Automatic Intelligent Movie Sentiment Analysis Model Creation for Box Office Prediction using Multiview Light Semi Supervised Convolution Neural Network

Chaitra Kulkarni, Manoj Manuja, R Suchithra

*Abstract***:** *With the rapid growth of e-commerce, online product and service monitoring is becoming more and more established as an important source of information for both sellers and customers. Emotional surveys and comments for online review analysis are gaining more and more attention as such studies help use information from online reviews for potential economic impacts. Twitter is a widely used social networking site and a trusted source of public opinion. The success of the film can be predicted by analyzing the tweets and researching the impact of the film. This report discusses the application of emotional analysis and in-depth machine learning methods to understand the relationship between online movie reviews, and this story is used to generate revenue at the movie box office. In this paper, this work present a Intelligent Extensive Information Rich Transfer Network (IEIRTN). It is modeled with information from sentences (i.e., reviews) and aspects simultaneously. First, IEIRTN extract 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 offers a superior predictive effect than other classifier.*

Keywords **:** *Twitter, Emotion Analysis, Sentiment examination, Box office Prediction, IEIRTN, MLSSCNN.*

I. INTRODUCTION

Emotional analysis is part of the research of the analysis of thoughts, feelings, assessments, behaviors and emotions from written language. Emotion analysis systems are used in almost every field because feedback is necessary for almost all human activities. They are a major influence of our behavior. Emotional analysis uses natural language processes and text analysis to identify and extract information about areas of interest. Due to the popularity of social media such as blogs and social networking sites such as Facebook, Twitter and others. Interest in emotional analysis has grown to a greater extent. There are many challenges in emotional

Chaitra Kulkarni*, Jain University, Bangalore , India.
Dr.Manoj Manuja, Chitkara University Rajpura, Punjab, India
Dr. R Suchithra, Jain University, Bangalore , India

analysis. The first is that words of opinion that are considered positive in one situation can be considered negative in another situation. The second challenge is that people do not always express themselves the same way. The simple text processing depends on the fact that the slight differences between the two parts of the text do not change much meaning. Emotional analysis helps to find words that express emotions and helps to understand the relationship between text reviews and the consequences of those reviews. One example is the online review of films that affect office collections. In this work, data retrieval techniques are applied to online movie reviews, which are used to predict a movie box collection based on reviews and analyze the impact of on-box reviews. To learn the sentiment of sentences from unmarked data, a classification of interpersonal emotions as a direction of success is proposed. It uses effective information in the source domain (with enough labeled data) to help classify emotions in the target domain (with or without labeled data). Because it is important to reduce the dependence on the amount of labeled data and is important for unbranded domains, a lot of attention is drawn from academia and industry. In the literature, many methods have been proposed to solve the cross-domain sentiment classification problem. But these approaches mainly concentrate on extracting common features between domains. Unfortunately, they cannot fully consider the effects of the aspect information of the sentences. To overcome this drawback the Intelligent Extensive Information Rich Transfer Network (IEIRTN) is developed. This IEIRTN considered the information from both sentences and aspects. Specifically, this network aimed to find common features across domains and then extracted information from the aspects with the help of common features. Additionally, it adopted an extensive information rich learning mechanism for the sentiment classification, which combined sentences and aspects together. In addition, cross-border emotional divisions can accept well-trained classifications from one source domain to another target domain, which reduces the time and effort of training new classifiers in these domains. Methods of cross-border emotional division require data or other information in the target domain to train their model. Therefore, the cost of labeling each domain separately is very high as well as time consuming. However, the collection and processing of new corporations takes a heavy toll.

Automatic Intelligent Movie Sentiment Analysis Model Creation for Box Office Prediction using Multiview Light Semi Supervised Convolution Neural Network

In addition, data in the target domain may be private and not always available for training. To overcome these drawbacks, a proposed method was developed that extracts and classifies comments from one domain, called the source domain, and predicts the comments of another domain, called the target domain, using Multiview Light Semi Supervised Convolution Neural Network (IEIRTN).

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

In the past, a lot of work has been done in the field of online digging reviews. Different research teams are looking for different ways to use quotes and emotional analysis as a new generation model. Pang et al [1] implements machine learning methods to divide online reviews of movies collected from online movie databases into positive or negative ones by obtaining a list of 14 powerful keywords (Amazing Love - Great, still, beautiful, ugly, boring, wasted and?) which is then applied in a clear way to lay the foundation for the accuracy of the rankings. The results suggest the use of learning methods such as SVM, Naive Bayes and others. Get significant improvements on a local basis. The document focuses on classification using sequential methods, and this leads to improved accuracy. Work is being done to understand how clinging improves ranking accuracy. . Jeffrey et al [2] studied the classification of traffic networks using the peculiarities of software when communicating through networks. This document discusses how to handle two unauthorized clusters, K-Means and DBSCAN, which were previously not used to classify network traffic. He evaluated two algorithms and compared them with the previous autoclass algorithms, and the results showed that K-Means and DBSCAN worked better and faster than autoglass. Antonio et al. [3] Solve learning problems to classify texts using information obtained from training and a set of tests. To achieve this grouping is used as an additional step to classify texts that apply to both training and test sets. Experiments have shown significant improvements in classification results, especially for small training equipment. Noam et al [4] reported the use of sequential information barriers (IB) for the classification of unchecked documents. He proposed a simple procedure for seizing clusters to increase accuracy. Work demonstrating this method can extract clusters that match almost perfectly to existing skin. These works show an improvement in the accuracy of the classification if gripping is used as an extra step during application. At [11]A. Sivasantoshreddy, P. Kasat, and A. Jain tries to predict the opening of a movie box office using ultra-analysis. This article basically focuses on Twitter data for over-analysis. The main logic of the hip-hop analysis is the success of the film, which relies heavily on its weekend earnings and how much hype is among the pre-release ones. First, they found a number of movie-related tweets using a web-based link. These tweets are collected by the hour. There are three factors for measuring excess. The first factor is to calculate the "number of relevant tweets per second". The second factor is "finding the number of individual users who have posted Twitter." The third factor is "calculating the scope of tweets". Here, the scope of tweets means that different people's tweets have different values. Suppose if a famous actor or director posted a positive tweet for a movie, it was worth more than a tweet posted by an ordinary person. To calculate the scope of tweets, they count specific user followers. They calculated the number of relevant tweets per second. The second factor was "finding and calculating the range of tweets as hip factors, averaging these three factors for each film." Their analysis is based on the excess factor of the number of film screens to be released and the average price of all tickets for screening. Patterns are usually very simple calculations and they just count the number of tweets related to the film, but they do not use any language type to know whether the tweets are positive or negative. The neural network is used to predict the financial success of a box office film before it reaches the cinema. This forecast is turned into a ranking problem, which is divided into 9 classes. This model is presented with minimal features. In [13] he attempted to improve the overall projection of the film by analyzing the information that quantitative information data generated by Lydia (high-speed word processing system for collecting and analyzing information). There are two different models (regression model and K-nearest model). But they just consider it a high-budget film. The model fails if the regular word is used as a name and it is not possible to predict whether there is information about the film. M.H Latif, H. Afzal [14] who use IMDB databases as the main source and their data is not clean. Again, their data is unstable and noiseless as mentioned. So they used the middle trend as a standard to fill in missing values for various attributes. K. Jonas, N. Stefan, S. Daniel, F. Kai used emotional analysis and social media for predictions, [15] their hypothesis was based on an analysis of the intensity and positivity of RBCs.

They see film critics as influential and their prophetic vision. They use word bags that give wrong results when some words are used for negative means. There are no awards of any kind and only awards are given for Best Director, Actor / Actor and Best Supporting Actor / Actor. In some cases, predictions of film success were made by neurological analysis ([10], [18]). Some researchers made predictions based on social media, social media analysis, and hip-hop ([16] [17] [19] [20]) in which they calculated the positivity and number of comments related to a particular film. Baek et al. [4] Classify popular social media sites (Twitter, Yahoo! Movies, YouTube and blogs) in relation to their impact based on actual eWOM data. They proved that Twitter was the most influential channel in the early stages of filmmaking, contributing to the predictive power of high urgency (insight through its stimuli) and high permeability.

(Fast capture). They also showed that the impact of online reviews was less than in the early stages of filmmaking. O et al. [9] Study the influence of social media on user behavior (engagement). Although the CEB on Facebook and YouTube were positively correlated with box office revenue, the same effect was not seen on Twitter, they said.

Nafafi and Millin [8] compared the traditional approach of customer surveys to predicting the return of cash registers from the first night by a tweet-based approach that deals with content (emotion) and tweet volume analysis. They suggest that for tweets released shortly before release, the method of social media is presented similar to traditional survey techniques. In subsequent studies, they have shown that this social media analysis clearly predicts the long-term success of the film and is as good as traditional studies (Nafafi and Miller [5]). In a later study, Nafafi and Miller confirmed that the amount of tweets was a better prediction than the content of Twitter. These studies involve the development of intentional, behavioral, and / or behavioral interactions in Twitter that are measured shortly before the broadcast of each film at night. Divakaran et al. [7] Expanded previous studies by investigating pre-predictions for the film's success, as opposed to showing it based on post-release data. They show that the level of variables such as awareness, speech, expectations and intent to be accepted in the online community for upcoming films has a direct impact on the success of the film.

Moreover, very few people have predicted the success of the film on TV based on Twitter tweets and comments on YouTube. In both cases, the accuracy of the forecast will be questioned and will not yield appropriate results. Small domains are not a good idea to measure. In previous work, most research was based on attributes available before or after the film was released. Although some researchers have considered both types of attributes, in this case there are very few reports. Better probability of success in prediction increases with more attributes.

III. PROPOSED MOVIE SENTIMENT ANALYSIS MODEL

This section describes the procedures for collecting data, processing schemes, obtaining data for presenting data sets, sentiment prediction schemes, and combinations of knowledge collection methods used in experimental research. A sketch of the work is shown in Fig.1.

1. Gathering of Dataset
2. Data Cleaning
3. Feature Retrieving
4. Sentiment Examination Model Generation

A. Gathering of Dataset

To evaluate the predictable performance of psychology along with the linguistic characteristics of emotional exploration, this paper examines the sequence of different sources such as Twitter, Facebook, websites. It contains positive and negative emotions along with neutral feelings. For the database compilation process, this works using the summary structure in [11]. This work uses Twitter4J. Used to collect tweets. All tweets are marked positive, negative or neutral. Later than collecting tweets, automatic filtering is

used to delete inappropriate and unnecessary tweets. As a result, this work received 6,188 negative points, 4,891 positive points and 4,252 tweets. To get a neutral character, our work dataset contains 4,200 positives as well as 4,200 negative tweets as well as positive tweets.

B. Data Cleaning

Due to the inaccurate and unofficial nature of clean Twitter posts, pre-creation is required to eliminate some issues (such as copying unnecessary words and using 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]:

• Delete quotes and responses from other users' tweets with a string beginning with "@".
• Delete URLs
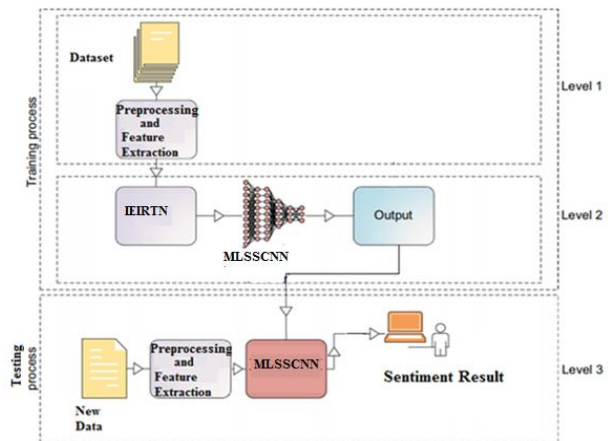• Delete "#" sign

Fig. 1. Sketch out of the Proposed Work

C. Feature Retrieving

Once the database is executed, the next step is to create a matrix of functions. Prior to retrieval, the tag function is first applied to pre-processed data. Tokenization is the process of dividing sentences into words. Then relax. The word stop is often used and often loses its basic meaning. Words like "continue" are just a few examples of words that stop.

There are two common download 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 the words / phrases that occur above 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 bookmark / bookmark list. In our work, we have implemented both methods for withdrawal. This end is followed by a continuous process of extinction. The purpose of the original is to reduce the form of reflection and sometimes to associate the derivative of the word with a common basic form.

This program is an easy way to narrow down some characters at the end of a word to their roots, using them just set the principle of cutting a few characters at the end of a word and hopefully they get Good results. And then features are retrieved by Intelligent Extensive Information Rich Transfer Network (IEIRTN). IEIRTN combined sentences and aspects together.

D. Sentiment Examination Model Generation

To create a highly effective model for emotional analysis from the Multiview Light Semi Supervised Convolution Neural Network (MVLSSCNN). At this stage, the extracted function matrix is analyzed as positive, neutral or negative to calculate the entire polar line. Multiview Sentiment Analysis (MSA) is used for this purpose. There are three main stages: emotional analysis, reverse emotional analysis, and final emotional analysis.

IV. RESULT AND ANALYSIS

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.

Precision Rate

The precision value is found by using the below formula

$$Precision = \frac{TP}{TP+FP} \quad (1)$$

Detection Accuracy

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

$$Accuracy = \frac{TP+TN}{TP+FP+TN+FN} \quad (2)$$

Specificity

Specificity is found by using the below formula

$$Specificity = \frac{TN}{(FP+TN)} \quad (3)$$

Recall Rate

The recall is found by using the below formula

$$Recall = \frac{TP}{TP+FN} \quad (4)$$

F-Measure

F-measure is found by using the below formula

$$F_m = (1+\alpha) * \frac{Precision * Recall}{\alpha * (Precision * Recall)} \quad (5)$$

Error Rate

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

$$Error\ Rate = \frac{No\ of\ Images\ of\ Falsely\ predicted\ sentiments}{Total\ No\ of\ statments} \quad (6)$$

Sensitivity

Sensitivity is found by using the below formula
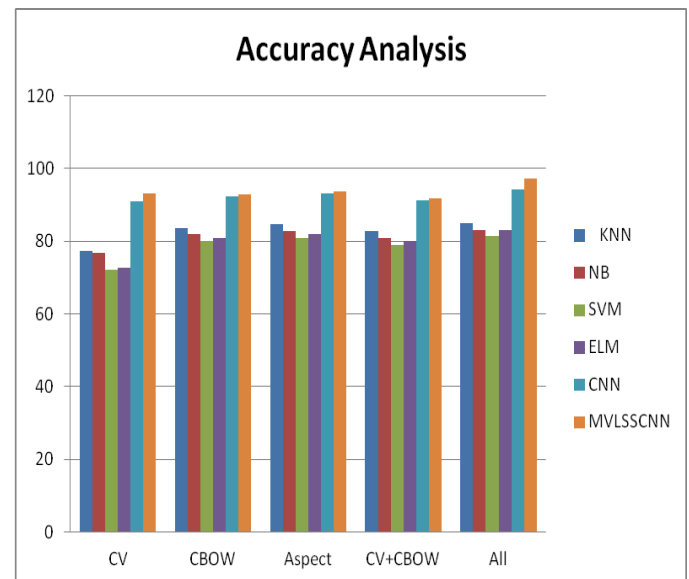
$$Sensitivity = \frac{TP}{(TP+FN)} \quad (7)$$

B. Trial No 1: Examination of MVLSSCNN using Accuracy

To examine the performance of this MVLSSCNN, 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 MVLSSCNN using Accuracy

Classifier						
Metrics	KNN	NB	SVM	ELM	CNN	MVLSSCNN
CV	77.3	76.7	72.2	72.7	91.1	93.21
CBOW	83.7	81.8	79.9	80.9	92.2	92.92
Aspect	84.6	82.7	80.8	81.8	93.1	93.83
CV+CBOW	82.7	80.8	78.9	79.9	91.2	91.91
All	85.1	83.1	81.3	83.0	94.2	97.23

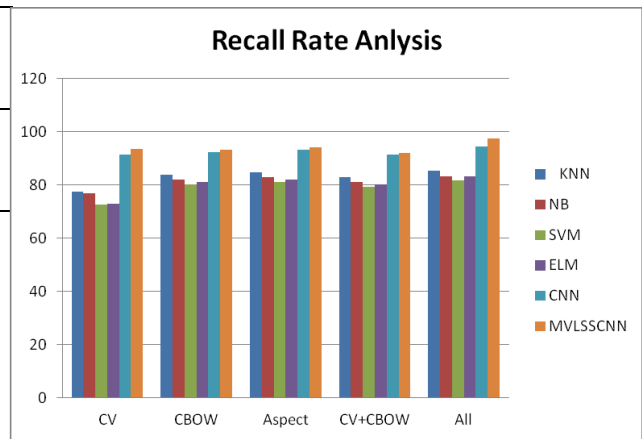### C. Trail No 2: Examination of MVLSSCNN by Precision Rate

To examine the performance of MVLSSCNN, 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 MVLSSCNN by Precision Rate

Metrics	KNN	NB	SVM	ELM	CNN	MVLSSCNN
All	85.33	83.35	81.6	83.25	94.45	97.45

Classifier						
Metrics	KNN	NB	SVM	ELM	CNN	MVLSSCNN
CV	77.6	77.0	72.6	73.1	91.4	93.52
CBOW	84.0	82.1	80.2	81.2	92.5	93.23
Aspect	84.9	83.0	81.1	82.2	93.4	94.14
CV+CBOW	83.0	81.1	79.2	80.2	91.5	92.22
All	85.4	83.4	81.6	83.3	94.5	97.54

Recall Rate Anlysis

E. Trial No 4: Examination of MVLSSCNN using Sensitivity

To examine the performance of MVLSSCNN, 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 MVLSSCNN using Sensitivity**
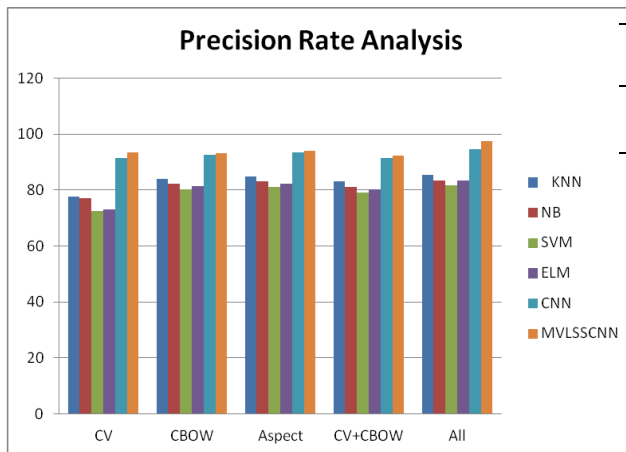
Classifier						
Metrics	KNN	NB	SVM	ELM	CNN	MVLSSCNN
CV	77.75	77.13	72.7	73.2	91.52	93.62
CBOW	84.16	82.26	80.36	81.39	92.65	93.33
Aspect	85.07	83.17	81.27	82.3	93.56	94.24
CV+CBOW	83.15	81.25	79.35	80.38	91.64	92.32
All	85.52	83.54	81.79	83.44	94.64	97.64

Precision Rate Analysis

D. Trial No 3 : Examination of MVLSSCNN through Recall Rate

To examine the performance of MVLSSCNN, 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 MVLSSCNN through Recall Rate

Classifier						
Metrics	KNN	NB	SVM	ELM	CNN	MVLSSCNN
CV	77.56	76.94	72.51	73.01	91.33	93.43
CBOW	83.97	82.07	80.17	81.2	92.46	93.14
Aspect	84.88	82.98	81.08	82.11	93.37	94.05
CV+CBOW	82.96	81.06	79.16	80.19	91.45	92.15

Sensitivity Analysis
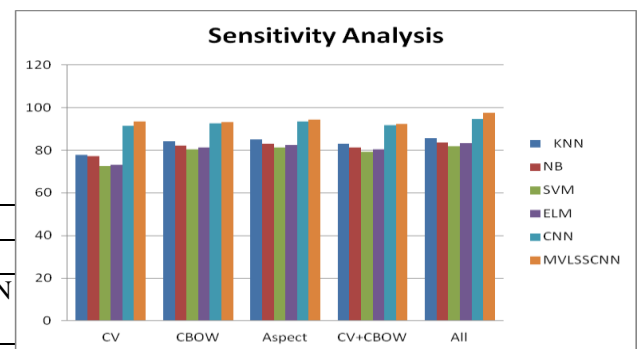
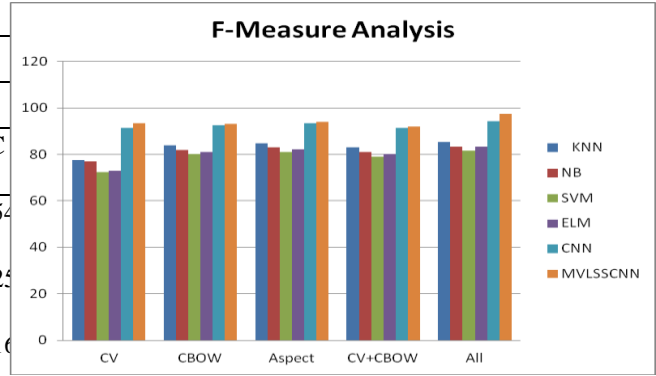## F. Trial No 5: Examination of MVLSSCNN by Specificity

To examine the performance of MVLSSCNN, 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 MVLSSCNN by Specificity

Metrics	Classifier KNN	NB	SVM	ELM	CNN	MVLSSCNN
CV	77.6	77.0	72.62	73.12	91.44	93.54
CBOW	84.08	82.1	80.28	81.31	92.57	93.25
Aspect	84.99	83.09	81.19	82.22	93.48	94.16
CV+CBOW	83.07	81.17	79.27	80.36	91.56	92.24
All	85.44	83.46	81.71	83.36	94.56	97.56

Specificity Analysis

G. Trial No 6 : Examination of MVLSSCNN through F-Measure

To examine the performance of MVLSSCNN, 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 MVLSSCNN through F-Measure

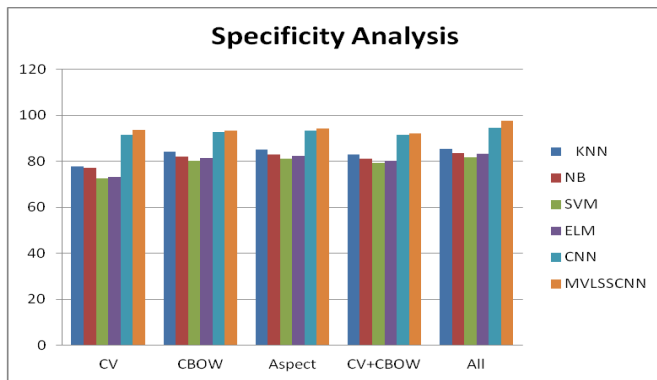| Metrics	Classifier KNN	NB	SVM	ELM	CNN	MVLSSCNN
CV	77.49	76.87	72.44	72.94	91.26	93.36
CBOW	83.9	82	80.1	81.13	92.39	93.07
Aspect	84.81	82.91	81.01	82.04	93.3	93.98
CV+CBOW	82.89	80.99	79.09	80.12	91.38	92.06
All	85.26	83.28	81.53	83.18	94.38	97.38

F-Measure Analysis

V. CONCLUSION

This article builds an intelligent emotional analysis system and in-depth learning approach to understand the relationship between online sentence review and aspects. This system is used to predict box office collections for movies. Experimental analysis shows that consideration of certain characteristics and aspects is more than individual. For the emotional analysis of the film, the highest predictive performance (98.31%) was achieved by a combination of MVLSSCNN language and psychological processes. Thus, the proposed approach and the combination of sentences and aspects are best illustrated in the analysis of the film's emotions.

REFERENCES

1. B. Pang, L. Lee, and S. Vaithyanathan, "Thumbs up?: sentiment classification using machine learning techniques," in Proceedings of the ACL-02 conference on Empirical methods in natural language processing-Volume 10. Association for Computational Linguistics, 2002, pp. 79–86.
2. J. Erman, M. Arlitt, and A. Mahanti, "Traffic classification using clustering algorithms," in Proceedings of the 2006 SIGCOMM workshop on Mining network data. ACM, 2006, pp. 281–286.
3. A. Kyriakopoulou and T. Kalamboukis, "Text classification using clustering," in Proceedings of The 17th European Conference on Machine Learning and the 10th European Conference on Principles and Practice of Knowledge Discovery in Databases (ECML-PKDD), Burlin, Germany, 2006, pp. 28–38.
4. Baek, H., Oh, S., Yang, HD. and Ahn, J. (2017). Electronic word-of-mouth, box office revenue and social media. Electronic Commerce Research and Applications Vol. 22, pp. 13-23.
5. N. Slonim, N. Friedman, and N. Tishby, "Unsupervised document classification using sequential information maximization," in Proceedings of the 25th annual international ACM SIGIR conference on Research and development in information retrieval. ACM, 2002, pp. 129–136.
6. Najafi, H.L. and Miller, D.W. (2016). Predicting long-term box office performance of movies using social media and its comparison with traditional survey method. International Journal of Research and Reviews in Applied Sciences, Vol. 26 No. 3, pp. 120-135.
7. Najafi, H.L and Miller, D.W. (2017). Assessing alternative methods for predicting box office success using tweets, Journal of Advancements in Applied Business Research, Vol 6 No. 1, pp. 26-38.
8. Divakaran P.K.P, Palmer A., Søndergaard H. A., and Matkovskyy, R. (2017). Pre-launch prediction of market performance for short lifecycle products using online community data. Journal of Interactive Marketing, Vol. 38, pp. 12-28.
9. Najafi, H. and Miller, D.W. (2015). Comparing analysis of social media content with traditional survey methods of predicting opening night box-office revenues for motion pictures. Journal of Digital & Social Media Marketing, Vol. 3 No. 3, pp. 262-278.

10. Oh, C., Roumani, Y., Nwankpa, J.K., , Hu, H.F. (2017). Beyond likes and tweets: Consumer engagement behavior and movie box office in social media. Information & Management, Vol. 54 No. 1, pp. 25-37.

11. R. Sharda and E. Meany, "Forecasting gate receipts using neural network and rough sets," in Proceedings of the International DSI Conference, 2000, pp. 1–5.

12. A. Sivasantoshreddy, P. Kasat, and A. Jain, "Box-Office Opening Prediction of Movies based on Hype Analysis through Data Mining," International Journal of Computer Applications, vol. 56, no. 1, pp. 1–5, 2012.

13. R. Sharda and D. Delen, "Predicting box-office success of motion pictures with neural networks," Expert Systems with Applications, vol. 30, no. 2, pp. 243–254, 2006.

14. W. Zhang and S. Skiena, "Improving Movie Gross Prediction through News Analysis," 2009 IEEE/WIC/ACM International Joint Conference on Web Intelligence and Intelligent Agent Technology, 2009.

15. M.H Latif, H. Afzal "Prediction of Movies popularity Using Machine Learning Techniques", National University of Sceinces and technology, H-12, ISB, Pakistan.

16. K. Jonas, N. Stefan, S. Daniel, F. Kai "Predicting Movie Success and Academi Awards through Sentiment and Social Network Analysis" University of Cologne, Pohligstrasse 1, Cologne, Germany.

17. J. Duan, X. Ding, and T. Liu, "A Gaussian Copula Regression Model for Movie Box-office Revenue Prediction with Social Media," Communications in Computer and Information Science Social Media Processing, pp. 28–37, 2015.

18. L. Doshi, J. Krauss, S. Nann, and P. Gloor, "Predicting Movie Prices Through Dynamic Social Network Analysis," Procedia - Social and Behavioral Sciences, vol. 2, no. 4, pp. 6423–6433, 2010.

19. T. G. Rhee and F. Zulkernine, "Predicting Movie Box Office Profitability: A Neural Network Approach," 2016 15th IEEE International Conference on Machine Learning and Applications (ICMLA), 2016.

20. T. Liu, X. Ding, Y. Chen, H. Chen, and M. Guo, "Predicting movie Box-office revenues by exploiting large-scale social media content," Multimedia Tools and Applications, vol. 75, no. 3, pp. 1509–1528, Feb. 2014.

21. Z. Zhang, B. Li, Z. Deng, J. Chai, Y. Wang, and M. An, "Research on Movie Box Office Forecasting Based on Internet Data," 2015 8th International Symposium on Computational Intelligence and Design (ISCID), 2015.

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.

Dr.Manoj Manuja is Pro-Vice Chancellor at Chitkara University Rajpura, Punjab, India . He has several years of experience in teaching and research .He has published more than 24 papers in reputed journals and his area of specialization is on Specialization in Internet of Things(IOT), machine learning and data science . He is guiding research scholars in the PhD programme

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