ISSN 2278-3091
Volume 9 No.2, March -April 2020
International Journal of Advanced Trends in Computer Science and Engineering
Available Online at http://www.warse.org/IJATCSE/static/pdf/file/ijatcse114922020.pdf
https://doi.org/10.30534/ijatcse/2020/114922020

Multi-Labelled Emotion with Intensity Based Sentiment Classification Model in Tweets using Convolution Neural Networks

V. Ganesh[1], Dr. M. Kamarason[2]
[1]Research Scholar, Department of Computer and Information Science, Annamalai University, Chidambaram, India, dinesh59@gmail.com
[2]Assistant Professor, Department of Computer and Information Science, Annamalai University, Chidambaram, India,smkrasan@yahoo.com

ABSTRACT

In present days, people utilize social networking sites like Twitter for sharing their emotions and opinions. Detection and examination of the emotions present in the social media becomes advantageous in several application areas like e-commerce, government, public health, entertainment and so on. Several earlier studies have on sentiments and emotions classification has concentrated only on single label classification and does not considered the co-occurrence of many emotion classes in one tweet or post. This study presents a new multi-labeled emotion with intensity based sentiment classification in twitter data using convolutional neural networks called CNN-EISC model. The proposed CNN-EISC method has been validated using SEMEVAL2018 Task-1Emotion Intensity Ordinal Classification dataset. The CNN-EISC model has classified a set of four intensities namely Anger, Sad, Fear and Joy under diverse classes (0, 1, 2 and 3). The obtained experimental outcome stated that the CNN-EISC model has exhibited maximum results with the average precision of 77.08%, recall of 62.52%, F-measure of 65.76% and accuracy of 82.72%.

Key words: CNN, Data Classification, Sentiment Analysis, Twitter

1. INTRODUCTION

Emotions are referred as a key or signal of human's opinions and feelings. Internet and social networking sites like Twitterhas made several adjustments in the way of human communication. At recent times, peoples share the knowledge about current facts, opinions, feelings, and emotional intensities on diverse topics. It has been widely applied in extensive applications like business, education, and so on. For example, it might be utilized in healthcare [1], public opinion prediction regarding politics [2-4], as well as stock exchange observation [5]. Emotion detection is defined as process of computing the behavior under diverse topics. The attitude may be either positive or negative or even emotional state like happy, sad, and angry [6]. Basically, a multi-label classifying issue has been existed because of wider application, such as text, video classification, as well as bioinformatics [7]. In contrary to the conventional single-label classifying issue, multiclass or binary classification is preferred which is related to a set of class labels.

Few classical problems of sentiment analysis (SA) and emotion recognition have intended on single label classification. Several earlier studies have does not considered the co-occurrence of many emotion labels in one tweet or post. Thus, in this approach, we mainly concentrates on multi-label emotion classification, that has the objective to deploy an automated system to compute the presence of different emotions of 8 [8] classes like joy, sad, anger, fear, trust, disgust, surprise, anticipation, optimism, pessimism, and love as depicted in Fig. 1.In order to report the issues, multi-label classification has been deployed in problem conversion. Under the application of this method, a multi-label issue were converted to more than one single-label problem. In general, single-label classification were learned and implemented; and the classifier predictions were modified as multi-label predictions. Various transformation modules are projected in multi-label study. A typical approach is named as binary relevance [9,10]. The suggestion behind binary relevance was easy and instinctive. A multilabel issue were converted as a multi-binary problem.

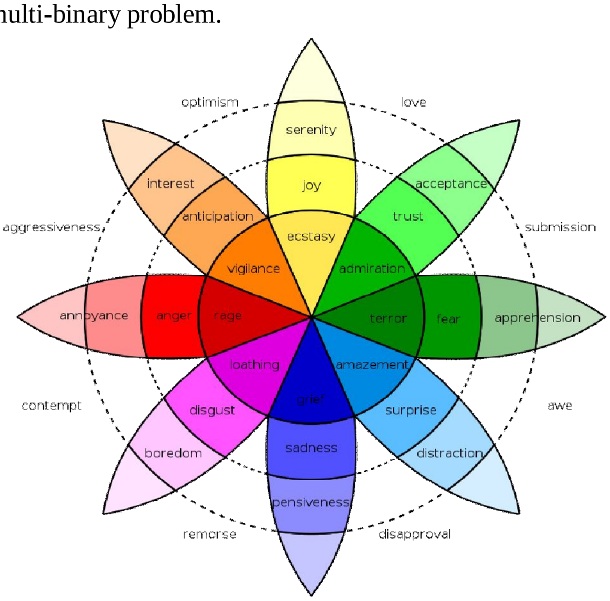

Figure 1: Different kinds of emotions [8]

Followed by, the autonomous binary classifier undergoes training for detecting the relevance of labels. Though binary relevance is a well-known model, because of the simplicity, it faces the straight forwarding modelling associations which exist among the labels.

Several Machine Learning (ML) methods have been presented for conventional emotion classification as well as multi-label emotion classification. Various traditional models are applied in resolving the issue as a text classifying issue. Supervised classification undergoes training on the collection of annotated corpora under the application of diverse hand-based features. The efficiency of these techniques depends upon 2 major aspects like maximum number of labelled data as well as smart deployment of features which could differentiate the samples. By the application of these approaches, several works concentrated on engineering the collection of effective features to attain optimal classifying function [11]. The main aim is to locate a group of applicable features to mimic the emotion depicted in content. Bag-of-Words (BoW) and difference, n-grams, are named as representation approach employed to classify text and SA.

Diverse models are integrated with BoW features like parts of speech (PoS) tags, the sentiment data obtained from lexicons, statistics data, and word structure to highlight the text expression. Even though BoW is a well-knownmodel for classifying textual data, it is also constrained with few limitations. Initially, it removes the word sequence which refers that 2 documents might have similar expression, though it has diverse semantics. The n-gram approach is applicable in solving these limitations of BoW by assuming the arrangement of words in length n. But, it faces the problems from sparsity as well as greater dimensionality. Alternatively, BoW is scarcely capable of developing semantics of words. For instance, the terms beautiful, wonderful and view are comprised with similar distance in BoW, in which the word beautiful is identical to word wonderful when compared with word view. The sentiment and emotion lexicons are more required in deploying successful sentiments as well as emotion analyzing scheme. Therefore, it is very hard to develop these lexicons. Furthermore, identifying the optimized integrations of lexicons and best collection of statistical features consumes higher duration. At recent times, Deep Learning (DL) framework is applied to deploy end-to-end systems in massive operations such as audio analysis, text and image classification, etc. It is shown that, these modules are capable of extracting high-level features from actual data in an automated manner [12, 13].

[14] performed a process of identifying multiple emotion states of SemEval-2018 Task1: Affects in Tweets. It has been trained with word2vec model with higher words obtained from a dataset comprises millions of tweets. The alternate position of SemEvalleaderboard provides the training to word-level bidirectional LSTM, and non-deep learning features are ensemble [15]. [16] trained 2 approaches to resolve this issue: regularized Linear Regression and Logistic Regression (LR) classification model [10]. It is applied to exploit labels' associations to process multi-label classification task. Using the former method, the developers formed a multilabel classification issue as a linear regression using distance of regularization term.

The main aim of this model is to assume the multi-label problem in the form of series of binary classification issues by existing classifier as additional input for upcoming classification process. Here, the DL-centric scheme was employed to create a system which is capable of extracting higher level expression of tweets and develop the inherent higher order relationship. Hence, the application of newly proposed model with transformation approach is used in training and resolves the issue of multi-label emotion classification in tweets. This paper introduces a multi-labeled deep learning based emotion with intensity based sentiment classification in twitter data. The proposed CNN-EISC method makes use of CNN model to identify the emotions and intensity like Anger, Sad, Fear and Joy. The presented CNN-EISC model has been validated using SEMEVAL2018 Task-1Emotion Intensity Ordinal Classification dataset.

The remaining section of the paper is formulated as follows. The CNN-EISC model is discussed in Section 3 followed by simulation analysis in Section 3. At last, the conclusions are drawn in Section 4.

2. THE PROPOSED METHOD

The overall working principle of the proposed CNN-EISC model is shown in Fig. 2. Initially, the tweets undergo preprocessing to make the tweets appropriate for further processing. Then, CNN based classification process takes place to identify the different intensities with diverse classes exist in the tweets. The processes are discussed in the following subsections.

2.1 Preprocessing

The actual tweets, obtained from Twitter, are composed with noise with respect to irregular and fuzzy words, URLs, stop-words, and so on, that has to be minimized in prior feature extraction. Hence, the presented model applies the pre-processing approach in 2 phases in prior to process the extract features:

A. Phase 1

The initial phase removes irregular noise from Twitter data set by the application of given procedure:

- Remove each URLs by regular expression mapping. A regular expression is referred as a textual pattern which describes the searching pattern for text. It may be applied in searching for URLs, email address, and so forth.
- Interchange "@Username" with "usr" by applying regular expression matching.

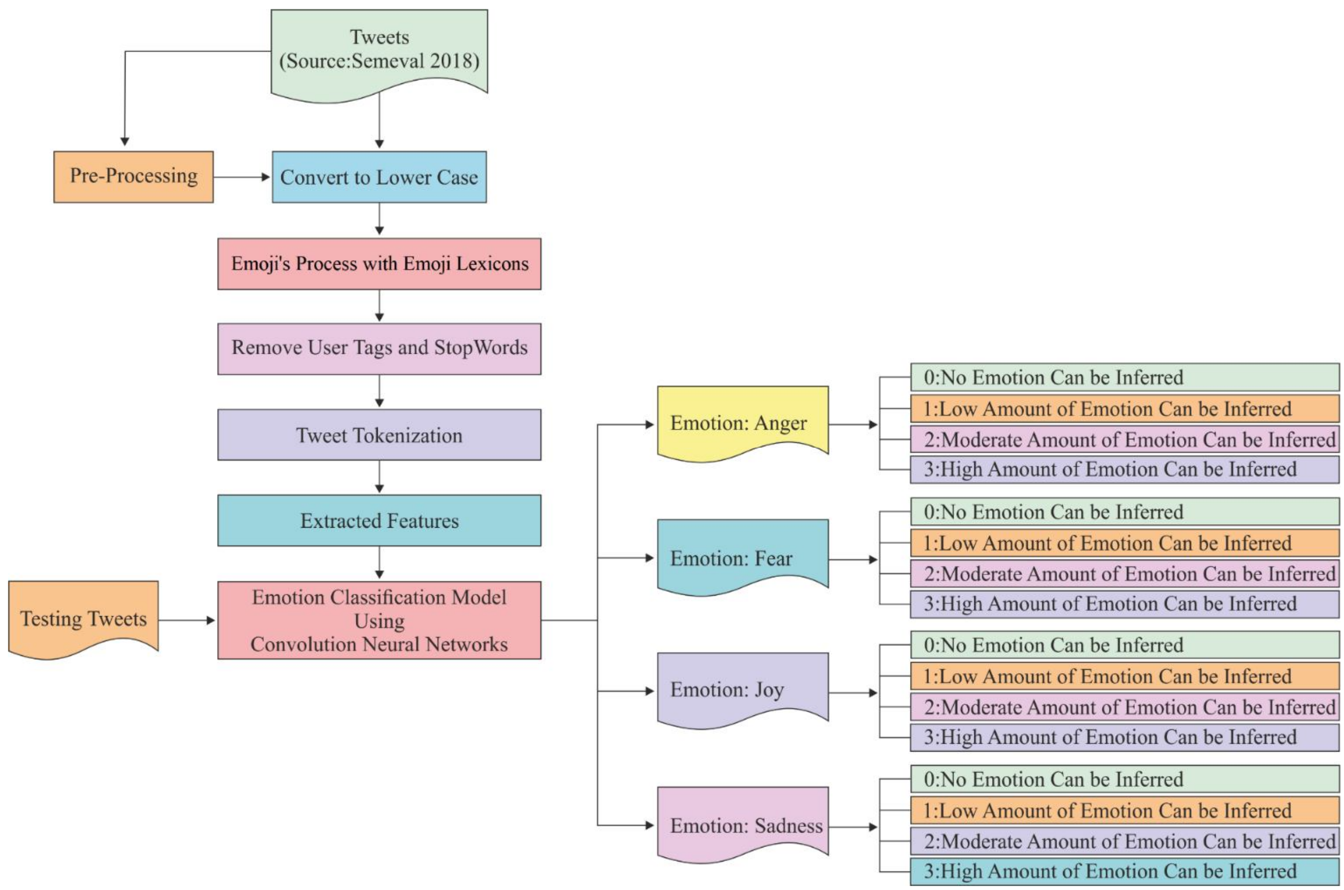

Figure 2: Block diagram of CNN-EISC Model

- As "hash-tag(#)" gives few informative data, hence eliminates #, by retaining the word. through, "#Lee" is interchanged with "Lee".
- Eliminate parenthesis, forward slash (/), backward slash (\), and dash from tweets.
- Replace several white spaces with single white space.

B. Phase 2

Here, 2 dictionaries such as; stop word and acronym has applied to enhance the precision of final Twitter dataset of Phase 1. The procedures of Phase 2 are given below:

- Replace the Emojis into Unicode and then Unicode representation is converted into lexicons. For instance, the emoji □ is converted into a Unicode of U+1F600 which is then assigned to a lexicon of "Happy Face"
- Transform every tweet word into lowercase.
- Reject every stop words like, a, is, the, etc. by relating with stop word dictionary.
- Interchange the series of sane characters in a word by single character by, "hellooooo" is transformed to "Hello".
- Replace every short forms in corresponding full forms by applying acronym dictionary

2.2. Tweet Tokenization

A token is a word in a sentence, and a sentence is defined as token in a paragraph. Tokenization is defined as the task of dividing a string into list of tokens**.**Here, Word2vec method is applied to get distributed representations of words. It is comprised with CBOW method as well as Skip-gram method. These approaches are constrained with input layer, projection layer and output layer. CBOW framework detects the desired words on the basis of context distribution. For word wk, it is represented as given in the following:

$$context\ (w_k) = \{w_{k-1}, w_{k-(t-1)}, \dots, w_{k+(t-1)}, w_{k+t}\} \quad (1)$$

Unlike, the Skip-gram approach is applied to detect the context according to target word wk. TF-IDF is evolved from the concatenation of TF and IDF weight estimation modules. It is generally applied weight estimation model in text classification. The sequence of a word from a individual document as well as the word distribution is assumed in this approach. Also, it has the significance of feature in classification task.

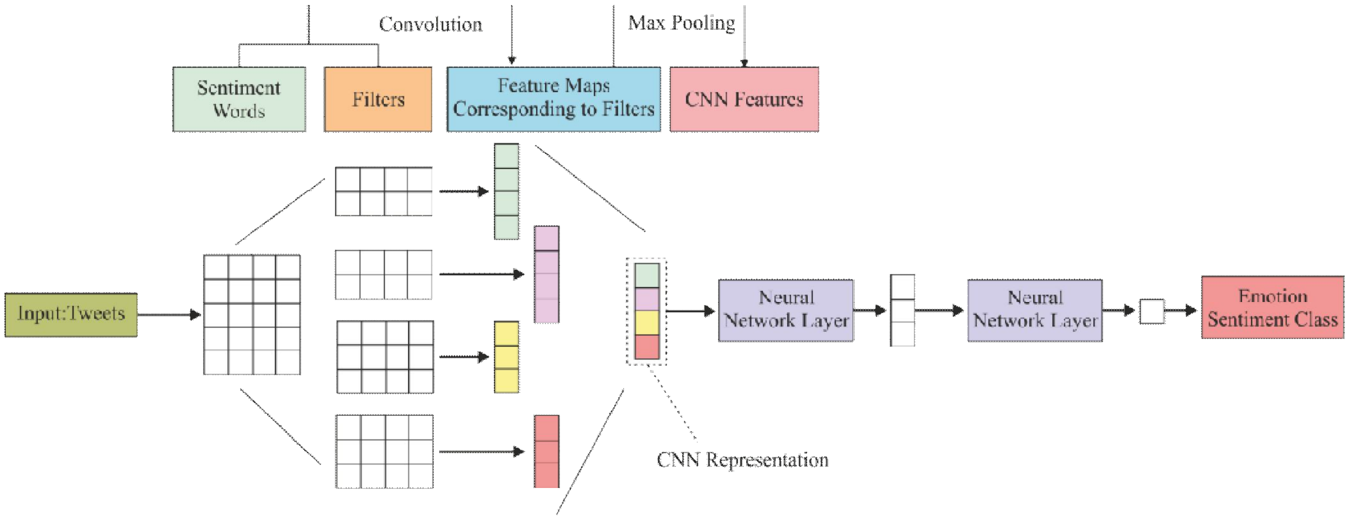

Figure 3: Architecture of CNN for emotion sentiment classification

The functions of TF-IDF are provided below:

$$w(t_i, d) = \frac{tf(t_i, d) \times idf(t_i)}{\sqrt{\sum_{t_i \in d}[tf(t_i, d) \times idf(t_i)]^2}} \quad (2)$$

$$idf(t_i) = \log(N/n_{ti}) + 1 \quad (3)$$

$w(t_i, d)$ is a weight of word t_i in document d, $tf(t_i, d)$ signifies the frequency, N implies the overall count of documents and n_{ti}represents the count of documents in word t_i.

2.3. CNN based Classification Process

CNN is extensively applied on image databases, which obtains the vital features in the form of "convolutional" filter, named as kernel that transmits by an image. When the input data are provides as 1D, the similar performance of CNN has been processed. At the time of moving a filter, local data were saved, and significant features were obtained. The application of CNN in text classification is efficient. Fig. 3 demonstrates a graphical presentation of newly developed CNN model for emotion classification. This model is composed with embedding layer, 2 convolutional, pooling, and fully-connected (FC) layers. The lengthy sentences were reduced to a definite type using a token.

A fixed length S has been allocated to higher length of a sentence. An embedding layer is applicable in mapping every word in a sentence to E-dimensional feature vector results in S × E matrix, where E is the embedding size. Besides, the embedding layer is a task of placing words which are in the form of input into well-developed space; in which words with identical meanings are placed nearby words with dissimilar meanings are placed far away.

Also, embedding is an operation of presenting 2D matrix into E-dimension to attain a word vector. The embedding vectors could be reached from alternate resources. In this work, the newly deployed embedding layer were extracted by a training process, and every word tokens for unknown words is transformed to arithmetical values with the application of embedding layer.

The S × E matrix, the result of embedding layer, has been projected as primary conv layer. It contains three CNNs, with conv layer and max-pooling layer. The CNNs applies similar activation function on conv layer and use the similar count of filters to incoming sentence matrix; therefore, diverse size of convolutional filters as well as size of max-pooling filters. The role of conv layer is to employ filtering for sentence matrix. Filters slide across complete rows of matrix words-to understand meaningful expressions. It is applicable in producing feature maps as same as filters' value as every filter generates a feature map, that is passed to a max-pooling layer. The first conv layer is C1 × E matrix, that saves the local data which has to be classified as sentiment class in S × E matrix and pass the data to upcoming conv layer. The C1 × E matrix slides each value of S × E matrix with random stride, that estimates the dot product, and provide the outcome of a dot product to the next layer. The alternate conv layer applies C2 × 1 matrix for extracting features from contextual data of major word depends upon local data recorded in a first conv layer. C1 and C2 implies the filter size of every conv layer, and 2 conv layers are constrained with K1 and K2 distinct filters, correspondingly, for capturing exclusive contextual data.

On the other hand, first conv layer has been applied to look at elegant contextual data whereas looking for S × E matrix, and second conv layer is employed for capturing main features and extracting poor, and optimal which has sentiments that affects classification task. The matrix which is passed by subsequent conv layer is utilized as input to pooling layer. During the average-pooling as well as L2-norm pooling is

applied as pooling layer position that is choosing maximum value as a representative of peripheral values.
As the sentiment has been computed using various words instead of showing sentiment in all words of a sentence, here it is applied with max-pooling method. The pooling layer slides every values of a matrix that is the result of second conv layer, using a random stride, that result is resultant vectors. As max-pooling is a layer which passes to a next layer, higher value between various values, it produce the vectors of tiny size. The conv layer seeks at the content and obtains major features, and pooling layer is responsible for selecting main featuresOnce the pooling layer is applied, a flattening process has been carried out to transform 2D feature map to 1D format and an F-dimensional FC layer. As the FC layer applies 1D vector as input, the 2D vector is provided from a pooling layer which has to be flattened.

The FC layer correlates every input as well as output neurons. A vector which has been passed by a FC layer creates an output which can be categorized as positive and negative. The activation function softmax performs for classifying various classes in FC layer. The softmax function produces the values with a probability rate, as produced for every class. Due to the network variation, it depends upon the data length and number of features, also, data passed by 2 subsequent conv layers it is passed by a pooling layer which is effective in saving the contextual data as well as obtaining important features.

3. PERFORMANCE VALIDATION

For ensuring the effective performance of the proposed model, a series of experiments were carried out using Python Programming language. The applied dataset and the attained results are examined under several subsections.

3.1. Dataset

To validate the outcome of the proposed CNN-EISC model, a benchmark SEMEVAL2018 Task-1Emotion Intensity Ordinal Classification dataset is utilized [17]. It comprises a collection of 4042 tweets with four intestines namely joy, fear, anger, and sadness. From the available tweets, 1101 tweets comes under the category of Joy, 983 tweets comes under the category of fear, 991 tweets comes under the category of anger and 967 tweets comes under the category of sadness. The dataset is split into training and testing data in the ratio of 7.5:2.5.

3.2. Results Analysis

Table 1 has provided the attained classification results of affect dimension using CNN-EISC model under several measures. The table values indicated that the instances under anger class are effectively classified with the precision of 65.18%, recall of 88.39%, F-measure of 75.03% and accuracy of 79.48%. At the same time, the instances under fear class are successfully classified with the precision of 76.17%, recall of 57.54%, F-measure of 65.55% and accuracy of 85.17%. Along with that, the instances under joy class are effectively classified with the precision of 65.49%, recall of 74.02%, F-measure of 69.49% and accuracy of 76.38%. Likewise, the instances under sadness class are effectively classified with the precision of 96.38%, recall of 38.78%, F-measure of 54.89% and accuracy of 86.57%.On average, the proposed model has achieved a higher precision of 75.81%, recall of 64.68%,F-measure of 66.24% and accuracy of 81.90%.

Table 1: Performance Measures of Affect Dimension using Proposed CNN

Measures	Precision	Recall	F-Measure	Accuracy
Anger	65.18	88.39	75.03	79.48
Fear	76.17	57.54	65.55	85.17
Joy	65.49	74.02	69.49	76.38
Sadness	96.38	38.78	54.89	86.57
Average	**75.81**	**64.68**	**66.24**	**81.90**

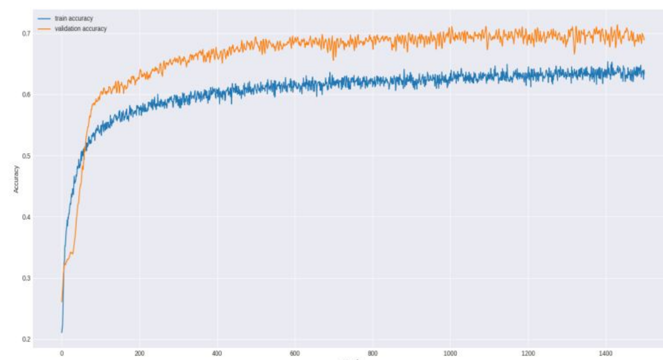

Figure 4: Accuracy Analysis of Training and Validation at the time of Model Creation

Fig. 4 shows the accuracy graph attained during the training and validation process of the model creation under varying number of epochs. The figure clearly stated that the validation accuracy is slightly higher than the training accuracy. Besides, it is observed that accuracy gets increased with a rise in number of epochs.

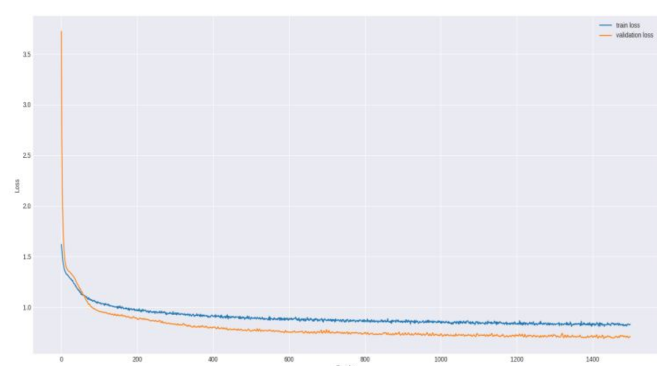

Figure 5: Loss Graph of Training and Validation at the time of Model Creation

Fig. 5 shows the loss graph attained during the training and validation process of the model creation under varying number of epochs. The figure clearly stated that the validation loss is slightly higher than the training loss. Besides, it is observed that loss gets increased with a rise in number of epochs.

An average analysis of the results produced by the CNN-EISC model has been analyzed under different intensities as provided in Figs. 6-7. It is shown that the affect dimensions are classified with the average precision of 75.81%, recall of 64.68%, F-measure of 66.24% and accuracy of 81.90%. Simultaneously, the CNN-EISC model has classified the classes under anger intensity with the average precision of 80.11%, recall of 59.73%, F-measure of 64.32% and accuracy of 83.64%. %. Concurrently, the CNN-EISC model has classified the classes under fear intensity with the average precision of 88.33%, recall of 69.14%, F-measure of 76.04% and accuracy of 90.97%. %.

In the same line, the CNN-EISC model has classified the classes under joy intensity with the average precision of 63.20%, recall of 49.75%, F-measure of 50.81% and accuracy of 72%. Along with that, the CNN-EISC model has classified the classes under sadness intensity with the average precision of 77.96%, recall of 69.32%, F-measure of 71.41% and accuracy of 85.09%. In overall, the CNN-EISC model has classified the applied dataset with the maximum average precision of 77.08%, recall of 62.52%, F-measure of 65.76% and accuracy of 82.72%.

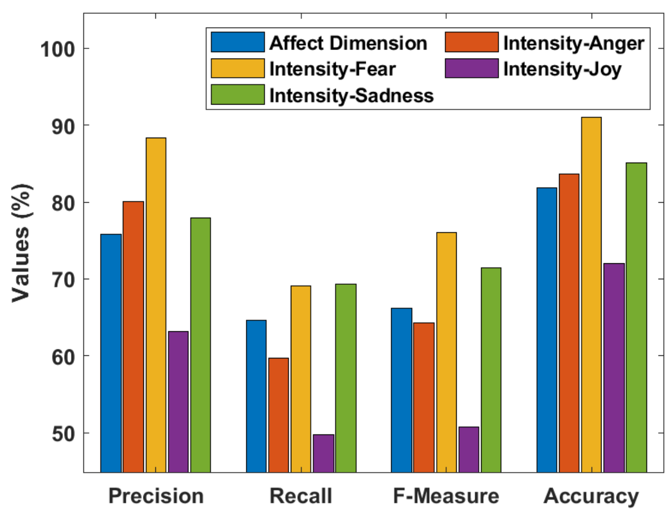

Figure 6: Classifier result analysis of CNN-EISC model

Fig. 8 has showcased the outcome of the detailed comparative analysis of the CNN-EISC model with the recently presented models in terms of accuracy. From the figure, it is obviously clear that the linear SVC model has provides worse outcome with the least accuracy value of 48.90%.

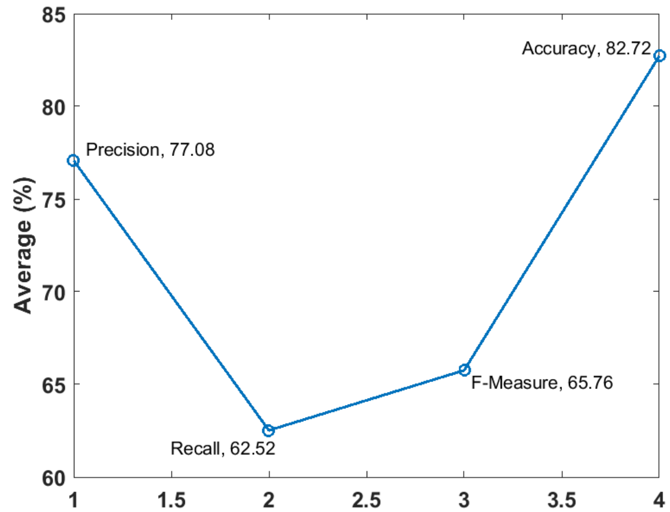

Figure 7: Average result analysis of CNN-EISC model

Besides, it is apparent that the GRU and Context Aware models has resulted to ineffective results with the minimal accuracy values of 52.40% and 53.20% respectively. Also, the MNB and RF models has tried to show slightly better and near identical results by offering accuracy values of 54.70% and 54% respectively. Likewise, the methods presented by MohameedJabreel et al., Mondher Bouazizi et al., and Malak Abdullah et al. leads to closer and moderate results with the accuracy of 59%, 60.20% and 59.90%. Moreover, the MLR model has outperformed all the earlier models by providing a high accuracy of 62.27%.However, the presented CNN-EISC model has reached to an optimal classification results and achieved a maximum accuracy of 82.72%.

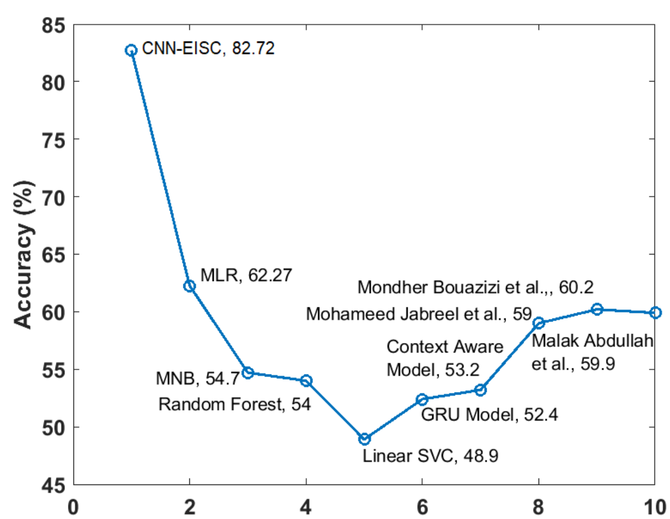

Figure 8: Comparative accuracy analysis of CNN-EISC with existing models

From the above-mentioned extensive experimental analysis, it is evident that the proposed CNN-EISC model has outperformed all the existing methods in a significant manner. Therefore, it can be considered as an appropriate tool for the emotion classification of content in social networking sites like Twitter, Facebook and so on.

4. CONCLUSION

This paper has developed a new multi-labeled emotion with intensity based sentiment classification in twitter data using convolutional neural networks called CNN-EISC model. The proposed CNN-EISC method makes use of CNN model to identify the emotions and intensity like Anger, Sad, Fear and Joy. Initially, the tweets undergo preprocessing to make the tweets appropriate for further processing. Then, CNN based classification process takes place to identify the different intensities with diverse classes exist in the tweets. The proposed CNN-EISC method has been validated using SEMEVAL2018 Task-1Emotion Intensity Ordinal Classification dataset. The obtained experimental outcome stated that the CNN-EISC model has exhibited maximum results with the average precision of 77.08%, recall of 62.52%, F-measure of 65.76% and accuracy of 82.72%. In future, the performance of the CNN-EISC model can be enhanced by the use of feature extraction techniques.

REFERENCES

1. Y. Chen, Y. Zhou, S. Zhu, and H. Xu. **Detecting Offensive Language in Social Media to Protect Adolescent Online Safety,***In Proceedings of the 2012 International Conference on Privacy, Security, Risk and Trust (PASSAT)*, Amsterdam, 2012, pp. 71–80. https://doi.org/10.1109/SocialCom-PASSAT.2012.55
2. S.M. Mohammad, X. Zhu, S. Kiritchenko, and J. Martin**. Sentiment, Emotion, Purpose, and Style in Electoral Tweets,***Inf. Process. Manag.,* Vol. 51, pp. 480–499, 2015. https://doi.org/10.1016/j.ipm.2014.09.003
3. E. Cambria. **Affective computing and sentiment analysis,***IEEE Intell. Syst.*, Vol. 31, pp. 102–107, 2016. https://doi.org/10.1109/MIS.2016.31
4. M. Jabreel, A. Moreno, and A. Huertas, **A. Do Local Residents and Visitors Express the Same Sentiments on Destinations Through Social Media?,***In Information and Communication Technologies in Tourism;* Springer: New York, 2017, pp. 655–668. https://doi.org/10.1007/978-3-319-51168-9_47
5. T. Rupa Rani, G. Lavanya, Dr. H. Shaheen, R. Anusha, and K. Nartkannai, **Open IOT Service Platform Technology with Semantic Web**, *International Journal of Advanced Trends in Computer Science and Engineering,* Vol. 9, pp.17-21, 2020. https://doi.org/10.30534/ijatcse/2020/0491.12020
6. Gleen A. Dalaorao, Ariel M. Sison, Ruji P. Medina,**Applying Modified TF-IDF with Collocation in Classifying Disaster-Related Tweets,***International Journal of Advanced Trends in Computer Science and Engineering,* Vo. 9, pp. 28-33, 2020. https://doi.org/10.30534/ijatcse/2020/0691.12020
7. Shuhaida Mohamed Shuhidan, Nur AiniShuhadaFadzil and SaidatulRahah Hamidi, **The Visualization of Sentiment Analysis for Malaysia Financial News Headlines,***International Journal of Advanced Trends in Computer Science and Engineering,* Vol. 9, pp. 589-594, 2020.
8. R. Plutchik, **A general psychoevolutionary theory of emotion,***In Theories of emotion,* Academic press, 1980, pp. 3-33.
9. G. Tsoumakas, and I. Katakis, **Multi-label Classification: An Overview,** Int. J. Data Warehous. Min. IJDWM, Vol. 3, pp. 1–13, 2007. https://doi.org/10.4018/jdwm.2007070101
10. J. Read, **Scalable Multi-label Classification**, Ph.D. Thesis, University of Waikato, Hamilton, New Zealand, 2010.
11. M. Jabreel, and A. Moreno, **SentiRich: Sentiment Analysis of Tweets Based on a Rich Set of Features,** *In Artificial Intelligence Research and Development,* Vol. 288, pp. 137–146, 2007.
12. Y. LeCun, Y. Bengio, and G. Hinton, **Deep Learning**, *Nature*, Vol. 521, pp. 436–444, 2015.
13. D. Tang, B. Qin, T. Liu, D**eep Learning for Sentiment Analysis: Successful Approaches and Future Challenges,** *Wiley Interdiscip. Rev. Data Min. Knowl. Discov.* Vol. 5, pp. 292–303, 2015. https://doi.org/10.1002/widm.1171
14. C. Baziotis, N. Athanasiou, A. Chronopoulou, A. Kolovou, G. Paraskevopoulos, N. Ellinas, S. Narayanan, and A. Potamianos, **NTUA-SLP at SemEval-2018 Task 1: Predicting Affective Content in Tweets with Deep Attentive RNNs and Transfer Learning,***In Proceedings of the 12th International Workshop on Semantic Evaluation*, New Orleans, LA, 2018, pp. 245–255.
15. H. Meisheri, and L. Dey, **TCS Research at Semeval2018 Task 1: Learning Robust Representations using Multi-Attention Architecture,***In Proceedings of the 12th International Workshop on Semantic Evaluation*, New Orleans, LA, USA, 2018, pp. 291–299. https://doi.org/10.18653/v1/S18-1043
16. J.H.Park, P. Xu, and P. Fung, **PlusEmo2Vec at SemEval-2018 Task 1: Exploiting Emotion Knowledge from Emoji and #hashtags,** In Proceedings of the 12th International Workshop on Semantic Evaluation, New Orleans, LA, 2018, pp. 264–272. https://doi.org/10.18653/v1/S18-1039