Movie Sentiment Analysis using Feature Dictionary and Multiview Light Semi Supervised Convolution Neural Network

Chaitra Kulkarni, R Suchithra

Abstract: ***Emotional information in film commentary is very important for emotional analysis. An emotional analysis that focuses on classifying opinions into positive and negative classes according to an emotional glossary is a study. Most existing research focuses on word synthesis and user evaluation, while users' attitudes toward feedback are ignored. To consider this point, this paper uses an emotional analysis and in-depth learning approach to examine the relationship between online film reviews, and this point is used for movie box revenue efficiency. In this paper, this work present a 11 different types of Feature Dictionary. It is modeled with information from sentences (i.e., reviews) and aspects simultaneously. First, Feature Dictionary is created with all aspects of the sentence. After obtaining the aspects, it utilize all data in the source domain and the target domain for training Multiview Light Semi Supervised Convolution Neural Network (MLSSCNN) classifier. To understand the predictive performance of this approach several performance metrics are used. The experimental result shows that the MLSSCNN provides a superior predictive effect than other classifier.***

Keywords : Natural Language Processing (NLP), Emotion Recognition, social media analysis, Sentiment Analysis, Box office Prediction, Twitter, Feature Dictionary, Aspect Ratio and MLSSCNN.

I. INTRODUCTION

In a rapidly changing dynamic world, digital media is constantly evolving in the way that feedback from viewers is presented in different forms across different platforms. The content of this genre continues to change as the decades-old film that may have been celebrated during its release may seem outdated due to changing viewers' preferences. Therefore, directors and production companies must keep their films fresh to ensure that the audience is satisfied. Getting feedback from the audience is done primarily to identify bugs in the current version and for further improvements in the future. Feedback can be classified as positive, negative, or neutral. Each sentence needs to be evaluated and may need to be broken down into smaller sections. Emotional analysis based on aspects adhere to what is in the sentence to achieve high accuracy [1]. When comments are taken from social media, downloading aspects is necessary because the feedback received from social media is not official and does not have a specific structure to follow. Thus, comments can be short or long and it is more difficult to analyze when the review is longer. Therefore, any analysis of data taken from social media must rely on an excerpt of the view and then a classification of opinions [2]. Sensitivity analysis is classified as a Natural Language Process (NLP) and is also known as a survey that extracts and refines data and identifies concepts from data extracted from the Internet [3]. Machine learning methods are used to classify ideas because traits and tastes are difficult to identify manually. Machine learning algorithms can easily classify characteristics, and they are very useful when there is a lot of data [4]. The main function of retrieving feedback is to retrieve the target feedback. The three main and essential functions that must be performed through the analysis of emotions are to extract the objective of the mind, to discover the type of aspect, and to form a polar line of emotion [5].

Emotional analysis methods based on the emotional dictionary rely primarily on open source emotional dictionaries or emotional expansion dictionaries [6], and the weight of emotions can be calculated in combination with a rule. The Chinese Open Source Emotional Dictionary includes HowNet [7], a Chinese dictionary database of sentences by the Dalian University of Technology [8] and NTUSD of the National University of Taiwan [9]. The English dictionary of the open source server contains WordNet [10]. Yi et al. Introduce emotional analysts (SA) who cite emotions (or comments) on topics from online text documents. To alleviate the shortcomings of purely statistical methods, they presented an analysis of the grammatical structure of sentences and phrases based on LP techniques. PANG et al. [11] Proposed methods for classifying texts based on the trend of emotional trends. Sun et al has developed a method for classifying emotions based on conditional fields and a dictionary of emotions. LUO et al. [13] Develop a vocabulary of two levels of emotion, and words for different levels will be improved differently. Methods based on the dictionary of sentences can be highly accurate for data sets that have real emotional tendencies. However, it would be very inefficient for a large collection of data in the trend of emotions. In addition, it does not work well in fields with specific words in the field, as it is almost impossible to create a dictionary that is applicable to all fields. In the past, it was difficult to identify because feedback had to be collected directly from the sample of the target population. The target population of the sample had to be identified, and the review requested an increase in artificial intelligence changes that allowed for more available reviews and easier data collection. Therefore, emotional analysis is the right time and it can be performed in different areas.

Revised Manuscript Received on October 10, 2020.

* Correspondence Author
Chaitra Kulkarni*, Jain University, Bangalore , India.
Dr. R Suchithra, Jain University, Bangalore , India

For the analysis of emotions, for high accuracy of identification, optimal hybrid models can be used. Semi-management methods in which training is gradually implemented have been studied by many researchers and there are still opportunities for research in this field.

The reminder of this paper is organized as below:

• In section 2, Literature Review of previous research works.

• In section 3, Methodology, system work flow, Data cleaning, Feature extraction, Data integration and transformation.

• In section 4, Analysis and Result discussion, sentiment analysis, performance analysis.

• In section 5, Conclusion and Discussion.

• In section 6, References.

II. RELATED WORK

Supervisors usually provide a summary in their review. So analyzing the emotions of watching a movie has been a difficult task since then. SS problems can be considered as a general classification task when examining into a positive or negative classification. Emotional analysis based on machine learning methods for IMDb examination was proposed by Pang et al., (2002) [14]. Classification of reviews based on common sense was made with Naïve Bayes, Maximum Entropy and Support Vector Machines (SVM) classifiers. The presentation of the feature is done using non-capitalized labels, the location of words and parts of speech (POS).

A prominent challenge in the analysis of emotions is the handling of performance data over time. Methods of emotional analysis by running data have been proposed in the work of Chamansingh et al. , (2016) [15]. The authors are advised to use machine learning tools such as Support Vector Machines (SVM) and Maximum Entropy for Twitter run time data. The work shows that significant reductions in data entry and processing can be achieved by maintaining the right level of usable data. Preparation of text data in advance is considered as an important method to improve ranking accuracy. Methods of extraction and function selection can be used to improve classification accuracy and reduce overall complexity. The work proposed by Ghosh et al. , (2017) [16] stressed the importance of pre-processing using text camera reviews for emotional analysis. Data processing is performed by marking, stopping, deleting and generating original words. The classifications for SVM, Max Entropy and Naive Bayes were used for classification. In this work, SVM gets the best results among them all. Chakrit et al. (2016) [17] informed of the importance of reducing the dimensions of the data during the previous process. The method of selecting the characteristics of the square is used to select the function and use of machine learning, combined with voice, gives better results compared to the existing method of machine learning. The use of hybrid methods in the selection of subset of good character was proposed by Yang et al, (2015) in [18].

Glossary of vocabulary and grammar is used, which is highly informative. This model was trained with six different classical classifiers, in which the Naïve Bayes Multinomial (NBM) proved to be very accurate. The following references highlight the importance of data disclosure or review. Tripathy et al. , (2016) [3] Describes the use of M-gram representations for characteristics. Word matrix (TDM) files are created using the letters n and g and use various technologies such as Naive Bayes, Maximum Entropy, Support Vector Machine (SVM) and Stochastic Gradient Descent to train models. The model showed the best results for SVM by combining non-large presentation presentation and credit. Sahu et al. , (2016) [19] focused their task on the IMDB database to classify film review skills and implement download and re-ranking functions. In addition to the usual word representation, 10 additional functions were selected based on the bias of the word in the review. Some of the unusual features used are positive and negative words for adaptation, combined with positive and negative adjectives of the spirit with repetitive letters. Information retrieval was used to select features from the N-gram presentation of the review. Selection of suitability for higher rating accuracy was proposed by Trivedi et al. , (2016) in [20]. The main purpose of the work is to analyze the emotions of the review of Indian films. Different methods of selecting functions, such as a gigabyte, thanks to the benefits and advantages, give them a good F value and a false positive reading. To reduce the computational complexity associated with quoting features, Go et al. , (2015) [21] used G-CT and PMI to select functions. Experiments were performed on two different corporations for SA-microblogs and e-commerce data to evaluate performance and emphasize the importance of functional selection for best results. In addition to the simple feature selection method contained in the literature, Tuba et al. , (2016) [22] proposed a new feature selection method called uery Expansion Ranking. This method showed higher ranking accuracy for three different Turkish control datasets compared to the different square and frequency methods of the file. Regularized Locality Preserving Indexing (RLPI) was proposed by Cai et al. , (2007) [24] An improved approach to basic indexing (LPI) for better presentation of text documents. Destroying file presentations using Eigen vector destruction and at least quadratic problems to select top vectors to represent file spaces makes RLPI more efficient and manages large matrices. Similar work has been proposed by Harish et al. , (2016) [23] where RLPI, a large Term Document Matrix (TDM) function selection technique, is used for text clipping. The results are very encouraging. Using reliable methods to select emotional analysis features and reduce complexity is an open challenge. Researchers are experimenting with different combinations of feature selection methods. This emphasizes the importance of functional selection in SA

III. PROPOSED APPROACH

This section illustrates the procedure of data collection, processing, feature retrieving schemes for presenting data sets, prediction schemes, and combining knowledge gathering methods employed in experimental examination. The sketch out of the work is depicted in Fig. 1. It consists of the following four modules.

1. Gathering of Dataset
2. Data Cleaning
3. Feature Dictionary Creation
4. Feature Aspect Ratio Calculation and Selection
5. Sentiment Analysis Model Generation

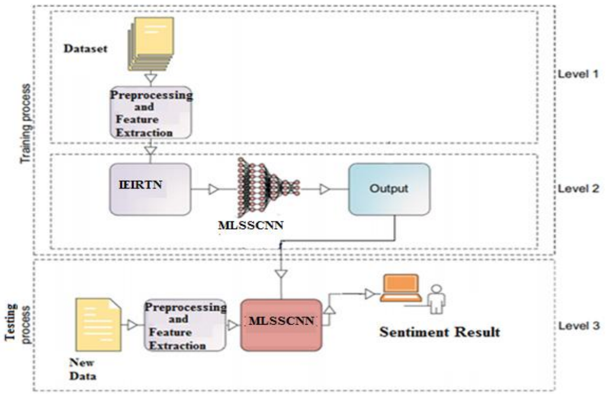


Fig. 1.Sketch out of the Proposed Work

A. Gathering of Dataset

To evaluate the predictable presentation of the psychological along with linguistic characteristics of emotional examination, this work examined the sequence of various sources such as twitter, facebook, website. It contain positive, negative along with neutral emotions. For database compiling process, this work used the structure summarized in [11]. This work employed Twitter4J. It is used to gather tweets. All tweet is captioned as either positive, negative, or neutral. Later than gathering tweets, automatic filtering is used to eliminate unsuitable and needless tweets. As a result this work get a collection of 6188 negatives, 4891 positives as well as 4,252 tweets. To get a unbiased body, dataset of our work contains 4,200 positive along with negative tweets as well as 4,200 positive tweets..

B. Data Cleaning

Due to the inaccurate and informal nature of pure Twitter posts, pre-creation is necessary to eliminate some issues (such as unnecessary duplication of words and use of incorrect text) [12]. At the pre-treatment stage, we adopted the framework presented in [32]. The pre-processing phase is primarily aimed at eliminating unnecessary symbols or sequences that are not valuable for emotional separation. To this end, the following tasks are performed for each tweet in this context [32]:

- Exclude mentions and replies from other users' tweets placed by a string starting with "@".
- Eliminate URLs (like sequences with "http: //").
- Clear "#" character

Once the database is run, the next step is to create a matrix of functions. Prior to retrieval, the token function is first applied to the pre-processed data. Tokenization is the process of dividing sentences into words. Then take a break. The word stop is commonly used and it is common that they lose their main meaning. Words like "on, are, that's, are" are just a few examples of words that stop.

There are two common drawl methods and both methods work equally well. Both are obvious. One way is to count all the words of the event and set a numeric value for each word / number of words and get rid of words / phrases that occur in excess of the specified value. Another way is to have a list of predefined word breaks that can create a list of short breaks that can be removed from the token / token list. In our work, we have implemented both methods for removing stop words.

This end is followed by the disappearance process continues. The purpose of the original was to reduce the reflection form and sometimes to associate the derivative of the word with a common basic form. This program is an easy way to truncate some characters at the end of a word to their roots, using them just set the principle for truncating some characters at the end of a word and hopefully they will get good results.

C. Feature Dictionary Creation

After the database has been processed, the next step is to create a feature dictionary. In this eleven different types of dictionaries are created.

Positive Word Dictionary

The dictionary of positive words consists of two parts: positive proverbs and rhetorical questions. It has been stated that when positive proverbs and rhetoric change the words of emotion, they will change the polarity of the emotion, but the rhetoric is stronger and the double positive will not change the polarity of the feeling of the words, but the tone will be louder.

Negative Word Dictionary

The dictionary of negative words consists of two parts, negative adverbs and rhetorical questions. It has been stated that when negative proverbs and rhetoric change the words of emotion, they will change the polarity of the emotion, but the rhetoric is stronger and the double negligence will not change the emotional smoothness of the words, but the tone will be louder.

Degree Adverb Dictionary

The Dialect Degree Dictionary is from the HowNet Dictionary Library. The term is divided into six levels. Levels are amazing, mostly a bit related and time consuming. Each of these six levels acquires a certain weight and the emotional intensity of the changed word expands by a multiple.

Positive Sentiment Tri-grams Dictionary

Three consecutive words with a positive emotional rating, according to SentiWordNet. For example, "interesting and fun and Intriguing."

Negative Sentiment Tri-grams Dictionary

All three consecutive words have negative emotions, according to SentiWordNet. For example, "gloomy and boring and uninteresting."

Positive Sentiment Bi-grams Dictionary

Two consecutive words, both of which have a positive emotional rating, according to SentiWordNet. For example, "interesting and fun".

Negative Sentiment Bi-grams Dictionary

Two consecutive words, both of which have a negative emotional rating, according to SentiWordNet. For example, "gloomy and boring."

Positive Sentiment words coupled with Adjective Dictionary

Words with a positive mood, preceded by an adjective. Example, "Great well-written movie."

Negative Sentiment words coupled with Adjective Dictionary

Negative-feeling words precede adjectives. For example, "Boring movie with weak script".

Positive Sentiment words with repeated letters Dictionary

Positive emotional words with repetitive words in letters. Example, Awwwesome.

Negative Sentiment words with repeated letters Dictionary

Negative emotional words with repetitive words in letters. Example, Awwwfull.

Feature Aspect Ratio Calculation

With the analysis of the observation characteristics, we need to find the effect of each character on the bias of the document to determine the scaling factor for each characteristic. To find out the impact, this document uses Point Mutual Information (PMI). PMI is used to calculate the similarity between new words and seed words, and finally to calculate the emotional clarity of unknown new words. Mutual information points are important for calculating word-for-word similarities. The formula for calculating the similarity between the new unknown word w1 and the seed word w2 is:

$$PMI(w_1, w_2) = \log_2 \frac{p(w_1, w_2)}{p(w_1) \times p(w_2)}$$

Where $p(w_1, w_2)$ is the probability that w1, w2 will co-occur and $p(w_1)$, $p(w_2)$ respectively represent the probability that w1, w2 will appear alone.

D. Sentiment Examination Model Generation

To build a high-efficiency Sentiment Analysis model by Multiview Light Semi Supervised Convolution Neural Network (MVLSSCNN). In this stage, the extracted feature matrices are analyzed for being positive, neutral or negative to calculate the overall polarity of tweets. Multiview Sentiment Analysis (MSA) is used for this purpose. It consists of three important stages such as Sentiment analysis, Reverse sentiment analysis and Final Sentiment Analysis.

Sentiment Examination

Sentiment examination of tweet is performed using Light Semi Supervised Convolution Neural Network (LSSCNN). The test results categorize tweets as positive, neutral or negative and also give them a measure of emotional confidence. Emotional confidence just shows positive, neutral or negative. Emotional evaluation is further used to calculate the final emotions of a tweet.

Reverse Sentiment Examination

MDSA not only considers how positive / neutral / negative the original tweet is, but also how negative / neutral / positive the reverse tweet is. This solves the problem of polar deformation. Therefore, the first tweet must be inverted for any part of the tweet speech to be identified. Adjectives are extracted from lexical databases such as WordNet. Reverse the mood there are the following steps

- Each tweet is checked for negative words such as "No".
- Negative words are removed and the words immediately after them are not changed.
- Other words do not change.

The steps outlined above provide reverse tweaking. The feeling of changing tweets was also obtained with the help of the same LSSDCNN emotional analyst who provided his feelings and confidence.

Final sentiment examination

Taking into account both the original and inverse of Twitter, its final sentiment is calculated. The positive of the original tweet is considered to be the negative of the inverse tweet to calculate the final positive of the tweet. Similarly, the negative of the original tweet is considered with the positive of the inverse tweet to determine the negative of the tweet. Assume $p(+|x)$ as the probability that Tweet x is positive and $p(-|\bar{x})$ the probability that the reverse side of the tweet x is negative. Similarly, assume $p(-|x)$ as he probability that the twitter x is negative and $p(+|\bar{x})$ the probability that the reverse side of the twitter x is positive. The relationship between these probabilities is then given by:

$$p(+|x, \bar{x}) = (1-\alpha).\ p(+|x) + \alpha\ .\ p\ (-|\bar{x})$$

$$p(-|x, \bar{x}) = (1-\alpha).\ p\ (-|x) + \alpha\ .\ p\ (+|\bar{x})$$

Light Semi Supervised Convolution Neural Network (LSSCNN)

The structure of Convolutional Neural Networks are similar to the structure of conventional neural networks. Each neural network is composed of neurons with training weights, biased values that are of studying weight along with bias. Generally a neuron obtains input, computes its point product moreover the behavior follows a non-linear characteristic. Each Convolutional Neuron Network is composed of single or additional convolutional layers as well as pooling layers or sub sampling layers. and this is depicted in Fig 2.

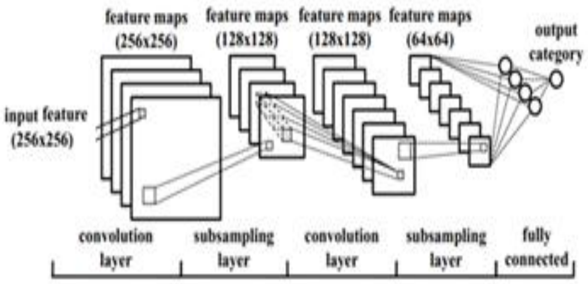


Fig.2. Sentiment Analyzing using CNN

The two types of pooling layers of CNN are Maximum and Average pooling in CNN and when united the biggest pixel values are calculated in max whereas in average pooling, the mean values are in use to explanation. The response of this layer is served as input to the subsequent layer of convolutional neural network. Generally convolutional neuron network is employed to sort the types of cancer cells and its severity.

Down sampling is accomplished by the employment of pooling layer of CNN. The time consumption is lessened by diminishing the features extraction in convolution layers. To overcome this drawback of CNN, Light Semi Supervised Convolution Neural Network (LSSCNN) is proposed. This reduces high computation cost and improves speed. The dimension reduction of image space is realized by vector of features that is created by PSO from multidimensional image space to low dimensional feature space. Fig 3 shows the sketch of LSSCNN.

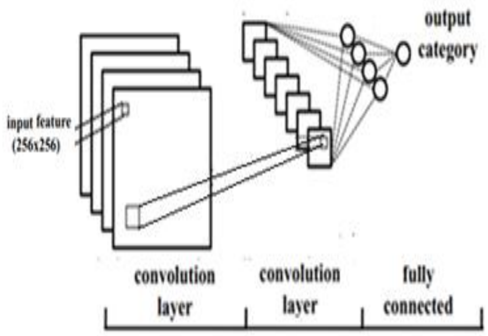


Fig.3. Sentiment Examination through LSSCNN

IV. RESULT AND EXAMINATION

A. Examination Parameters

To examine the effectiveness of the sentiment identification methods, a number of examination parameters are available. This work considers the Detection Accuracy, Precision Rate, Recall Rate, Sensitivity, Specificity, F-Measure and Error Rate to examine the effectiveness.
Detection Accuracy

Detection Accuracy metric finds the percentage of truthiness between the original sentiments and the predicted sentiments.

$$\text{Accuracy} = \frac{TP+TN}{TP+FP+TN+FN}$$

Error Rate

Error Rate finds the percentage of falseness between the original sentiments and the predicted sentiments.

$$\text{Error Rate} = \frac{\text{No of Images of Falsely predicted sentiments}}{\text{Total No of statments}}$$

Precision Rate

The precision value is found by using the below formula

$$\text{Precision} = \frac{TP}{TP+FP}$$

Recall Rate

The recall is found by using the below formula

$$\text{Recall} = \frac{TP}{TP+FN}$$

Sensitivity

Sensitivity is found by using the below formula

$$\text{Sensitivity} = \frac{TP}{(TP+FN)}$$

Specificity

Specificity is found by using the below formula

$$\text{Specificity} = \frac{TN}{(FP+TN)}$$

F-Measure

F-measure is found by using the below formula

$$F_m = (1+\alpha) * \frac{\text{Precision} * \text{Recall}}{\alpha * (\text{Precision} * \text{Recall})}$$

B. Units

In this section the performance of the proposed method is evaluated in various experiments. To evaluated the efficiency of this sentiment examination scheme, the Detection Accuracy, Precision Rate, Recall Rate, Sensitivity, Specificity, F-Measure and Error Rate measures are used.

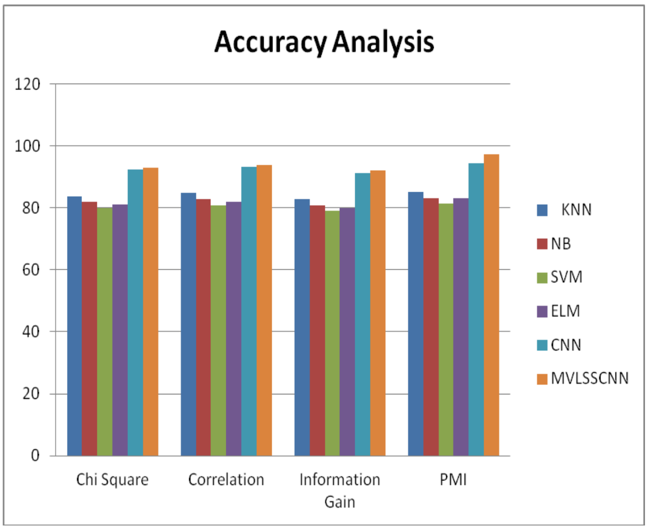
Trial No 1: Examination of Sentiment Examination Approaches using Accuracy

To examine the performance of this sentiment examination scheme, it is compared with different techniques using the operating indicators which are mentioned in Section 4.1. The output of these indicators are tabulated in Table 1.

Table 1: Examination of Sentiment Examination Techniques using Accuracy

Sentiment Examination Approaches						
Examination Parameters	KNN	NB	SVM	ELM	CNN	MVLSSCNN
Chi Square	83.83	81.93	80.03	81.06	92.32	93
Correlation	84.74	82.84	80.94	81.97	93.23	93.91
Information Gain	82.82	80.92	79.02	80.05	91.31	91.99
PMI	85.19	83.21	81.46	83.11	94.31	97.31

The influence of Sentiment Examination Techniques that are employed in this test using accuracy are successfully assessed. Table 1 illustrates that the highest accuracy value is 97.23 is found for MVLSSCNN and it is more powerful as it is the highest value when compared with other techniques. The output of these indicators are tabulated in Fig.8.



The influence of Sentiment Examination Techniques that are employed in this test using accuracy are successfully assessed. Fig.8 illustrates that the highest accuracy value is 97.23 is found for MVLSSCNN and it is more powerful as it is the highest value when compared with other techniques.

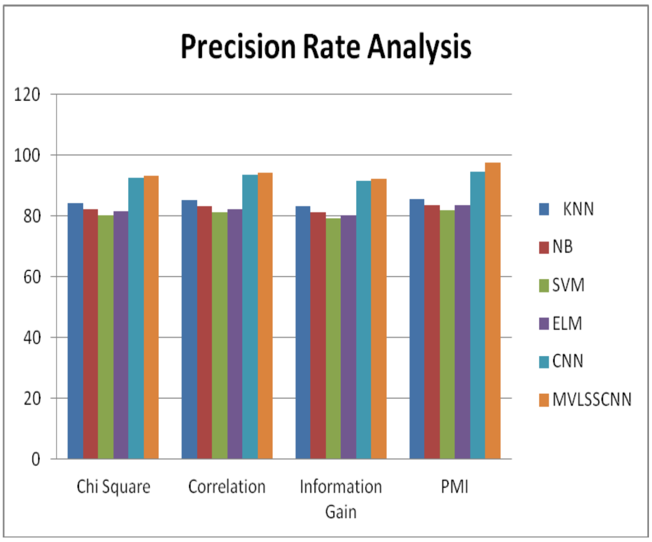
Trail No 2: Examination of Sentiment Examination Approaches using Precision Rate

To examine the performance of this sentiment examination scheme, it is compared with different techniques using the operating indicators which are mentioned in Section 4.1. The output of these indicators are tabulated in Table 2.

Table 2: Examination of Sentiment Examination Techniques using Precision Rate

Sentiment Examination Approaches						
Examination Parameters	KNN	NB	SVM	ELM	CNN	MVLSSCNN
Chi Square	84.1	82.2	80.3	81.3	92.6	93.3
Correlation	85.0	83.1	81.2	82.2	93.5	94.21
Information Gain	83.1	81.2	79.3	80.3	91.6	92.29
PMI	85.4	83.5	81.7	83.4	94.6	97.61

The influence of Sentiment Examination Techniques that are employed in this test using precision rate are successfully assessed. Table 2 illustrates that the highest precision value is 97.54 is found for MVLSSCNN and it is more powerful as it is the highest value when compared with other techniques. The output of these indicators are tabulated in Fig.9.



The influence of Sentiment Examination Techniques that are employed in this test using precision rate are successfully assessed. Fig.9 illustrates that the highest precision value is 97.54 is found for MVLSSCNN and it is more powerful as it is the highest value when compared with other techniques. The output of these indicators are tabulated in Fig.9.

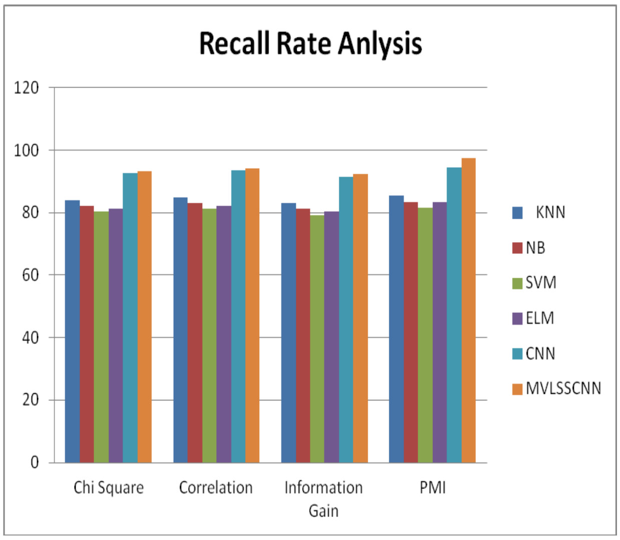
Trial No 3 : Examination of Sentiment Examination Approaches using Recall Rate

To examine the performance of this sentiment examination scheme, it is compared with different techniques using the operating indicators which are mentioned in Section 4.1. The output of these indicators are tabulated in Table 3.

Table 3: Examination of Sentiment Examination Techniques using Recall Rate

Sentiment Examination Approaches						
Examination Parameters	KNN	NB	SVM	ELM	CNN	MVLSSCNN
Chi Square	84.05	82.15	80.25	81.28	92.54	93.22
Correlation	84.96	83.06	81.16	82.19	93.45	94.13
Information Gain	83.04	81.14	79.24	80.27	91.53	92.21
PMI	85.41	83.43	81.68	83.33	94.53	97.53

The influence of Sentiment Examination Techniques that are employed in this test using recall rate are successfully assessed. Table 3 illustrates that the highest recall value is 97.45 is found for MVLSSCNN and it is more powerful as it is the highest value when compared with other techniques. The output of these indicators are tabulated in Fig.10.



The influence of Sentiment Examination Techniques that are employed in this test using recall rate are successfully assessed. Fig.10 illustrates that the highest recall value is 97.45 is found for MVLSSCNN and it is more powerful as it is the highest value when compared with other techniques.

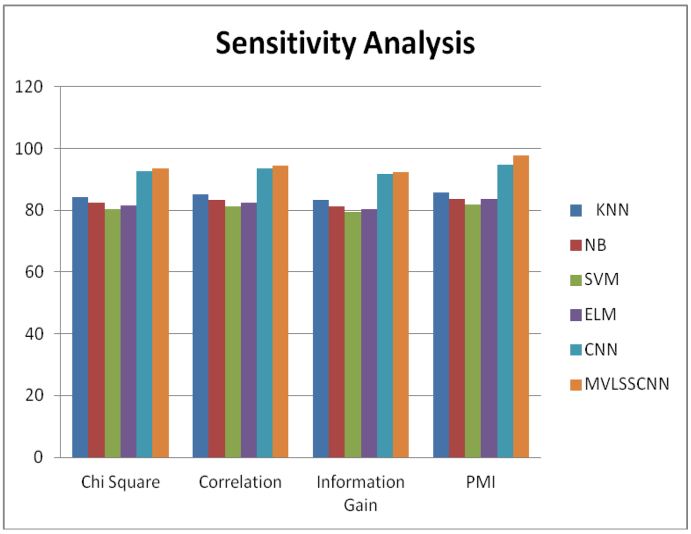
Trial No 4: Examination of Sentiment Examination Approaches using Sensitivity

To examine the performance of this sentiment examination scheme, it is compared with different techniques using the operating indicators which are mentioned in Section 4.1. The output of these indicators are tabulated in Table 4.

Table 4: Examination of Sentiment Examination Techniques using Sensitivity

Sentiment Examination Approaches						
Examination Parameters	KNN	NB	SVM	ELM	CNN	MVLSSCNN
Chi Square	84.23	82.33	80.43	81.46	92.72	93.4
Correlation	85.14	83.24	81.34	82.37	93.63	94.31
Information Gain	83.22	81.32	79.42	80.45	91.71	92.39
PMI	85.59	83.61	\81.86	83.51	94.71	97.71

The influence of Sentiment Examination Techniques that are employed in this test using sensitivity are successfully assessed. Table 4 illustrates that the highest sensitivity value is 97.64 is found for MVLSSCNN and it is more powerful as it is the highest value when compared with other techniques. The output of these indicators are tabulated in Fig.11.



The influence of Sentiment Examination Techniques that are employed in this test using sensitivity are successfully assessed. Fig.11 illustrates that the highest sensitivity value is 97.64 is found for MVLSSCNN and it is more powerful as it is the highest value when compared with other techniques.

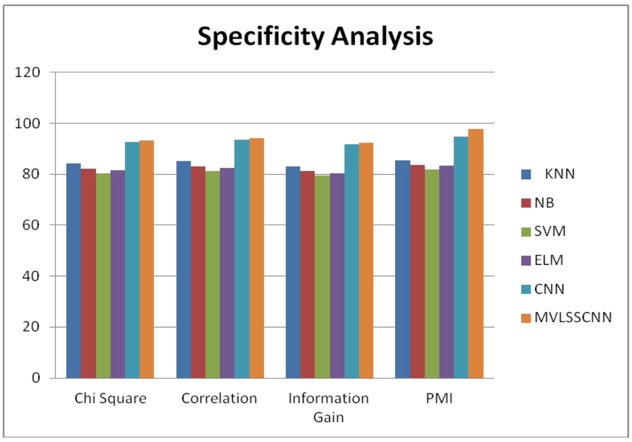
Trial No 5: Examination of Sentiment Examination Approaches using Specificity

To examine the performance of this sentiment examination scheme, it is compared with different techniques using the operating indicators which are mentioned in Section 4.1. The output of these indicators are tabulated in Table 5.

Table 5: Examination of Sentiment Examination Techniques using Specificity

Sentiment Examination Approaches						
Examination Parameters	KNN	NB	SVM	ELM	CNN	MVLSSCNN
Chi Square	84.16	82.26	80.36	81.39	92.65	93.33
Correlation	85.07	83.17	81.27	82.3	93.56	94.24
Information Gain	83.15	81.25	79.35	80.38	91.64	92.32
PMI	85.52	83.54	81.79	83.44	94.64	97.64

The influence of Sentiment Examination Techniques that are employed in this test using specificity are successfully assessed. Table 5 illustrates that the highest specificity value is 97.56 is found for MVLSSCNN and it is more powerful as it is the highest value when compared with other techniques. The output of these indicators are tabulated in Fig.12.



The influence of Sentiment Examination Techniques that are employed in this test using specificity are successfully assessed. Table 5 illustrates that the highest specificity value is 97.56 is found for MVLSSCNN and it is more powerful as it is the highest value when compared with other techniques.

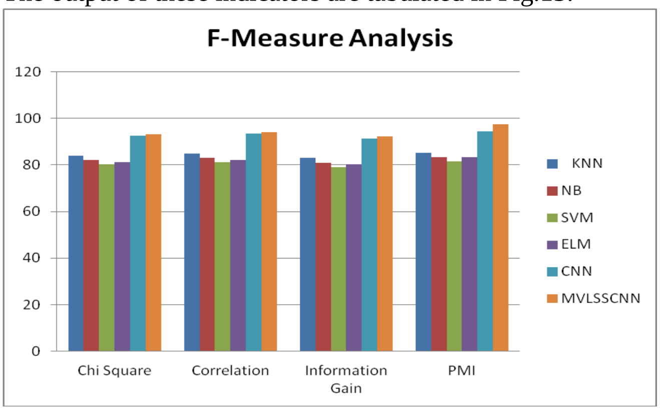
Trial No 5 : Examination of Sentiment Examination Approaches using F-Measure

To examine the performance of this sentiment examination scheme, it is compared with different techniques using the operating indicators which are mentioned in Section 4.1. The output of these indicators are tabulated in Table 6.

Table 6: Examination of Sentiment Examination Techniques using F-Measure

Sentiment Examination Approaches						
Examination Parameters	KNN	NB	SVM	ELM	CNN	MVLSSCNN
Chi Square	83.96	82.06	80.16	81.19	92.45	93.13
Correlation	84.87	82.97	81.07	82.1	93.36	94.04
Information Gain	82.95	81.05	79.15	80.18	91.44	92.12
PMI	85.32	83.34	81.59	83.24	94.44	97.44

The influence of Sentiment Examination Techniques that are employed in this test using f-measure are successfully assessed. Table 6 illustrates that the highest f-measure value is 97.38 is found for MVLSSCNN and it is more powerful as it is the highest value when compared with other techniques. The output of these indicators are tabulated in Fig.13.



The influence of Sentiment Examination Techniques that are employed in this test using f-measure are successfully assessed. Table 6 illustrates that the highest f-measure value is 97.38 is found for MVLSSCNN and it is more powerful as it is the highest value when compared with other techniques

V. CONCLUSION

Dictionary and MVLSSCNN approach can improve the accuracy of the feeling of film review and can be better applied to the evaluation and recommendation of film preferences and play a leading role in promoting the film on the media platform. The four methods based on the dictionary of emotions proposed in this document have a better effect on the analysis of emotions than the methods based on the basic dictionary of emotions. Experimental analysis shows that the considering several feature dictionary and aspects are more than individual. For movie sentiment analysis, the highest predictive performance (97.44%) was achieved by combining the language and mental processes of MVLSSCNN. Therefore, this proposed method and combination of sentence and aspects performs best in emotion analysis of movie.

REFERENCES

1. W. Xue and T. Li, "Aspect Based Sentiment Analysis with Gated Convolutional Networks," Florida International University, 2018.
2. Q. Liu, H. Zhang, Y. Zeng, Z. Huang, and Z. Wu, "Content Attention Model for Aspect Based Sentiment Analysis," in Track: Web Content Analysis, Semantics and Knowledge, 2018, pp. 1023–1032, doi: 10.1145/3178876.3186001.
3. E. Cambria, B. Schuller, Y. Xia, and C. Havasi, "New avenues in opinion mining and sentiment analysis," 2013.
4. A. P. Jain and P. Dandannavar, "Application of machine learning techniques to sentiment analysis," in 2016 2nd International Conference on Applied and Theoretical Computing and Communication Technology (iCATccT), 2016, pp. 628– 632, doi: 10.1109/ICATCCT.2016.7912076.
5. M. Pontiki et al., "SemEval-2016 Task 5: Aspect Based Sentiment Analysis," in SemEval-2016 Task 5: Aspect Based Sentiment Analysis, 2016, pp. 342–349, doi: 10.18653/v1/s16-1055.
6. W. Du, S. Tan, X. Cheng, et al., Adapting information bottleneck method for automatic construction of domain-oriented sentiment lexicon, ACM Int. Conf. Web Search Data Min. (2010) 111–120.
7. Y.Y. Zhao, B. Qin, T. Liu, Sentiment analysis, J. Softw. 21(08) (2010)1834–1848.
8. J.M. Chen, H.F. Lin, Z.H. Yang, Automatic acquisition of sentiment vocabulary based on syntax, CAAI Trans. Intell. Syst. 4 (2) (2009) 100–106.
9. A.P. Li, P. Di, G.L. Duan, Document sentiment orientation analysis based on sentence weighted algorithm, J. Chin. Comput. Syst. 36 (10) (2015)2252–2256.
10. Y. Lu, M. Castellanos, U. Dayal, Automatic construction of a context-aware sentiment lexicon: An optimization approach. International Conference on World Wide Web, 2011, pp. 347-356.
11. J. Yi, T. Nasukawa, R. Bunescu, et al., Sentiment analyzer: Extracting sentiments about a given topic using natural language processing techniques, IEEE Int. Conf. Data Min. (2003)427–434.
12. PANG Zhen-jun, GAO Li-bo, YAO Tian-fang. Web text tendency classification based on sentiment phrase [C] Chinese Opinion Analysis Evaluation (COAE 2014). Kunming: CIPS, 2014: 179–186.(in Chinese).
13. LUO Yi, LI Li, TAN Song-bo, CHEN Xue-qi. Sentiment analysis on Chinese micro-Blog corpus [C] Chinese Opinion Analysis Evaluation (COAE 2014). Kunming: CIPS, 2014: 123–130. (in Chinese).
14. B. Pang, L. Lee, and S. Vaithyanathan, "Thumbs up? Sentiment classification using machine learning techniques," in Proc. Conference on Empirical Methods in Natural Language Processing, Philadelphia, July 2002, pp. 79-86.
15. N. Chamansingh and P. Hosein, "Efficient sentiment classification of twitter feeds," in Proc. IEEE International Conference on Knowledge Engineering and Applications, 2016.

16. M. Ghosh and G. Sanyal, “Preprocessing and feature selection approach for efficient sentiment analysis on product reviews,” in Advances in Intelligent Systems and Computing, S. Satapathy, V. Bhateja, S. Udgata, and P. Pattnaik Eds., Springer, 2017, vol. 515.
17. C. Pong-Inwong and K. Kaewmak, “Improved sentiment analysis for teaching evaluation using feature selection and voting ensemble learning integration,” in Proc. 2nd IEEE International Conference on Computer and Communications, 2016.
18. A. Yang, J. Zhang, L. Pan, and Y. Xiang. “Enhanced twitter sentiment analysis by using feature selection and combination,” in Proc. International Symposium on Security and Privacy in Social Networks and Big Data, 2015.
19. T. P. Sahu and S. Ahuja, “Sentiment analysis of movie reviews: A study on feature selection & classification algorithms,” in Proc. International Conference on Microelectronics, Computing and Communications, 2016.
20. S. K. Trivedi and A. Tripathi, “Sentiment analysis of Indian movie review with various feature selection techniques,” in Proc. IEEE International Conference on Advances in Computer Applications, 2016.
21. K. Gao, S. Su, and J. Wang, “A sentiment analysis hybrid approach for microblogging and e-commerce corpus,” in Proc. 7th International Conference on Modelling, Identification and Control, Sousse, Tunisia, December 18-20, 2015.
22. T. Parlar and S. Ayşe Özel, “A new feature selection method for sentiment analysis of Turkish reviews,” in Proc. International Symposium on Innovations in Intelligent Systems and Applications, 2016.
23. B. S. Harish, M. B. Revanasiddappa, and S. V. A. Kumar, “A modified support vector clustering method for document ategorization,” in Proc. IEEE International Conference on Knowledge Engineering and Applications, 2016.
24. D. Cai, X. He, W. V. Zhang, and J. Han, “Regularized locality preserving indexing via spectral regression,” in Proc. Sixteenth ACM Conference on Conference on Information and Knowledge Management, 2007, pp. 741–750.

AUTHORS PROFILE

Chaitra Kulkarni is a research scholar in Jain University ,Bangalore . He has published more than 8 papers in national and international journals . His area of specialisations are Cloud Resource management and Software engineering . He is a java programmer and give seminar in various institutions on cloud computing.

R.Suchithra is the director of MCA department in Jain Deemed to be University,Bangalore . She has 16 years of experience in teaching and research .She has published more than 24 papers in reputed journals and her area of specialization is on cloud computing , machine learning and data science . She is guiding research scholars in the PhD programme

.