermining which information to keep or discard. Thereby, it can pass relevant information down the long chain of sequences to make predictions.

RNNs are considered superior toFeedforward neural network as RNNs has feedback connections thatare missing in Feedforward neural network. A major limitation of Recurrent Neural Networks is that it suffers from short-term memory. Hence, when a lengthier sentence/paragraph is being processed for prediction, RNNs may leave out the vital information that was collected atthe start.

The neural network's weights are updated with the Gradient values. As the gradient value reduces over some time, it is referred to as the Vanishing gradient problem. As the gradient value diminishes, its learning contribution also decreases. As the learning of RNN layers diminishes, it forgot what it has learned in the beginning due to its short-term memory.

Long Short-Term Memory (LSTM) was created to overcome the issue faced due to the short-term memory problem of the RNNs. LSTMs are superior to CNN (including conventional feed-forward neural networks) and RNN inmultiple ways.

LSTMs can remember the old inputs for a long duration. Further, To regulate the flow of information, LSTM uses gates. The gates aid in deciding which input it has learned has to be kept or discarded. Thereby the relevant information it has learned over some time is used to make predictions.

Support Vector Machines

Support Vector Machines (SVM), is one of the most popular supervised learning. SVM has a wide range of applications in both classification and regression problems. Given a dataset, every element is aligned to one of the two categories; the SVM training algorithm assigns new elements/nodes to one of the two categories.

The fact that SVM may function in infinite dimensions is one of its main advantages. In multidimensional space, it establishes a margin or boundary between the data points. Finding a flat boundary "hyperplane" that results in a homogenous data division is the objective. The hyperplane with the highest separation to the nearest training data point of any class performs the optimal partitioning since the bigger the margin, the smaller the classifier's generalisation error. The goal then would be to increase the gap.

SVM algorithm can be applied to classifications or regressions. Figure 3 depicts the SVM with hyperplanes separating the data. It can be inferred that the hyperplane H3 separating the data points is the optimal one.

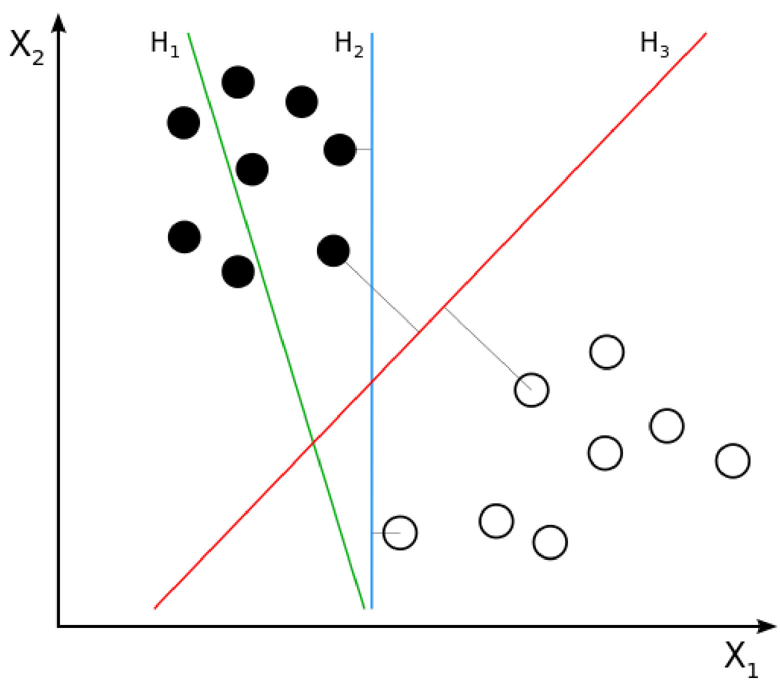

Figure 3: Support Vector Classifier

Table 1: Advantages and Disadvantages of Machine Learning Algorithms

Method	Approach	Advantages	Disadvantages
Naïve Bayes	Probabilistic Classifier	• Simple and easy to implement. • Not as much training data is needed. • Handle continuously as well as discrete data. • Effective and fast for making real-time data predictions. • Not sensitive to irrelevant features.	• Limits the applicability of the algorithm in real-world scenarios. • Faces the "zero-frequency problem," where it gives a categorical variable with zero probability.
Random Forest	Decision Tree Classifier	• Performs regression as well as classification tasks. • Handles large datasets efficiently. • Gives a higher level of accuracy in predicting.	• A large number of trees can slow down the process and render it ineffectual for making predictions in real-time. • Fast to train, but slow to create predictions.
LSTM	Deep Learning based	• Manages biases in entire data streams as well as single data points. • Identifies long-term relationships between words or numbers in sequences	• Solve the problem of vanishing gradients. • Requires a lot of memory and time to train to be prepared for use in the real world.
SVM	Linear Classifier	• When there is a significant degree of class separation, SVM performs reasonably well. • SVM performs better in large dimensional spaces • well when there are more dimensions than samples. • Uses memory rather efficiently.	• Not suitable for large data sets. • Gives poor perform when there are more training data samples than features. • Does not perform when target classes overlap and there is greater noise in the data set. • There is no probabilistic justification for the classification.

Table 1 shows the advantages and disadvantages of various machine learning algorithms

RRWMD measure calculates the distance between the words based on the semantic relationship. The semantic similarity measures closer to ‘0’ is classified as relevant comments and the semantic similarity measure closer to ‘1’ is classified as irrelevant comments. The irrelevant comments are ignored and only the relevant comments are captured for further processing. The review comments which are conveying the expressions or opinions are identified as subjective sentences. The classification algorithms of Machine Learning such as Naïve Bayes, SVM, and the deep learning algorithm of LSTM are applied to the subjective sentences for classifying the sentences as polarity intention and expression intention, which is further classified into more positive, more negative, neutral, positive and negative for fine grained analysis and expression intention, which furtherly classifies the textual messages as happy, anger and hate. The sarcastic sentences are classified as sarcastic and non-sarcastic sentences by ‘1’ and ‘0’.

The best model has to be chosen based on the accuracy, precision, and recall for sentiment prediction. The confusion matrix is constructed to evaluate the correctness of the algorithms. Finally, the user will get unbiased results which are more useful for decision-making.

Related Relaxed Word Mover's Distance (RRWMD)

Equation (1) shows the semantic dependency among words in a sentence which is calculated using RRWMD.

$$max\left\{\sum_{i=1}^{|D|} D_i \,.\sum_{i=1}^{|D|} D_i \,.\min_{w'_j \in R'(w_i)} c\left(w_{i,} w'_j\right),\ \ \sum_{j=1}^{|D'|} D'_j \,.\ \sum_{j=1}^{|D'|} D'_j \,.\min_{w_i \in R(w'_j)} c(w_i, w'_j)\right\} \tag{1}$$

w_i -ith word in Document D_i
w_j -jth word in Document D_j

To find the semantic relationship, the words list dependent to w in cache C for each word w' in the list are compared. For the large set of document collection, RRWMD has shown considerable improvement in execution time.

Pseudo Code

Input: Movie Data Set

1. Collect the dataset of reviews.
2. Pre-processing collected dataset.
3. Apply the RRWMD similarity measure
4. Apply Naive Bayes, Random Forest, LSTM
5. Apply SVM for emotional classification
6. Predicting the accuracy of the algorithms.
7. comparing the accuracy to know the best one for modeling.
8. Using bipolar classification, the intention is predicted as 0's and 1's.
9. The reviews are finely grained as multi-class values.
10. Detecting the sarcastic words in the reviews.

Pseudo Code for Distance Measure

Input: words in a sentence

for each word w in Document, D do

calculate the distance between words

if wi in D matches wi' in D' and the Similarity score is closer to 0 then

set wi and wi' are relevant

else if wi in D does not match wi' in D' and the Similarity score is closer to 1 then

set wi and wi' are irrelevant

end

Output: semantic Similarity score for words

4 Results and Discussion

The 40,000 movie reviews from the dataset including positive (1) and negative (0) sentiments were collected from Kaggle, out of which 28000 movie reviews were taken for training and 12000 movie reviews were taken for testing. The dataset which has the most finely grained sentiment movie reviews were denoted by numerical values such as 0 for negative,1 for strong negative, 2 for neutral, 3 for positive, and 4 denotes strong positive.

After training the model, the sentiment is predicted using a unique test dataset of the movie reviews which provides well clear and enhanced details to other people who need to know more about the movies.

The sarcastic words or statements are difficult to detect as it denotes the opposite meaning. In this paper, the intention of the users is predicted accuratelyby overcoming the challenge of analyzing the sentiment. These sarcastic reviews are detected by training the dataset and testing it with the movie reviews dataset and the training data set evaluation metric is given in table 2.

In this way, the correct meaning of the people's reviews about the movies was obtained which will be useful for the person who is in need to get the correct information about the movies.

With the result obtained after predicting the accuracy of machine learning algorithms, a more suitable algorithm has been found which is used for modeling the movie review datasets to predict the user intention. The prediction of user intention can be done in three ways or processes, that is bipolar analysis, fine-grained or Multi-class analysis, and the sarcasm detection of the reviews. The results obtained from the bipolar sentiment analysis contain the binary-valued sentiment for the reviews. The results obtained from the fine-grained analysis provide improved information to the users about the movies by predicting the sentiment by five discrete values from 0-4. The results obtained from the sarcasm detection process of analysis show that the reviews are sarcastic or not which helps know the correct intention of the users.

Table 2: Training Data Set Evaluation Metric

Algorithms	Without RRWMD				With RRWMD			
	Accuracy	Recall	Precision	F1score	Accuracy	Recall	Precision	F1 score
Naïve Bayes	82.2	0.87	0.68	0.76	84.36	0.9	0.71	0.78
SVM	83.94	0.77	0.72	0.74	86	0.8	0.74	0.76
Random Forest	85.5	0.85	0.82	0.79	88	0.87	0.85	0.81
LSTM	89.45	0.88	0.84	0.82	92	0.91	0.86	0.85

Table 3: Test Data Set Evaluation Metric

Algorithms	Without RRWMD				With RRWMD			
	Accuracy	Recall	Precision	F1score	Accuracy	Recall	Precision	F1score
Naïve Bayes	76.5	0.7	0.5	0.68	79	0.73	0.52	0.7
SVM	78.34	0.81	0.62	0.63	80.3	0.83	0.64	0.65
Random Forest	75.4	0.74	0.73	0.65	78	0.77	0.75	0.67
LSTM	80.35	0.8	0.79	0.78	82.6	0.82	0.81	0.82

Table 3 shows the analysis of the different algorithmsfor the test data set and table 4 shows the accuracy prediction by LSTM.

Table 4: Accuracy Predicted by LSTM

PREDICTIONS	ACCURACY (IN %)
Bipolar Sentiment Prediction	88.75
Fine-Grained Sentiment Analysis	71.51
Sarcasm Detection	75.76

Figure 4 shows the plot between intent mining processes and the accuracy of that by the LSTM algorithm. The best classification model that perfectly classifies the intent emotion classification of social media data using multi-class classification for testing datasets is LSTM, which is a deep learning architecture model that gives an accuracy of 88.75% for Bipolar Sentiment Prediction, 71.51% for Fine-Grained Sentiment Analysis and gives an accuracy of 75.76% for sarcasm detection as shown in Table 3.

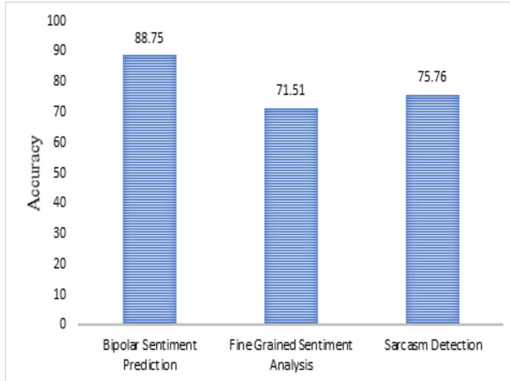

Figure 4: Accuracy of the LSTM Algorithm

In addition, these classification models with plots are compared using the Precision, Recall, F1-score values, and expressive intention classification with actual predicted values for estimating the performance of the dataset. Hence from the above values, it is concluded that the LSTM model performs well during testing for the given dataset and the SVM model performs well for intent directive class expressive classification in social media data.

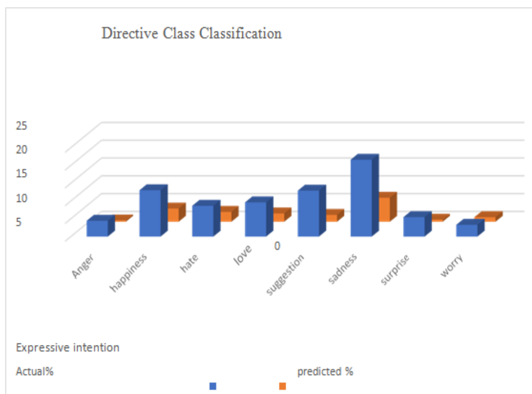

Figure 5: Directive Class Classification for Expressions

Figure 5 shows the Emotion Classification of text messages, where the directive classes of expression were conveyed in the text reviews as anger, happiness, hate, love, suggestion, sadness, surprise, and worry by using an emotional dictionary.

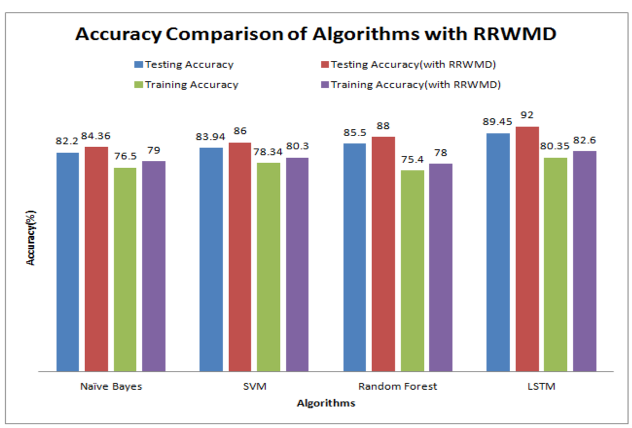

Figure 6: Accuracy – Comparing the Algorithms After Applying RRWMD

Figure 6 shows the accuracy comparison of the algorithms after applying RRWMD.The algorithms Naïve Bayes, SVM, Random Forest, and LSTM are taken and the review's context level similarity is calculated using RRWMD. After applying the RRWMD the Naïve Bayes, Random Forest, and LSTM algorithms produced 3% accuracy and 2% accuracy in the SVM algorithm during the training. In the testing period, all the algorithms produced 3% accuracy after applying RRWMD.

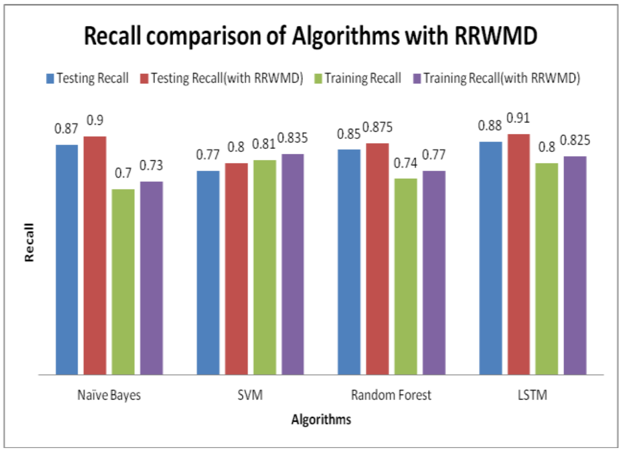

Figure 7: Recall the Comparison of Algorithms with RRWMD

Figure 7 shows the recall comparison of algorithms after applying RRWMD, where the algorithms Naïve Bayes, and Random Forest produces 3% recall and 3% recall improvement in LSTM and SVM algorithm respectively after applying the RRWMD during the training period. In the testing period, Naïve Bayes and Random Forest produced 4% recall and 3% recall improvement in SVM and LSTM, which proves that the algorithm gives better results after applying RRWMD.

Figure 8 shows that the result of precision value when comparing existing algorithms with RRWMD.

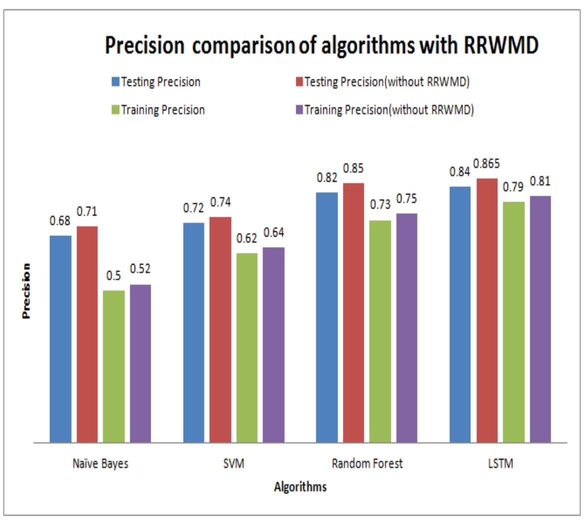

Figure 8: Precision - Comparison of Algorithms with RRWMD

Figure 8 shows the precision of comparison algorithms after applying RRWMD, where the Naïve Bayes and Random Forest algorithms achieve 4% precision and 3% precision improvement in SVM and LSTM algorithms respectively during the training period. In the testing period, 4% precision and 3% precision improvement have been achieved by the rest of the algorithms which proves that after applying RRWMD, there is good improvement in the precision for all the algorithms.

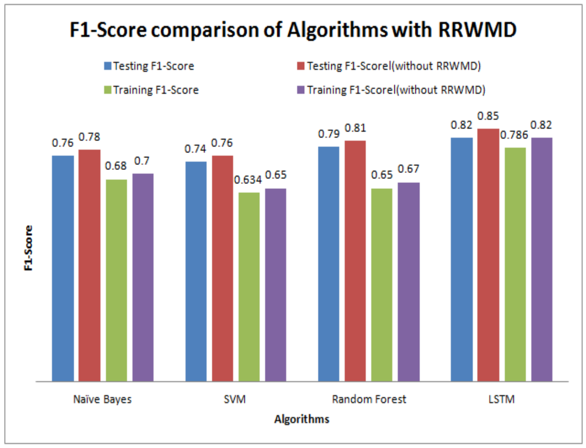

Figure 9: F1-Score Comparison of Algorithms with RRWMD

Figure 9 shows the F1-Score comparison of algorithms after applying RRWMD, where the F1-score has shown a 3% improvement for the Naïve Bayes algorithm in the SVM algorithm and a 4% improvement for the Random Forest Algorithm. The LSTM algorithm after applying RRWMD during the training and testing period shows good achievement in F1-score after applying RRWMD in the algorithms of Naïve Bayes, SVM, Random Forest, and LSTM. Table 5. Shows the summarization of approaches with ROUGE-2

Table 5: Summarization of Approaches with ROUGE-2

Methods	**Precision**	**Recall**	**F-Measure**
LSTM+RRWMD (Proposed)	60.2	61.3	60.1
Naïve Bayes + SKM (Atif Khan et al., 2020)	40.1	40.3	40.2
Lex Rank (Erkan and Radev, 2004)	30.2	30.1	30.5
Text Rank (Mihalcea and Tarau, 2004)	13.5	14.2	13.9

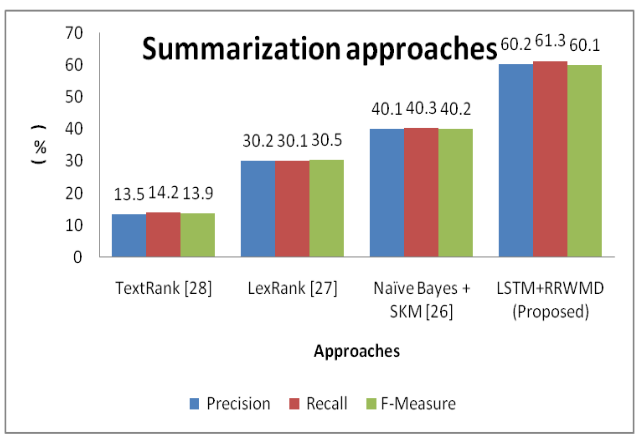

Figure 10: Summarization of Approaches

Atif Khan et al proposed the method Naïve Bayes with SKM for the movie review data set with ROGUE – 2 measures. Since our proposed approach used LSTM with RRWMD to measure the semantic similarity measure between words in sentences, Precision is improved from 40.1% to 60.2%, Recall is improved from 40.3% to 61.3% and F-Measure improved from 40.2% to 60.1%. Our proposed method is shown figure 10 gives better results in terms of precision, recall and F-Measure compared to the existing methods Lex Rank and Tex Rank.

5 Conclusion and Future Work

The process of modeling is done by using the machine learning algorithm since it is predicted to provide better accuracy and sentiment for each review in the dataset. Based on data from movie reviews, we developed a deep neural network LSTM with RRWMD effectively. On the training dataset, our model achieves 92% accuracy, 91% recall, 85% F1-Score, and 86.5% precision. On the test dataset, the model achieves 82.6% accuracy, 82.5% recall, 82% F1-Score, and 81% precision. The accuracy obtained by the LSTM, in the bipolar sentiment prediction is 88.75%, in the fine-grained sentiment analysis is 71.51%, and in the sarcasm detection is 75.76%. The analysis done on finely-grained sentiments helped the users to gain more details by comparing the movie reviews. The challenging sarcastic reviews of the movies are detected to obtain the exact intention of the users which helps in updating future movies. The RRWMD semantic similarity measures have achieved better performance which has been evaluated by the different metrics such as accuracy, precision, recall and F1 score. The SVM machine learning classification algorithm gives good results for intent emotion classification whose predictions provide detailed and accurate information about the movies before spending the time and money, which is in the form of sentiment analysis which may not satisfy the informational needs. The social interaction between different kinds of people will be more useful for the organization to satisfy economic benefits. There is still more possibility to improve the performance by changing the learning rates, increasing epochs, using extra features, enriching embeddings, and removing misspellings. Aside from this, the etymological and vocabulary-based split of feature extraction can be taken rather than TF-IDF features which are usedas scores. For further improvement, we can use hierarchical ensemble learning techniques or multi-class deep learning models to get better accuracy.

References

[1] Adetunji, O., Hadiza, M., & Otuneme, N. (2020). Design of a Movie Review Rating Prediction (MR2P) Algorithm. *International Journal of Scientific Research in Computer Science, Engineering and Information Technology, 6*(4), 423-432.

[2] Araújo, M., Pereira, A., & Benevenuto, F. (2020). A comparative study of machine translation for multilingual sentence-level sentiment analysis. *Information Sciences, 512*, 1078-1102.

[3] Bavakhani, M., Yari, A., & Sharifi, A. (2019). A deep learning approach for extracting polarity from customers' reviews. *In IEEE 5th International Conference on Web Research*, 276-280

[4] Bhavsar, H., & Manglani, R. (2019). Sentiment analysis of Twitter data using Python. *International Research Journal of Engineering and Technology (IRJET), 6*(3), 510-511.

[5] Bouazizi, M., & Ohtsuki, T. (2019). Multi-class sentiment analysis on twitter: Classification performance and challenges. *Big Data Mining and Analytics, 2*(3), 181-194.

[6] Chatterjee, A., Gupta, U., Chinnakotla, M.K., Srikanth, R., Galley, M., & Agrawal, P. (2019). Understanding emotions in text using deep learning and big data. *Computers in Human Behavior, 93*, 309-317.

[7] Erkan, G., & Radev, D.R. (2004). Lexrank: Graph-based lexical centrality as salience in text summarization. *Journal of artificial intelligence research, 22*, 457-479.

[8] Faisal Muhammad Shah, Abdus Sayef Riyadh, AsifImtiazShaafi, Sifat Ahmed & Fatima Tabsun Sithil. (2019), *5th International Conference on Advances in Electrical Engineering (ICAEE).*

[9] Gan, C., Wang, L., Zhang, Z., & Wang, Z. (2020). Sparse attention based separable dilated convolutional neural network for targeted sentiment analysis. *Knowledge-Based Systems, 188*.

[10] Ireland, R., & Liu, A. (2018). Application of data analytics for product design: Sentiment analysis of online product reviews. *CIRP Journal of Manufacturing Science and Technology, 23*, 128-144.

[11] Johnson, C., Khadka, B., Basnet, R.B., & Doleck, T. (2020). Towards Detecting and Classifying Malicious URLs Using Deep Learning. *Journal of Wireless Mobile Networks, Ubiquitous Computing, and Dependable Applications*, *11*(4), 31-48.

[12] Khan, A., Gul, M.A., Uddin, M.I., Ali Shah, S.A., Ahmad, S., Al Firdausi, M.D., & Zaindin, M. (2020). Summarizing online movie reviews: a machine learning approach to big data analytics. *Scientific Programming*, *2020*, 1-14.

[13] Kumar, A., Narapareddy, V.T., Srikanth, V.A., Malapati, A., & Neti, L.B.M. (2020). Sarcasm detection using multi-head attention based bidirectional LSTM. *IEEE Access*, *8*, 6388-6397.

[14] Lin, B., Zampetti, F., Bavota, G., Di Penta, M., Lanza, M., & Oliveto, R. (2018). Sentiment analysis for software engineering: How far can we go? *In Proceedings of the 40th international conference on software engineering*, 94-104.

[15] Mars, A., &Gouider, M.S. (2017). Big data analysis to features opinions extraction of customer. *Procedia computer science*, *112*, 906-916.

[16] Mashal, S.X., &Asnani, K. (2017). Emotion intensity detection for social media data. *In IEEE International Conference on computing methodologies and communication (ICCMC)*, 155-158.

[17] Mihalcea, R., & Tarau, P. (2004). Textrank: Bringing order into text. *In Proceedings of the conference on empirical methods in natural language processing*, 404-411.

[18] Morente-Molinera, J.A., Cabrerizo, F.J., Mezei, J., Carlsson, C., & Herrera-Viedma, E. (2020). A dynamic group decision making process for high number of alternatives using hesitant Fuzzy Ontologies and sentiment analysis. *Knowledge-Based Systems*, *195*.

[19] Nizamani, Z.A., Liu, H., Chen, D.M., &Niu, Z. (2018). Automatic approval prediction for software enhancement requests. *Automated Software Engineering*, *25*, 347-381.

[20] Ranganath, K., & Sallauddin, M.D. (2017). Recent Trends in Sarcasm Detection on online social networks. *International Journal of Computer Sciences and Engineering*, *5*(10), 235-239.

[21] Riaz, S., Fatima, M., Kamran, M., &Nisar, M.W. (2019). Opinion mining on large scale data using sentiment analysis and k-means clustering. *Cluster Computing*, *22*, 7149-7164.

[22] Ruz, G.A., Henríquez, P.A., & Mascareño, A. (2020). Sentiment analysis of Twitter data during critical events through Bayesian networks classifiers. *Future Generation Computer Systems*, *106*, 92-104.

[23] Sahithya, N.M.S., Prathyusha, M., Rachana, N., Priyanka, P., & Jyothi, P.J. (2019). Deep Embedding Sentiment Analysis on Product Reviews Using Naive Bayesian Classifier, *5*(2), 858-864.

[24] Shrivastava, K., Kumar, S., & Jain, D.K. (2019). An effective approach for emotion detection in multimedia text data using sequence based convolutional neural network. *Multimedia tools and applications*, *78*, 29607-29639.

[25] Umer, Q., Liu, H., & Sultan, Y. (2018). Emotion based automated priority prediction for bug reports. *IEEE Access*, *6*, 35743-35752.

[26] Umer, Q., Liu, H., & Sultan, Y. (2019). Sentiment based approval prediction for enhancement reports. *Journal of Systems and Software*, *155*, 57-69.

[27] Werner, M., & Laber, E. (2019). Speeding up Word Mover's Distance and its variants via properties of distances between embeddings.

[28] Wikarsa, L., & Thahir, S.N. (2015). A text mining application of emotion classifications of Twitter's users using Naive Bayes method. *In IEEE 1st International Conference on Wireless and Telematics (ICWT)*, 1-6.

[29] Williams, G., & Mahmoud, A. (2017). Analyzing, classifying, and interpreting emotions in software users' tweets. *In IEEE/ACM 2nd International Workshop on Emotion Awareness in Software Engineering (SEmotion)*, 2-7.

Author's Biography

Mrs.R. Akila is a Asst. Professor in B.S. Abdur Rahman Crescent Institute of Science and Technology. She did her B.E from Manonmaniam Sundaranar University in Computer Science and Engineering. She completed her M.E. from Anna University. She is pursuing her Phd in the field of Datamining and machine learning. She has 20 years of experience in teaching.

Dr.S. Revathi is a professor in B.S. Abdur Rahman Crescent Institute of Science and Technology. Her qualifications are as mentioned. Ph.D. (Information and Communication Engineering), M.E. (Computer Science & Engg.), B.E. (Computer Science & Engg.). She has 25 years of teaching experience and her areas of interest include Machine Learning, Networking, Internet of Things Total number of research publications are 30.