An Intelligent High Performance Automatic Sentiment Analysis Model Creation using Deep Convolution Neural Network

Srinidhi B. S., Suchithra R.

Abstract: With the advancement of data and communications technology, social media platforms and small news blogs serve as significant sources of data. In a small blogging forum, people can share their opinions, complaints, feelings and behaviors about the topic, current problems, and products. Emotional examination is an significant examination area in natural language processing that intends to target the emotion of the source material. Twitter is a well-liked stage where people around the globe can interrelate through user-produced messages. Data received from Twitter can give out as a primary source for many applications, together with event recognition, news recommendations as well as emergency supervision. In the categorization of emotions, recognition of suitable sub feature set acts an significant role. LIWC (Linguistic Inquiry and Word Count) is a research program for text examination to retrieve psychometric features from text documents. In this article this work present a psychometric method called the intelligent high performance automatic sentiment analysis model (IHPASAM) for Twitter emotion analysis. In this scheme, this work employed two main types of LIWC (linguistic processes along with psychological) as feature sets. To discover the predictive efficiency of dissimilar feature engineering systems, five supervised learning techniques (Naïve Bayes, logistic regression, k-nearest neighbor algorithm, support vector machines as well as convolution neural network) along with proposed Intelligent Deep Convolution Neural Network (IDCNN) are employed. Investigational outcome show that the ensemble feature sets provides a superior predictive efficiency than the individual set.

Keywords: Natural Language Processing (NLP), Data scraping, social media analysis, Sentiment Analysis, psychological feature sets, Twitter, Machine Learning and IDCNN.

I. INTRODUCTION

With the growth of social media, a few applications that use emotional examination have evolved quickly in modern days. Most people view forums, public opinion sites, or social networks that want to gain experience from further users earlier than making a conclusion. As the amount of reviews increases dramatically on the web, the job of understanding all of these reviews becomes impracticable for users. Therefore, it is necessary to automate the analysis and classification of these opinions.

* Correspondence Author
B. S. Srinidhi*, research scholar, Jain Deemed to be University, Bangalore, India.
R. Suchithra, director, department of MCA, Jain Deemed to be University, Bangalore, India.

Emotional analysis has turn out to be a well-liked topic to recognize public opinion from unstructured web information. Emotional analysis is an significant investigate area in natural language processing that intend to target the emotion of the source material. Emotional examination can be employed to obtain data about innovative products as well as services. It can be additionally used to recognize the positive as well as negative feature of a exact product or service. Emotional analysis techniques can be alienated into two classes: lexicon-based techniques and the machine-based techniques. In totaling up, emotional examination can be performed in dissimilar ways, depending on the level of detail. Depending on the level of detail, emotion examination techniques are clustered into three groups: sentence, aspect-level as well as document-level sentiment examination. In this circumstance, dissimilar methods rooted in supervised training and uncontrolled training can be employed . For example, the issues investigated focused on identifying, [1] analyzing and creating emotional phonics [9], [8] evaluating and categorizing Twitter posts [3], [12], [19], features selection [6], negation [5]. A number of other suggestions have attempted to suggest problems with the classification of emotions at dissimilar levels, namely document-level [13] [20], sentence-level [11], word-level along with aspect-level [16]. On the other hand, one of the key problems in emotion categorization is the presence of several theoretical system that rules the expression of the language of emotion. Human psychology, which deals with cultural, biological aspects, personal as well as social, can be an significant trait that should be taken into account in emotional analysis. To this end, LIWC and MatlabNLP-master text examination tools are helpful tools because they help extract psychometric and linguistic features from natural language texts. Another option that has attracted many researchers in recent years is to create a limited range of emotions and to create predictable patterns for each emotion. A multi-tag version that allows documents to reflect more than one sense uses a single template. Unlike sentimental analysis, this approach modeled the structure of higher-order human emotions. There are some notable difficulties with the above method. First, it is difficult to capture the complex statistical relationship between the number of binary variables (representing emotions) and one large vector (representing a document). It is also difficult to imagine a reliable procedure for compiling a limited list of possible human emotions. Lastly, it is unclear how to use a document that expresses emotions like exhaustion when installing predictive models

Retrieval Number: B4510129219/2020©BEIESP
DOI: 10.35940/ijeat.B4510.029320
Journal Website: www.ijeat.org

1288

Published By:
Blue Eyes Intelligence Engineering
& Sciences Publication

for something like drowsiness. The use of tagged files only when mounting models with their labeled tags ignores the relationship between emotions and problems with a few notes.

Not only recent sentiment analysis models that attempt to determine the precision and intensity of data. Algorithms split text into text with clarity and learn the meaning of words, including each polarizer. Although previous studies have addressed many issues with emotional analysis by design, architecture, and its algorithms, this model can predict one's emotional state rather than textual data. Studies should focus on finding the emotions of the person who wrote the word, not on the feelings of the words or sentences.

To overcome these shortcomings, existing models for emotion analysis have developed proposed methods that extract both psychological and linguistic characteristics of emotion classification. The proposed method takes into account the larger dimension of the expansion of the concept of emotion, which represents a broader range of human emotions. The proposed method goes beyond the previous work, it varies continuously, not the limits of human emotion.

The rest of the paper is prearranged as follows. Section 2 presents the work related to emotional analysis. The first steps to identify pathogens and to update and improve pathology are established in Section III. Section 4 explains the outcome of the proposed methodological information. In this segment, the proposed system is assessed through obtainable approaches to demonstrate the efficiency of the new method. Conclusion along with future work of this learning is proposed in Section 4.

II. RELATED WORK

The analysis of emotions on Twitter information has drawn considerable research concentration. This segment momentarily examine the already implemented work on sentiment examination on Twitter information. Twitter's sentiment examination creates many difficulties because of its informal, informal, and informal message and feature. Therefore, identifying appropriate traits is an significant area of research. For example, Go et al. [7] Understand the use of abbreviations, large and large diagrams, and sections of spoken labels as a function. In the classification stage, Naive Bayes, maximum entropy, along with host sustain groups were employed . The empirical examination shows that increasing the number of complexes and grams gives better predictability than other features of the engineering scheme. Some speech markers are not a useful function, and the highest classification efficiency is achieved by the maximum number of learners.

In another information, Barbosa and Feng [4] examined n-gram prediction performance and the function of tweets. The empirical examination of the vector sustain engine shows that the tweet syntax features improve the anticipation of emotion schemas on the tweets, and L-gram cannot display the full text message. Likewise, Pak and Parbbe observed the use of the microgram as well as sections of speech markers in a systematic manner. In the empirical examination , the polynomials of the Nam Weyer, the support vector machine, and the classical field classifier are employed . Empirical examination has shown that the use of ngram-related speeches provides a better predictive effect in sentiment

examination of Twitter data. In another information, Koulumpis et al. [10] examined the use of M-gram features, glossary functions, the section of labels, speech and data block features in the examination of Twitter emotion information.

Experimental examination showed that the highest predictive efficiency among the various features of the engineering scheme was obtained from the n-g characteristics in terms of semantic and microscopic features. Moreover, the results show that integrating parts of the speech feature reduces the predictability of emotion ratings. Likewise, Agarwal et al. [2] studied the use of speech, vocabulary, as well as microscopy to analyze data sentiment through Twitter. In addition, they introduced tree representation to expand functional diagrams effectively. Experimental examination has shown that the use of prefixes of words in terms of their speech labels provides the maximum categorization accuracy. In another learning,

Saif et al. [18] Learn about the use of non-functional features of speech as well as subjective emotion examination on Twitter. Experimental examination has shown that a rigorous approach based on a set of functions provides better predictive results than other engineering schemes for the function. Onan [14] investigated the predictive representation of dissimilar M-gram models (such as dimensions, grammars, and triggers) and their combinations in analyzing Turkish emotions on Twitter. In the empirical examination , the maximum prognostic efficacy was attained through the grouping of the uremic along with grammatical features.

In an additional learning, Salas-Zarate et al. [17] observed the effectiveness of features of psychology in the investigation of product reviews. This learning considers language processing, psychosocial processes, personal anxiety, speech types, and punctuation. While the examination of Twitter sentiment information focused on n-grams as part of microphone speech along with features, this learning aimed to inspect the psychological and linguistic predictive performance obtained by LIWC in the examination . Emotions. On Twitter.

III. PROPOSED APPROACH

Recent research activities related to recall, conceptual analysis, along with emotion recognition from natural language texts are all beneath the umbrella of emotional computation. There is now a enormous quantity of textual information on social media (for example, forums, social media as well as blogs) about consumers' ideas about buying products and service experiences. Feeling analyzed or recalling ideas is part of an investigation that analyzes people's thoughts and feelings from essays available online. In this document, this work develop an wide experiment to assess the usefulness of psychometric as well as linguistic features in emotion classification. To this end, this work propose one intelligent high performance automatic sentiment analysis model (IHPASAM) using Intelligent Deep Convolution Neural Network (IDCNN).

The outline of our proposed work is shown in Figure 1. The proposed work consists of four modules. They are
1. Data Preprocessing
2. Feature Extraction
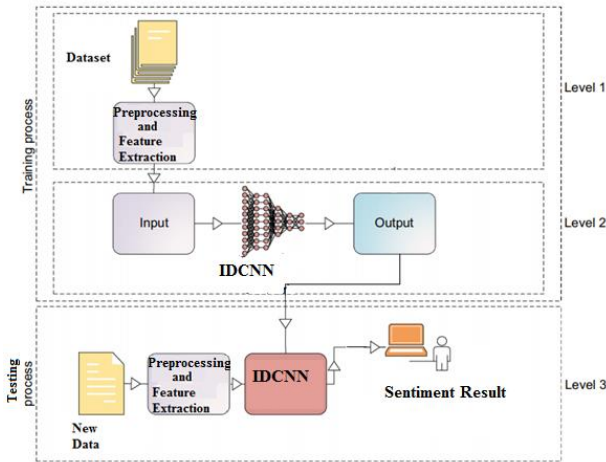3. Sentiment Analysis Model Generation

Fig. 1. Outline of the Proposed Work

A. Preprocessing

Initially, Twitter account data was collected from Twitter accounts. The data is then reformatted at the data conversion stage to prepare for the following algorithm. In the process of data clearance, duplicate records are deleted and missing data is deleted or filled in with a value of zero.

B. Feature Extraction

After the database has been processed, the next step is to create a feature matrix. Before feature extraction, tokens are first applied to pre-processed data. Tokenization is merely the procedure of separating a sentence into words to detect stop words. Stop words a repeatedly employed word as well as is so ordinary that they drop their semantic meaning. There are two general techniques of eliminating words, as well as together are obvious.

Individual method is to calculate each and every event words as well as allocate a numeric value for the amount and throw away every terms/words that happen additional than the particular value. An additional method is to have a already defined word smash list that can be created by a tiny list of stop words, which can be removed from the list of tokens/tokenized sentences. n our work, we have implemented both methods to remove stop words.

Finally, the process of stemming is continued. The purpose of is to reduce the reflective form and sometimes relate the derivative of the word to a common basic form. Stemming is easy way to reduce conditions to their roots, just set a rule for cutting some characters at the end of a word and hope they get the best of them. And then the features of the matrix are created.

And then the features of the matrix are created. In our work, psychosocial and linguistic features were developed using four word-level procedures, such as dividing the word extent, the average amount of words for each sentence, the distribution of word lengths by the incidence of one word to the incidence of words 14. . These world-class measures are calculated using a combination of terms. This approach focuses on assessing documents primarily by searching at separate words. Because of user-defined dictionaries, all word included in the dictionary gives to the average effeteness of the file in one (or additional) dimensions. Controlled machine learning techniques typically take account as functions that can be useful in document classification. However, the term is rarely employed in isolation. The words in the text are often lined up to transmit a sense that not every word can give itself. In this sense, an article is additional than the amount of the words available. Individual method to consider this is to concentrate on a group of words that be inclined to come collectively in a exact context

LIWC (linguistic investigation and word counting) to retrieve psychometric features from the dataset. LIWC is a software program that offers successful tools for the emotional, cognitive, as well as structural learning of word-for-word languages. An early approach to psychology concerns almost exclusively qualitative philosophical analysis. Intended for this learning, this work employed the English dictionary, LIWC2007. It was collected of 7515 words along with originals. Every term is grouped into one or additional of the 72 classes, which are incorporated with predefined in LIWC. In addition, these classes are grouped into four sizes: standard linguistic procedure, relativity, personal concerns as well as psychological procedure.

C. Sentiment Analysis Model Generation

To build a high-efficiency Sentiment Analysis model by intelligent deep convolutional neural network CNN . Convolutional Neural Networks are like to normal neural networks. They are composed of neurons that have learnable weights and biases. that are of learning weight and bias. Each neuron receives some input, executes a point product, and is willing to follow with a non-linearity. A CNN consists of one or more convolutional layers and pooling layers. Pooling layers are also called sub sampling layers. Normally CNN are employed for classification purpose. The pooling layer is employed to perform down sampling. It is employed to reduce the amount of computation time by reducing the extracted features in convolution layer.

Max and Average pooling are two types of pooling layers in CNN. When combined, the largest pixel values are counted in max pooling. The average of all values is taken into account in average pooling. The result of the this layer is given as input to the next layer of CNN. CNN has very high computational cost for large feature maps. CNN is slow to train large feature maps. To overcome the drawback of CNN, a optimization technique Lion Pulse Coupled Neural Network (LPCNN) is projected. This reduces high computation cost and improves speed. The size diminution of image space is comprehend by vector of features that is created by LPCNN from multidimensional image space to short dimensional feature space. The overall diagram of IDCNN is shown in below.

An Intelligent High Performance Automatic Sentiment Analysis Model Creation using Deep Convolution Neural Network
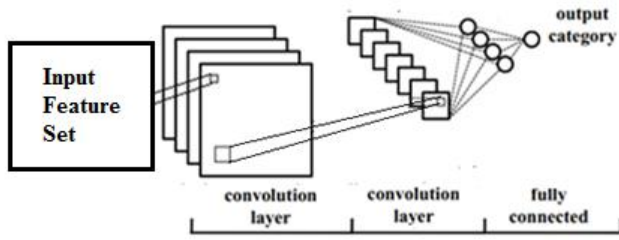

Fig. 2. Sentiment Analysis using IDCNN

The shortcomings of conventional lion schemes, for instance awful accuracy, short convergence value, as well as effortlessness to drop into local best value. To overcome the drawback lion scheme is combined with PCNN. The purpose of LPCNN scheme is to find a subset of relevant attributes that guide to a diminish the volume of the dataset, speed up the execution time and reduces the prediction error. Our proposed LPCNN feature selection procedure contains of four phases, specifically, the creation of the subset, subset assessment, stopping condition, and result specification. This is shown in Fig.3. In the procedure for creating a subset, candidate subsets are created and evaluated. The newcomer compares favorably with the previous one according to certain evaluation criteria. If the new subset is found to be better, it replaces the previous best subset. This process is repeated until the criteria are stopped. The standard LPCNN model is described as iteration by the following equations. From the below formulations,

S_{ij} is the input stimulus such as the best values from Lion
$F_{ij}[n]$ – Feedback input of the neuron in (i,j)
$L_{ij}[n]$ – Linking item

$$F_{ij}[n] = e^{-\alpha_F}F_{ij}[n-1] + V_F \sum_{K,I} W_{i,j,K,I} Y_{ij}[n-1] + S_{ij}$$
(1)

$$L_{ij}[n] = e^{-\alpha_L}L_{ij}[n-1] + V_L \sum_{K,I} M_{i,j,K,I} Y_{ij}[n-1]$$
(2)

$$U_{ij}[n] = F_{ij}[n](1 + \beta L_{ij}[n])$$
(3)

$$Y_{ij}[n] = \begin{cases} 1 & U_{ij}[n] > T_{ij}[n] \\ 0 & \text{otherwise} \end{cases}$$
(4)

$$T_{ij}[n] = e^{-\alpha_t}T_{ij}[n-1] + V_T Y_{ij}[n-1]$$
(5)

Then the selected best subset is authenticated because of previous information regarding the data sets. Initially, the subset generation is started with empty set. And then take the first two attribute from input and previous layer output and added in subset. After that finding the Correlation between the attributes in subsets are calculated. The subset evaluation of the proposed scheme is rooted in the filter model. So the correlation values of the subset is compared with threshold value. The steps for subset estimation assign status to individual attributes based on correlation coefficients. Because highly correlated attributes cannot be part of a dataset together, only those attributes can participate in the LPCNN search operation, which must look for little correlation coefficients. Otherwise, the attributes are put into the subset.
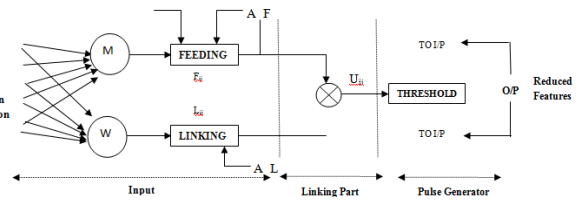

Fig.3. Structure of LPCNN

This operation continuous until all the other attribute are combined and checked with the attribute in the subset. So that the proposed classifier is called Intelligent Deep classifier. After train the model the sentiment label value is shown.

IV. RESULT AND ANALYSIS

A. Dataset Collection

To assess the extrapolative analysis of the psychosocial along with linguistic features of emotion analysis, this work analyzed Twitter posts that contained negative, neutral and positive emotions. This work employed Twitter4J, an open source Java library, to use Twitter Streaming API, to gather tweets. All tweet is labeled with a group of labels such as negative, neutral or positive. Later than gathering tweets, automatic filtering is performed to eliminate inappropriate as well as unnecessary tweets (copy with duplicate). So this work got a group of 6,218 negative, 4891 positive as well as 4252 tweets. To achieve a balanced corpus, the final dataset includes 4,200 tweets of negative tweets and 4,200 positive tweets.

B. Efficiency Parameters

Several performance indicators are available to assess the effectiveness of the proposed sentiment labeling procedure. This article uses the Precision Rate, Detection Accuracy, Sensitivity, Recall Rate, F-Measure, Specificity to assesses the performance.

Detection Accuracy

Detection accuracy is a quantity system that detects the amount of nearness among the original captioned text as well as the properly captioned text.

$$Accuracy = \frac{TP+TN}{TP+FP+TN+FN} \quad (4.1)$$

Precision Rate

The precision is the portion of derived instances that are applicable to the find.

$$Precision = \frac{TP}{TP+FP} \quad (4.2)$$

Recall Rate

The recall is one of the significant examples that is extracted in relation to the input image.

$$Recall = \frac{TP}{TP+FN} \quad (4.3)$$

Sensitivity

Sensitivity is named the true positive frequency, or the call rate in some sectors detects the fraction of positive effects.

$$Sensitivity = \frac{TP}{(TP+FN)} \quad (4.4)$$

Specificity

Specificity detects the percentage of negatives that are rightly defined as a fraction.

$$Specificity = \frac{TN}{(FP+TN)} \quad (4.5)$$

F-Measure

The F-Measure is the proportion of the product of accuracy in recalling the addition of accuracy with recall. This measure can be intended as,

$$F_m = (1 + \alpha) * \frac{Precision \cdot Recall}{\alpha \cdot (Precision \cdot Recall)} \quad (4.6)$$
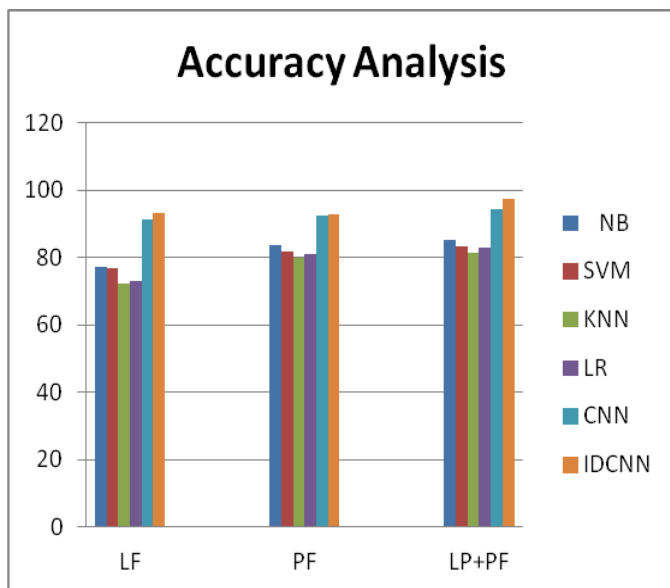
Experiment No #1 : Accuracy Analysis of Proposed Sentiment Analysis Model

In this experiment, the contribution of each classifier method is evaluated. Measures of Precision Rate, Detection Accuracy, Sensitivity, Recall Rate, F-Measure, Specificity, as well as Error Rate measures were employed to assess the effectiveness of this emotion analysis scheme. Ideally, the best scheme for emotion analysis is adopted to have a peak Accuracy value. Table 1 shows the detection accuracy values of IDCNN.

Table 1: Detection Accuracy Values of Proposed Sentiment Analysis Model

Classifier Metrics	NB	SVM	KNN	LR	CNN	IDCNN
LF	77.35	76.73	72.3	72.8	91.12	93.22
PF	83.76	81.86	79.96	80.99	92.25	92.93
LP+PF	85.12	83.14	81.39	83.04	94.24	97.24

As can be perceptible from Table 1, the accuracy of IDCN in the range 93-97 is higher than the other methods. Therefore, the IDCNN classifier is considered best for emotional analysis. Figure 8 shows the accuracy of the classification approach.

As can be perceptible from the figure above, the accuracy of IDCN in the range 93-97 is higher than that of the other

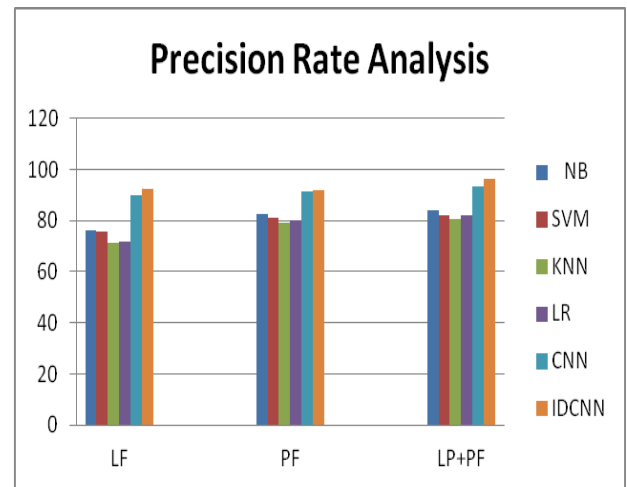methods. Therefore, the IDCNN classifier is best suited for emotion analysis.

Experiment No #2 : Precision Rate Analysis of Proposed Sentiment Analysis Model

In this experiment, the contribution of each classifier method is evaluated. Measures of Precision Rate, Detection Accuracy, Sensitivity, Recall Rate, F-Measure, Specificity, as well as Error Rate measures were employed to assess the effectiveness of this emotion analysis scheme. Ideally, the best scheme for emotion analysis is adopted to have a peak Precision value. Table 2 lists the Precision Rate values of IDCNN.

Table 2: Precision Rate Values of Proposed Sentiment Analysis Model

Classifier Metrics	NB	SVM	KNN	LR	CNN	IDCNN
LF	76.33	75.71	71.28	71.78	90.1	92.2
PF	82.74	80.84	78.94	79.97	91.23	91.91
LP+PF	84.1	82.12	80.37	82.02	93.22	96.22

As can be perceptible from Table 2, the precision rate of ICNNs in the range 92-96 is higher than the other methods. Therefore, the IDCNN classifier is considered best for emotional analysis. Figure 8 describes the precision measure of classification.

As can be perceptible from the figure above, the precision value of IDCN in the range 92-96 is higher than that of the other methods. Therefore, the IDCNN classifier is best suited for emotion analysis.

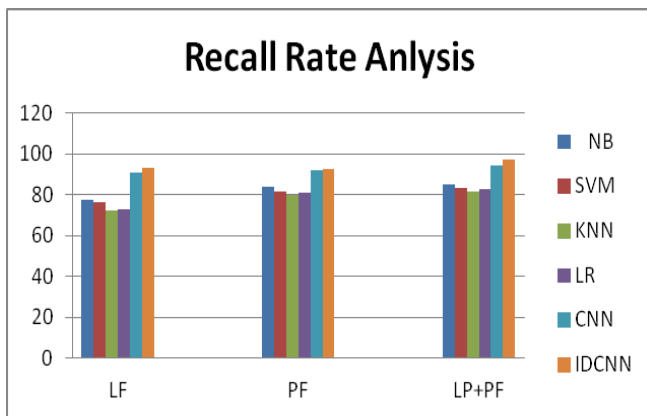**Experiment No #3 : Recall Rate Analysis of Proposed Sentiment Analysis Model**

In this experiment, the contribution of each classifier method is evaluated. Measures of Precision Rate, Detection Accuracy, Sensitivity, Recall Rate, F-Measure, Specificity, as well as Error Rate measures were employed to assess the effectiveness of this emotion analysis scheme. Ideally, the best scheme for emotion analysis is adopted to have a peak Recall Rate value. Table 1 shows the Recall Rate value of IDCNN.

Table 3: Recall Rate Analysis of Proposed Sentiment Analysis Model

Classifier Metrics	NB	SVM	KNN	LR	CNN	IDCNN
LF	77.24	76.62	72.19	72.69	91.01	93.11
PF	83.65	81.75	79.85	80.88	92.14	92.82
LP+PF	85.01	83.03	81.28	82.93	94.13	97.13

As can be perceptible from Table 3, the recall value of IDCNN in the range 93-97 is higher than the other methods. Therefore, the IDCNN classifier is considered best for emotional analysis. Figure 8 depicts the recall rate measures of the classification method.

Recall Rate Anlysis

As can be perceptible from the figure above, the recall value of IDCN in the range 93-97 is higher than that of the other methods. Therefore, the IDCNN classifier is best suited for emotion analysis.
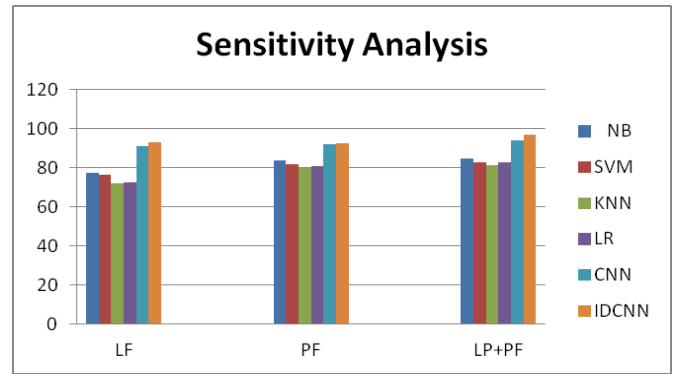
Experiment No #4 : Sensitivity Analysis of Proposed Sentiment Analysis Model

In this experiment, the contribution of each classifier method is evaluated. Measures of Precision Rate, Detection Accuracy, Sensitivity, Recall Rate, F-Measure, Specificity, as well as Error Rate measures were employed to assess the effectiveness of this emotion analysis scheme. Ideally, the best scheme for emotion analysis is adopted to have a peak sensitivity value. Table 4 shows the sensitivity values of IDCNN.

Table 4: Sensitivity Values of Proposed Sentiment Analysis Model

Classifier Metrics	NB	SVM	KNN	LR	CNN	IDCNN
LF	77.19	76.57	72.14	72.64	90.96	93.06
PF	83.6	81.7	79.8	80.83	92.09	92.77
LP+PF	84.96	82.98	81.23	82.88	94.08	97.08

As can be perceptible from Table 4, the sensitivity rate of ICNNs in the range 93-97 is higher than the other methods. Therefore, the IDCNN classifier is considered best for emotional analysis. Figure 8 describes the sensitivity measure of classification.

Sensitivity Analysis

As can be perceptible from the figure above, the sensitivity value of IDCN in the range 93-97 is higher than that of the other methods. Therefore, the IDCNN classifier is best suited for emotion analysis.
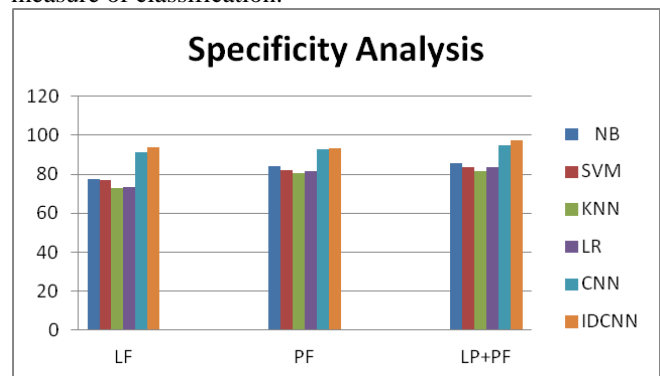
Experiment No #5 : Specificity Analysis of Proposed Sentiment Analysis Model

In this experiment, the contribution of each classifier method is evaluated. Measures of Precision Rate, Detection Accuracy, Sensitivity, Recall Rate, F-Measure, Specificity, as well as Error Rate measures were employed to assess the effectiveness of this emotion analysis scheme. Ideally, the best scheme for emotion analysis is adopted to have a peak Specificity value. Table 5 shows the Specificity values of IDCNN.

Table 5: Specificity Values of Proposed Sentiment Analysis Model

Classifier Metrics	NB	SVM	KNN	LR	CNN	IDCNN
LF	77.7	77.08	72.65	73.15	91.47	93.57
PF	84.11	82.21	80.31	81.34	92.6	93.28
LP+PF	85.47	83.49	81.74	83.39	94.59	97.59

As can be perceptible from Table 5, the specificity rate of ICNNs in the range 93-97 is higher than the other methods. Therefore, the IDCNN classifier is considered best for emotional analysis. Figure 8 describes the sensitivity measure of classification.

Specificity Analysis

As can be perceptible from the figure above, the specificity value of IDCN in the range 93-97 is higher than that of the other methods. Therefore, the IDCNN classifier is best suited for emotion analysis.

Experiment No #5 : F-Measure Analysis of Proposed Sentiment Analysis Model

In this experiment, the contribution of each classifier method is evaluated. Measures of Precision Rate, Detection Accuracy, Sensitivity, Recall Rate, F-Measure, Specificity, as well as Error Rate measures were employed to assess the effectiveness of this emotion analysis scheme. Ideally, the best scheme for emotion analysis is adopted to have a peak F-Measure value. Table 5 shows the F-Measure values of IDCNN.

Table 6: F-Measure Values of Proposed Sentiment Analysis Model

Classifier Metrics	NB	SVM	KNN	LR	CNN	IDCNN
LF	78.42	77.8	73.37	73.87	92.19	94.29
PF	84.83	82.93	81.03	82.06	93.32	94.11
LP+PF	86.19	84.21	82.46	84.11	95.31	98.31

As can be perceptible from Table 6, the F-Measure values of ICNNs in the range 93-97 is higher than the other methods. Therefore, the IDCNN classifier is considered best for emotional analysis. Figure 8 describes the F-Measure values of classification.
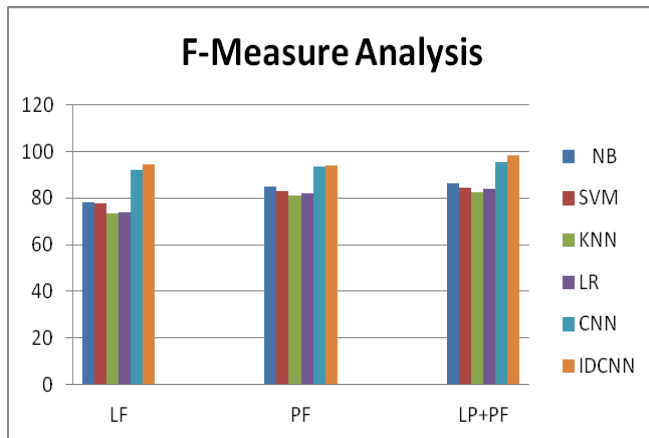
As can be perceptible from the figure above, the F-Measure value of IDCN in the range 94-98 is higher than that of the other methods. Therefore, the IDCNN classifier is best suited for emotion analysis.

V. CONCLUSION

In this paper, this work review the psychological and linguistic predictive performance that LIWC receives in emotional analysis on Twitter. To this end, two main types of LIWC (namely, linguistic processes and psychological processes) as well as their groupings are considered as feature arrays. Investigational information among the ranking algorithm shows that the psychopathology set can provide encouraging results when analyzing the sentiment of Twitter data. Experimental analysis shows that the band is more than individual. For Twitter sentiment analysis, the maximum predictive performance (98.31%) was attained by combining the language and mental processes of IDCNN. Therefore, this proposed method and combination of language processes, psychological processes, perform best in emotion analysis.

REFERENCES

1. Abdul-Mageed M, M. Diab, and S. Kübler, 'SAMAR: Subjectivity and sentiment analysis for Arabic social media', Comput. Speech Lang., vol. 28, no. 1, pp. 20–37, Jan. 2014.
2. Agarwal A, B.Xie, I.Vovsha, O.Rambow and R. Passonneau, "Sentiment analysis of twitter data", in Proceedings of ACL 2011, USA, 2011, pp. 30-38.
3. Bae Y and H. Lee, 'Sentiment analysis of twitter audiences: Measuring the positive or negative influence of popular twitterers', J. Am. Soc. Inf. Sci. Technol., vol. 63, no. 12, pp. 2521–2535, Dec. 2012.
4. Barbosa L and J. Feng, "Robust sentiment detection on twitter from biased and noisy data", in Proceedings of ACL, USA, 2010, pp. 36-44.
5. Cruz N. P, M. Taboada, and R. Mitkov, 'A machine-learning approach to negation and speculation detection for sentiment analysis', J. Assoc. Inf. Sci. Technol., p. n/a–n/a, Jun. 2015.
6. Duric A and F. Song, 'Feature selection for sentiment analysis based on content and syntax models', Decis. Support Syst., vol. 53, no. 4, pp. 704–711, Nov. 2012.
7. Go A, R. Bhayani, and L. Huang, "Twitter sentiment classification using distant supervision", CS224N Project Report, 2009.
8. Hogenboom A, B. Heerschop, F. Frasincar, U. Kaymak, and F. de Jong, 'Multi-lingual support for lexicon-based sentiment analysis guided by semantics', Decis. Support Syst., vol. 62, pp. 43–53, Jun. 2014.
9. Huang S, Z. Niu, and C. Shi, 'Automatic construction of domain-specific sentiment lexicon based on constrained label propagation', Knowl.-Based Syst., vol. 56, pp. 191–200, Jan. 2014.
10. Kouloumpis E, T.Wilson and J.D.Moore, "Twitter sentiment analysis: the good, the bad and the omg!", in Proceedings of ICWSM 2011, USA, 2011, pp. 538-541.
11. Liu Y, X. Yu, B. Liu, and Z. Chen, 'Sentence-Level Sentiment Analysis in the Presence of Modalities', in Computational Linguistics and Intelligent Text Processing, A. Gelbukh, Ed. Springer Berlin Heidelberg, 2014, pp. 1–16.
12. Montejo-Ráez A, E. Martínez-Cámara, M. T. Martín-Valdivia, and L. A. Ureña-López, 'A knowledge-based approach for polarity classification in Twitter', J. Assoc. Inf. Sci. Technol., vol. 65, no. 2, pp. 414–425, Feb. 2014.
13. Moraes R, J. F. Valiati, and W. P. GaviãoNeto, 'Document-level sentiment classification: An empirical comparison between SVM and ANN', Expert Syst. Appl., vol. 40, no. 2, pp. 621– 633, Feb. 2013.
14. Onan A, "Twitter mesajları üzerinde makine öğrenmesi yöntemlerine dayalı duygu analizi", Yönetim Bilişim Sistemleri Dergisi, Vol. 3, No. 2, 2017, pp. 1-14.
15. Pak A and P.Paroubek, "Twitter as a corpus for sentiment analysis and opinion mining", in Proceedings of LREC 2010, USA, 2010, pp. 1320-1326.
16. Peñalver-Martinez I, F. Garcia-Sanchez, R. Valencia-Garcia, M. Á. Rodríguez-García, V. Moreno, A. Fraga, and J. L. Sánchez-Cervantes, 'Feature-based opinion mining through ontologies', Expert Syst. Appl., vol. 41, no. 13, pp. 5995–6008, Oct. 2014.
17. Salas-Zarate M.P, E.Lopez-Lopez, R.Valencia-Garcia, N. Gilles, A.Almela and G.Alor-Hernandez, "A study on LIWC categories for opinion mining in Spanish reviews", Journal of Data Science, Vol.40, No.6, 2014, pp.749-760.
18. Saif H, Y.He and H.Alani, "Semantic sentiment analysis of twitter", in Proceedings of ISWC 2012, USA, 2012, pp.508-524.
19. Singhal K, B. Agrawal, and N. Mittal, 'Modeling Indian General Elections: Sentiment Analysis of Political Twitter Data', in Data Systems Design and Intelligent Applications, J. K. Mandal, S. C. Satapathy, M. K. Sanyal, P. P. Sarkar, and A. Mukhopadhyay, Eds. Springer India, 2015, pp. 469–477.
20. Xia R, F. Xu, J. Yu, Y. Qi, and E. Cambria, 'Polarity shift detection, elimination and ensemble: A three-stage model for document-level sentiment analysis', Inf. Process. Manag., 2015.

AUTHORS PROFILE

B. S. **Srinidhi**, is a research scholar in Jain University. He is working on the text classification and Sentiment analysis. He is also a web developer with .Net technologies His interests are in the area of natural language processing and Emotional Intelligence , He find it very interesting to discover patterns in the data relating to Emotional behavior of people.

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