Amazon Fake or Spam Review Classification Using Short Text Processing Technique

Sunny M. Ramchandani[1], Dr. Hemant H. Patel[2]

[1]PG Student, Department of Computer Science and Engineering, Dr. Subhash University, Junagadh, Gujarat, India.
[2]Associate Professor and Head, Department of Computer Science and Engineering, Dr. Subhash University, Junagadh, Gujarat, India.

---***---

Abstract - *Text classification is a classical research domain utilized for various applications in education, medical and government. However, traditional text classification task is different from the new age text classification problem. In this presented work, the text classification is investigated for classifying the fake or spam review in an e-commerce platform. The e-commerce product vendors are sometimes utilizing fake and partial reviews to boost sales of their own low-quality products in e-commerce platform. This act will waste the consumer's time, and also negatively impact the e-commerce credibility. Therefore, identification and removal of such misleading reviews from the e-commerce platform. In this work, the text classification technique is used for classifying the spam reviews in e-commerce platform. For this purpose, first text pre-processing technique is used to make clean the reviews, next the word two vector technique is used to prepare the training and validation samples. For conducting the experiments amazon product review dataset has been used. In this dataset different category of product reviews is available, among them the toy and game product category is considered. Additionally, to perform the classification task, two popular machine learning algorithms namely Convolutional Neural Network (CNN) and Long Short-Term Memory (LSTM) have been used. After successfully implementation the performance in terms of precision, recall, f-measure and accuracy has been measured. Based on experimental results CNN is providing 90% classification accuracy then LSTM, which provides 89.31% accuracy. Additionally, the CNN is efficient than LSTM model in terms of training time. Therefore, it is recommended to use CNN model for future implementations.*

Key Words: Deep Learning, E-commerce review, Machine Learning, small text classification, Text classification.

1.INTRODUCTION

The text analysis is a traditional domain of research in academics. There are a number of applications has been developed using text analysis techniques such as information retrieval (IR) and search engine [1]. But, due to increase in the communication mediums various sources of text information have been developed such as social media, and e-commerce [2]. In both the channels a significant amount of text has been generated and analysis of such information can help in various real-world applications such as disaster management, social gathering and others [3]. However, the traditional text analysis task is different from the social media or e-commerce-based text analysis. The amount of text in these type of text data is fewer but having high impact on social media and also in e-commerce platform [4].

In this presented work, a task of sentiment analysis has been done on e-commerce platform data specifically on reviews. The aim is to identify the spam reviews which are influencing user's or buyers' decisions [5]. Because sometimes of the time e-commerce sellers are utilizing the fake or spam review to lure the buyers during new product launch. Most of the time a new buyer in e-commerce platform utilizes the reviews for making a buying decision [6]. In this context, the fake review can impact on the buying decisions. Therefore, it is essential to identify and remove such kind of fake reviews using the sentiment analysis task. In this section, the basic overview of the proposed work has been discussed. Additionally, the next section includes the motivation of the proposed work.

2. PROPOSED WORK

In this section, the architecture of the proposed system has been discussed. Additionally, the components of the proposed system are also explained. Using these components the functional requirement of each component is discussed.

2.1 System overview

In order to understand the working of the sentiment-based text classification techniques recently a review has been carried out. In this review, the different research articles are involved based on machine learning and text classification. The aim is to study different available technique of classifying the spam reviews in an e-commerce platform. Therefore, a detailed investigation of the collected research articles has been performed. Additionally, based on the analysis it is recognized that, the traditional text classification techniques are different from the sentiment-based text classification. In addition, the sentiment-based text classification techniques require a smaller number of features as compared to the traditional methods of text classification. Therefore, why the traditional text feature selection techniques are used in the emotion classification. It is required to investigate.

In this context, the proposed study has been focused on investigating the different text data pre-processing and feature selection techniques. The aim is to find the appropriate text data analysis methods for emotion-based text classification. Therefore, an experimental investigation of the sentiment-based text classification is proposed in this study. By motivation of the required technique, a model has been proposed for future design and implementation. The given model is providing the understanding about the sentiment text analysis process. In addition, the basic requirements of the components.

2.2 Proposed system

The main aim of the proposed wok is to accurately classify the spam review in an e-commerce platform. The identification and elimination of fake reviews are essential to deal with the partial reviews which are falsely describing the quality of product. Therefore, a machine learning model has been presented in this section. Additionally, the concept for implementation is demonstrated in Fig-1.

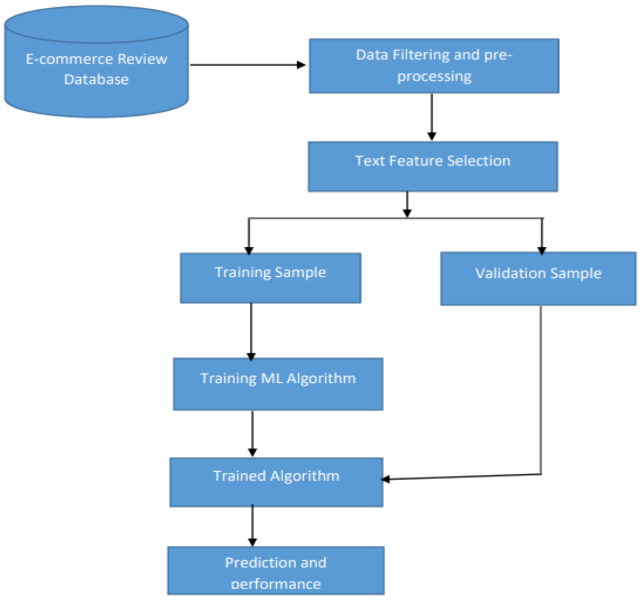

Fig- 1: Proposed system architecture

2.2.1 E-commerce Review Database

According to this diagram, the system is initiated with the dataset consist of user reviews. Actually, in any machine learning model for training a set of historical records are required. The algorithms are utilizing the historical records to train and build a model. In this presented work amazon product dataset has been used, which is downloaded from the Kaggle repository. This dataset has multiple categories of products and their reviews. In this work, the toy and games category of product reviews have been downloaded. The downloaded file has more than 1.25 lacks instances. Due to limitations of computational resources the 30k instances of the dataset has been used for experimentation. The dataset instances have made with 12 attributes. Among them 11 attributes are defining the properties of product and review and 1 is class label for training. The raw sample of the dataset is given in Fig - 2.

_id	reviewerID	asin	reviewerName	helpful	reviewText	overall	summary	unixReviewTime	reviewTime	category	class
{'$oid': ...13282b741a2384e879a620'}	A3C9CSW3TJITGT	0005069491	Renee	[0, 0]	I love these felt nursery rhyme characters and...	4	Charming characters but busy work required	1377561600	08 27, 2013	Toys_and_Games	1
{'$oid': ...13282b741a2384e879a621'}	A31POTIYCKSZ9G	0076561046	So CA Teacher	[0, 0]	I see no directions for its use. Therefore I h...	3	No directions for use...	1404864000	07 9, 2014	Toys_and_Games	0
{'$oid': ...13282b741a2384e879a622'}	A2GGHHME9B6W4O	0131358936	Dalilah G.	[0, 0]	This is a great tool for any teacher using the...	5	Great CD-ROM	1382400000	10 22, 2013	Toys_and_Games	1
{'$oid':	AMEVO2LY6VEJA	0000191639	Nicole Soeder	[0, 0]	Great product, thank you!	5	Puzzles	1388016000	12 26, 2013	Toys_and_Games	1

Fig- 2: Dataset first four instances

2.2.2 Data pre-processing and filtering

For better training of the machine learning algorithms, it is essential to utilize high quality and authentic data for learning. Therefore, the dataset is used with different cleaning operations. This process of data cleaning is known as pre-processing steps, which is responsible for improving the quality of learning dataset.

	reviewText	class
0	I love these felt nursery rhyme characters and...	1
1	I see no directions for its use. Therefore I h...	0
2	This is a great tool for any teacher using the...	1
3	Great product, thank you! Our son loved the pu...	1
4	Although not as streamlined as the Algebra I m...	1

Fig- 3: separated attributes for experiments

The pre-processing techniques are also helping to transform data for suitable data representation. Therefore, first the unwanted dataset attributes have been eliminated from the dataset. Additionally, the attributes namely 'review text' and 'class' is separated for further pre-processing. The extracted data is given in Fig-3. The review text has been also pre-processed using the text pre-processing technique. Therefore, two functions have been implemented first is used for eliminating the punctuations from the text and second function is used for removing the numerical digits. After applying both the functions on data the final prepared data is given in Fig-4.

	reviewText	class
0	I love these felt nursery rhyme characters and...	1
1	I see no directions for its use Therefore I ha...	0
2	This is a great tool for any teacher using the...	1
3	Great product thank you Our son loved the puzz...	1
4	Although not as streamlined as the Algebra I m...	1

Fig- 4: pre-processed dataset

2.2.3 Feature selection

After cleaning and filtering the dataset is used with different feature selection techniques. The aim of feature selection technique is to identify the most valuable features which are highly contributing in recognizing the target object. Therefore, different feature selection techniques have been adopted and then planned for experiment. In sentiment-based text analysis techniques different feature selection techniques can be used such as Term frequency – inverse document frequency (TF-IDF), Part of Speech Tagging (POS), word to vector and others. In the text classification problems, the feature selection approaches are also used for representation of the learning and validation data. In this presented work, the deep learning algorithms are trying to implement for dealing with the large dataset. Therefore, the word embedding technique has been used with this work. The word embedding technique is efficient and easily adoptable with the deep learning systems.

In this process, first the dataset review text has been tokenized by using the word tokenizer method available. A total of 10k tokens has been extracted from the text, where the total number of tokens 58058 has been available. Next the entire review text has been needed to pad for preparing the similar size of text samples for utilizing with the deep learning model. The sequence length is 100 considered in this work. Next for preparing the embedding we need a dictionary of words, therefore Google News Word2Vec model has been used for preparing the embedding. After preparing the word embedding the embedded vector is prepared and used for next step of the data processing. Further, the selected features are subdivided into two parts training samples and validation samples. There are training samples are selected randomly and has a composition of 70% of entire samples. Additionally, the validation samples have a total of 30% of entire samples.

2.2.4 Training ML model

The training samples are used with the machine learning algorithms namely Long Short-Term Memory (LSTM) and 1 dimensional Convolutional Neural Network (1D-CNN). Both the models are variants of Artificial Neural Network (ANN) architecture. Additionally belongs to deep learning models for improved learning performance with large training data. Both the models are briefly discussed as:

Convolutional Neural Network (CNN) model is frequently used. The CNN can also be used in classification task. The CNN is combination of different kinds of filters and layers. In order to understand, the working of the CNN a basic architecture of CNN is given in Fig-5. The architecture includes a number of different components; thus, an overview of the CNN components is given as:

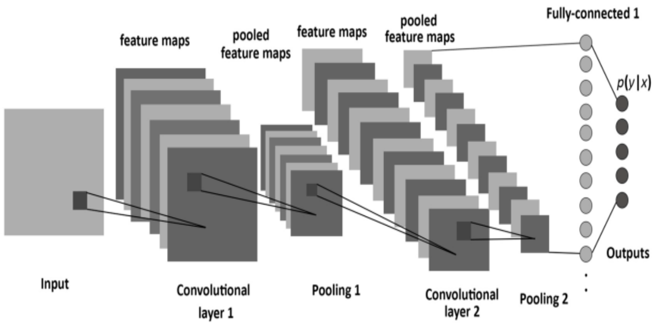

Fig- 5: CNN Architecture

2.2.5 Convolutional layer

The main aim of this layer is to extract features from the input image. The Convolutional layer is always first step in a CNN. In this step an input image is processed using a feature detector and a feature map is produced. The different kind of filters can be utilized in this phase. These filters are working pixel by pixel to the input image. In order to perform this operation, the multiplication of the matrices is used. The CNN learns the values of the filters on its own during the training process. There are a lot of filters options available, therefore there are multiple ways to make the best image classifier. Here, it is a better practice to pad the input image matrix with zeros, before applying the filter into image. This also allows controlling the size. Adding zero padding is wide convolution. Not adding zero padding is narrow convolution.

2.2.6 ReLU layer

The ReLU (rectified linear unit) layer is another step to convolution layer. This applying an activation function onto feature maps to increase non-linearity because images are highly non-linear. Here nonlinearity means the transition between pixels, the borders, the colours, etc. It removes negative values from an activation map by setting them to zero. The rectifier function is removing all the black elements from image, keeping only those carrying a positive value. The essential difference between the non-rectified version of the image and the rectified one is the progression of colours. After rectify the image, we will find the colours changing more abruptly. The gradual change is no longer there. That indicates that the linearity has been disposed of. Therefore, that is not a different layer from Convolutional layer it is a part of it. After this process we advance the input image into max pooling layer.

2.2.7 Pooling

The last thing is in network to get one specific feature. The aim of the pooling layer is to enable CNN to detect the leaf image in any manner. Because for training we need lots of images so that network can recognize leaf in images. Pooling progressively reduces the size of the input representation. It makes it possible to detect leaf in an image no matter where and how they're located. Pooling helps to reduce the number of parameters and computation. It also helps to control over-fitting. There different kinds of pooling techniques available for processing an image namely:

1. Mean pooling
2. Max pooling
3. Sum pooling

In order to obtain good performance in terms of time and accuracy in recognizing the healthy and diseased leaf image we are proposed to investigating all three kinds of technique for pooling. However, in most of the literature max pooling is frequently used technique. According to the figure 6 of max pooling if we have image pixels as given in first block, we need to have a window which scan and select the maximum values in the window. According to diagram we have 2X2 window size to reduce the image features. And the resultant image block is shown in similar image.

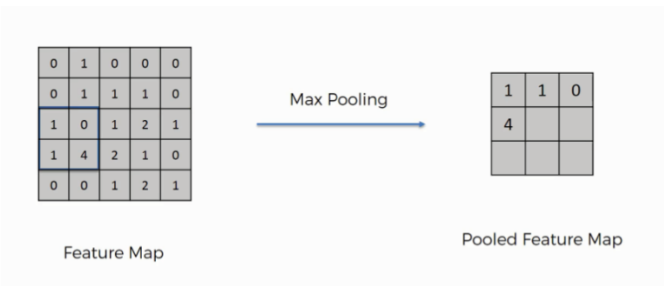

Fig- 6: Example of max pooling

After reducing the image data using the max pooling layer, we need a fully connected layer to learn the features.

2.2.8 Fully connected layer

An artificial neural network to CNN is implemented here the main purpose of this layer is to combine features into attributes. Fully Connected Layer is a feed forward neural network. Fully Connected Layers form the last few layers in the network. The input to the fully connected layer is the output from the final max Pooling Layer. That is flattened and then fed into the fully connected layer. These will predict the classes.

This Flattened vector is then connected to a few fully connected layers which are same as Artificial Neural Networks and perform the same mathematical operations. For each layer of the Artificial Neural Network, the following calculation takes place

$$g(Wx + b)$$

Where, x is the input vector of dimension $[p_l, 1]$, W is the weight matrix of dimensions $[p_l, n_l]$, p_l is the number of neurons in the previous layer and $n_l$is the number of neurons in the current layer, b is the bias vector of dimension $[p_l, 1]$, and g is the activation function. The Fig-7 shows the example of fully connected layer and the above given calculation is repeated for each layer. After passing the image into the fully connected layers, the final layer uses the soft-max activation function. That is used to get probabilities of the input being in a particular class. And so finally, we have the probabilities of the object in the image belonging to the different classes.

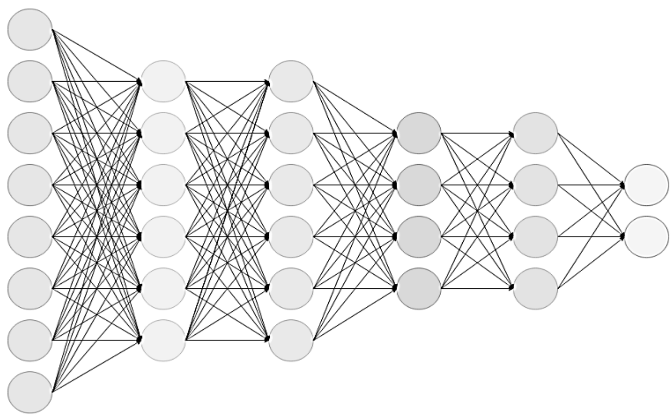

Fig- 7: fully connected layer

2.2.9 Long Short-Term Memory

LSTM is a type of Recurrent Neural Network (RNN) but is better than RNN in terms of memory. It is better for memorizing patterns. LSTM may have multiple layers and information will pass through every layer. The relevant information is kept and the irrelevant information gets discarded. LSTMs efficiently improves performance by memorizing the relevant information and important for finding the patterns. In order to learn LSTM includes 3 types of gates: FORGET Gate, INPUT Gate, and OUTPUT Gate. An Understanding of the LSTM model is demonstrated in Fig-8.

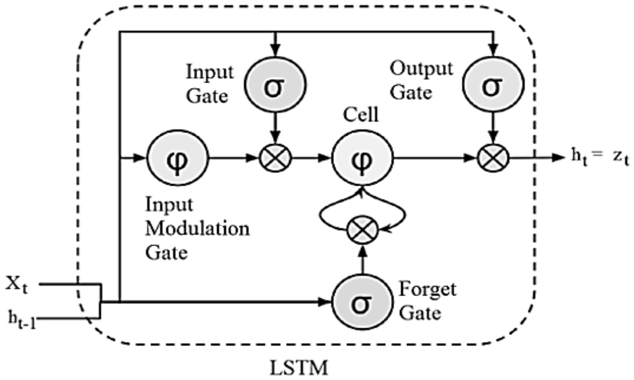

Fig- 8: LSTM Gates

1. FORGET Gate is used to deciding which information is kept and which is not. The h_{t-1} is the information from the previous cell and x_t is the information of current cell. These 2 are inputs of the Forget gate. The inputs are passed through a sigmoid function and input towards 0 are eliminated, and others are passed further layer.

2. INPUT Gate updates the cell state and decides the given information is important or not, and store information in the memory. h_{t-1} and x_t are the inputs for this gate and both inputs are passed through sigmoid and tanh functions respectively. These functions are helpful to control and reduce bias. All the information is used to calculate the new cell state. The cell state is multiplied with the output of the forget gate. Then a point wise addition with the output from the input gate updates the cell state to new values.

3. OUTPUT gate is last gate, which decides "what the next hidden state should be". h_{t-1} and x_t are passed to a sigmoid function. Then the new cell state is passed through the tanh function and is multiplied with the sigmoid output to decide what information should carry.

The 1D CNN model has the following architecture as given in table 1.

Table -1: Architecture of 1D-CNN model

Model type	Sequential
Layer 1	Type dense, number of neurons 128, input dimension (100, 1), activation function 'Relu'
Layer 2	Type dense, number of neurons 64, activation function 'Relu'
Layer 3	Type dense, number of neurons 32, activation function 'Relu'
Layer 4	Type dense, number of neurons 16, activation function 'Relu'
Layer 5	Type dense, number of neurons 2, activation function 'SoftMax'

Additionally for compilation of the 1D-CNN model the categorical cross entropy loss function used with the 'Adam' optimization function. Finally for measuring the performance of training and validation model the 'Accuracy' as the performance matrix has been used. Similarly, the LSTM model has been implemented for classifying the word 2 vector based selected feature. The LSTM has the following layers in the implemented architecture as given in table 2.

Table -2: LSTM architecture

Model type	Sequential
Layer 1	Type embedding, maximum vocab length is 10k, input size is 300
Layer 2	Neurons 100
Layer 3	Type dense, number of neurons 1, activation 'Sigmoid'

Further the LSTM model is compiled with the binary cross entropy loss function. Additionally, the optimizer 'Adam' has been used with the performance matrix 'Accuracy'. The CNN model has been trained for 100 epoch cycles and LSTM classifier is trained for 10 epoch cycles.

2.2.10 Trained Model

Finally, after training of the machine learning algorithms are being trained to accept validation data. The validation data is used with the trained machine learning algorithm for classifying the data. At that time the predicted class labels and actual validation data class labels has been compared to measure performance of the machine learning algorithms.

2.3 Proposed algorithm

The above given methodology of e-commerce review classification can be summarized using the algorithm steps. The steps of implementing the review classification system is given in table 3. According to the given steps, the algorithm is accepting the dataset D and selected machine learning algorithm M. after successful execution of the algorithm the classification labels of validation samples, accuracy and error rate has been obtained.

Table -3: shows the algorithm steps to be implement

Input: Dataset D, Selected ML algorithm M Output: Classified samples C, Performance P
Process: 1. $R_n = ReadDataset(D)$ 2. $P_n = preprocessData(R_n)$ 3. $V = word2Vector(P_n)$ 4. $[Tr, Vl] = Split(V, 70, 30)$ 5. $T_{model} = M.Train(Tr)$ 6. $for(i = 1; i < Vl.length; i++)$ a. $C_i = T_{model}.Classify(Vl_i)$ b. $if(C_i == A_i)$ i. $accuracy++$ c. $else$ i. Error ++ d. End if 7. $end\ for$ 8. Return C, Accuracy, Error

The dataset is read and stored in a variable R_n, where the n is the number of samples in dataset. The prepared variable R_n is pre-processed using the steps as given in previous section. The pre-processed dataset is stored in a new variable P_n. Next for the feature representation the word to vector model has been used. The prepared vector is next split into two parts training samples Tr and validation samples Vl. The Tr has 70% of sample which is used with the selected machine learning algorithm M. after taking training of the model M it becomes T_{model}. Further the validation of the model has been done using the trained model T_{model} and validation samples Vl. For each instance of validation, the trained model T_{model} predict the class label C_i. Further, the predicted class labels C_i are compared with the actual class labels A_i for measuring the performance. If the predicted labels are equal to actual labels, then accuracy is incremented and is not then error is increased. All these three outcomes are provided by the algorithm.

The proposed work is aimed to accurately classify the e-commerce fake reviews to help the consumers to get authentic and actual reviews. By which they can take appropriate decisions related to the product. Therefore, a machine learning model has been presented in this work. Additionally, two different ML models have been proposed to implement for experimental analysis and obtaining most relevant application. Therefore, an architecture of the model has been presented and algorithm steps are provided. Further the presented algorithm is implemented and performance has been studied. Next section includes the implementation and results analysis.

3. RESULTS ANALYSIS

In this section, the proposed fake review classification model has been evaluated on the different performance parameters. Additionally, the individual performance matrix has been discussed which are used in analyzing machine learning algorithms.

3.1 Precision

Precision is also known as positive predictive value. That is the fraction of relevant instances among the retrieved instances. Precision is defined as follows:

$$precision = \frac{TP}{TP + FP}$$

Where, TP indicates the True Positive, and FP shows the False Positive ratio.

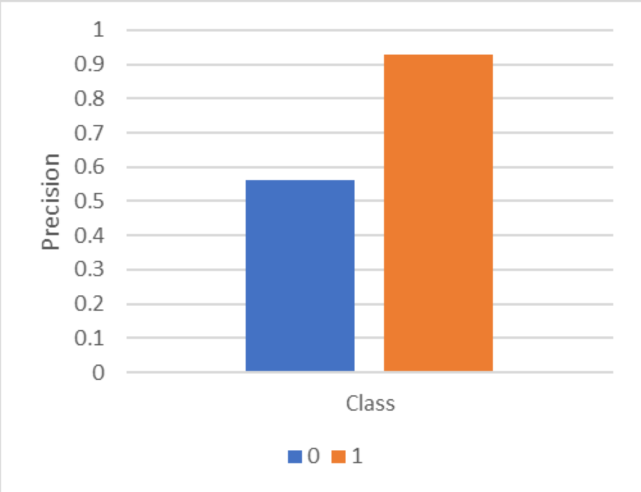

Fig- 9: Class wise precision of CNN model

The class wise precision of the CNN model is given in Fig-9. In this diagram, X axis shows the classes to be classify and Y axis shows precision of the CNN algorithm. According to the obtained results the model is providing higher accurate results for classifying the class label '1' and provides less precision for class '0'. Precision is similar to the accuracy but by only using the precision we cannot decide the best performing algorithm. Therefore, recall is also needed to be measure. The recall of the CNN model is given in next section.

3.2 Recall

Recall is also known as sensitivity or true positive rate and is defined as follows:

$$recall = \frac{TP}{TP + FN}$$

Where, FN shows the False Negative ratio.

Recall is an alternate method to evaluate the model's correctness. Figure 10 shows the recall of the implemented CNN model according to the class to be classify. There two classes are available for recognize and the recall for both the class recognition is given in this figure. In the X axis of the diagram number of classes are given and Y axis shows the recall.

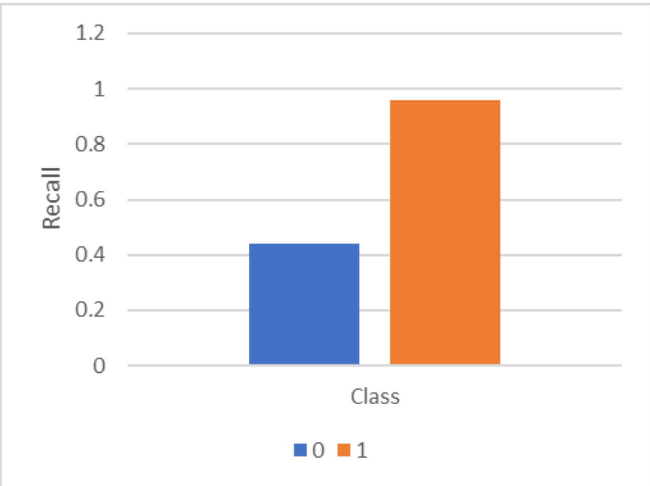

Fig- 10: Recall of CNN model class wise

According to the obtained results the model is provide more accurate classification for classifying the class label 1 and it achieve up to 96% correctness. On the other hand, for recognizing the class 0, the model provides the 46% correctness. Therefore, normally it is essential to make more improvement in the model for obtaining more accurate results.

3.3 F-Score

F1-score is a metric which takes into account both precision and recall therefore that is a harmonic mean of precision and recall in order to describe the quality of classification outcomes. It is defined as:

$$F1 - Score = 2 * \frac{precision * recall}{precision + recall}$$

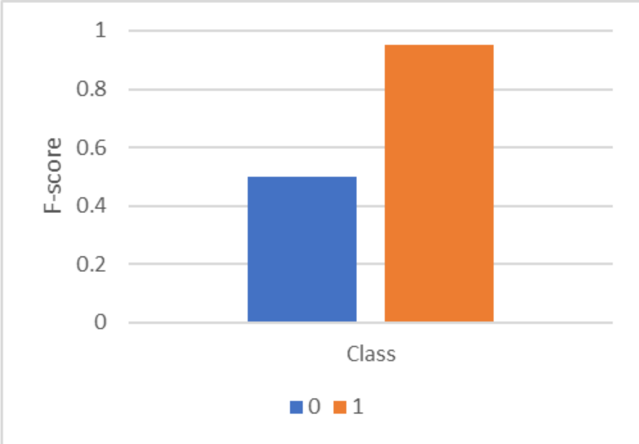

Fig- 11: F-score of CNN model class wise

Fig-11 shows the class wise f-score of the CNN model. In this diagram, X axis shows the classes to recognize and Y axis shows the calculated f-score. Basically, the measure f-score shows the mean of the precision and recall. Therefore, this parameter is used for deciding the better performing classifier. Due to mean of the precision and recall that parameter is also demonstrating the similar performance.

3.4 Accuracy

The accuracy is the ratio of correctly recognized information and total information produced for recognition. The accuracy can be measured using the following equation:

$$accuracy = \frac{correctly\ recognized}{total\ samples}$$

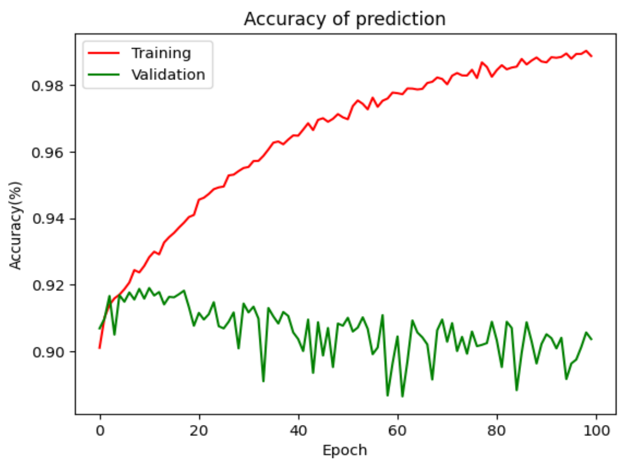

Fig- 12: Training and validation accuracy for 1D-CNN Model

There are two type of classification techniques has been used and both the algorithms are trained and validated using the same dataset. The accuracy for CNN model is given in Fig-12 and the LSTM model's accuracy is given in Fig-13.

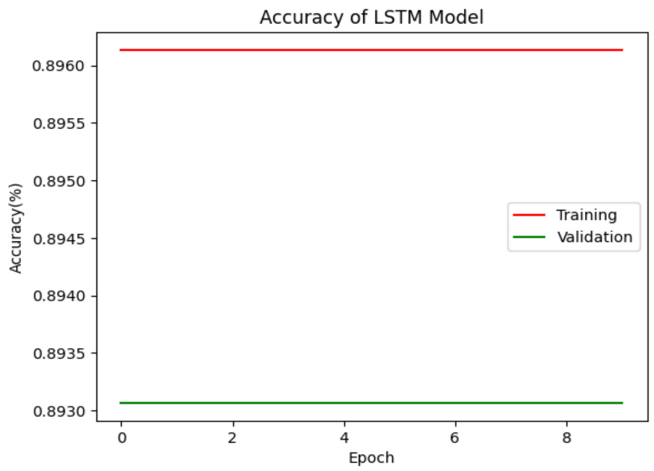

Fig- 13: Training and Validation accuracy of LSTM Model

In both the figures, X axis shows the number of epoch and Y axis shows the accuracy in percentage (%). For CNN model the neural network is trained for 100 epoch and for LSTM model the training is performed for 10 epoch. In this context, when we considering the CNN then training accuracy is improving with the epoch and validation accuracy is degrading with the epoch. However, after some epoch the validation accuracy becomes consistent. On the other hand, when the performance of LSTM model has been considered than it is observed the training and validation accuracy remains consistent for all the epoch cycles. Therefore, in LSTM the training and validation is not improving with the provided training. Further for comparing both the algorithms mean accuracy of the models have also been calculated. The mean accuracy of both the models are given in figure 14 and table 4. In this diagram a bar graph is provided, where X axis shows the algorithms implemented and Y axis shows the accuracy in terms of percentage (%).

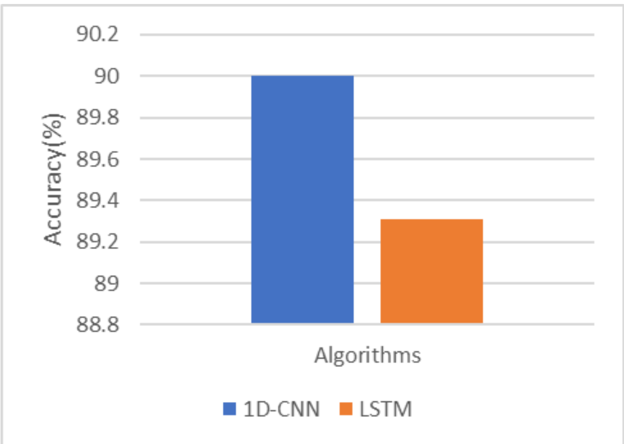

Fig- 14: Comparative Mean Accuracy

Based on the performance in terms of accuracy of both the models both the algorithms are performing similar but the 1D-CNN is providing high accurate results than LSTM.

Table -4: Comparative Mean Accuracy

	1D-CNN	LSTM
Mean Accuracy	90	89.31

3.5 Loss

The loss is not measured in terms of percentage or any other kind of scale. That indicates the difference between predicted and desired outcome obtained from the machine learning algorithm.

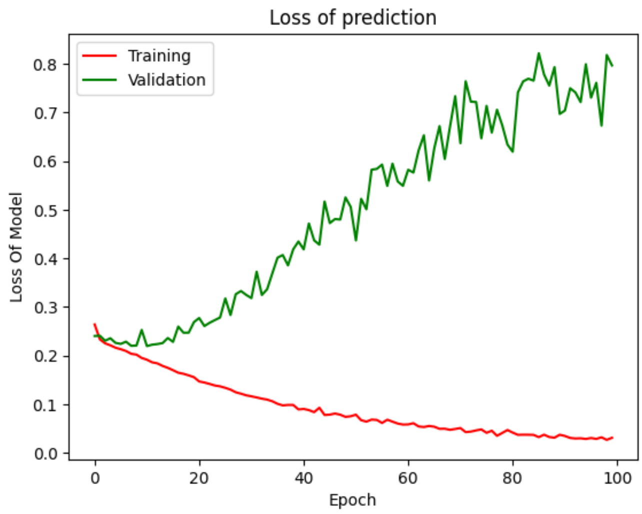

Fig- 15: Loss for CNN training and validation

In this context, the loss functions have been used to measure the difference, using this loss value the optimizer function is performing the update operation on the neural network weights.

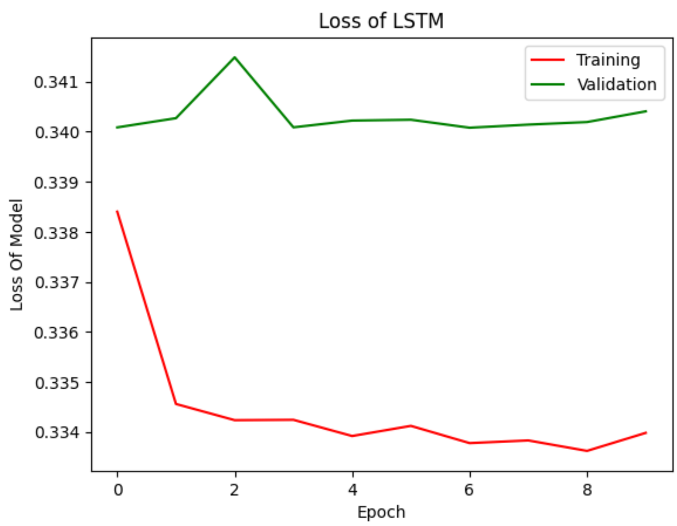

Fig- 16: Loss of LSTM training and validation

The calculated loss of CNN model for training and validation is given in Fig-15. In this diagram, X axis contains the epoch and Y axis contains the measured loss. Similarly, Fig-16 contains the training and validation of LSTM model. According to both the diagrams the training loss is decreasing with the increasing number of epochs on the other hand the validation loss for both the models are increasing with the epoch. However, in both the diagrams the training and validation losses are becomes consistent after some epoch. Normally, it is observed that when the loss become consistent than the model stops the improvement.

4. CONCLUSION AND FUTURE WORK

This section provides the conclusion of the entire work performed. Therefore, the observations and experimental outcomes are given as the conclusion of the work. First provides a brief of the findings and work. Additionally, the future study plan of the presented work has also been discussed.

4.1 Conclusion

The text classification is a classical domain of research and development. However, the text analysis has a number of different applications in real world. The proposed work is intended to explore and investigate the power of text classification based on sentiment analysis. The traditional text classification is different from the sentiment and social media-based text classification task. The current age of text classification contains a large number of instances but the less amount of content available. Therefore, current method of text classification and feature extraction is different. Thus, the presented work is dealing with the small text data analysis problem. Here, the term small is concerned with the social media, review or micro-blog text. The proposed work is mainly applied on review dataset for identifying the fake or misleading reviews from the e-commerce platform.

In this context, as a first task a review has been performed. The review is performed on the basis of small text classification problem based on machine learning algorithms. Using this review the proposed work is trying to identifying the key machine learning algorithms, experimental datasets and relevant results. Next, for supporting fairness in ecommerce platform a spam review classification system has been proposed. The presented spam review classification technique is prepared using the identified techniques in this review. The proposed model is a machine learning based model to conduct text analysis to recognize spam e-commerce reviews. Thus, a complete spam product review classification model is demonstrated. Additionally, the basic overview of the model with the essential components has been also given. The model incorporates the process of text pre-processing, feature selection and two popular deep learning techniques namely LSTM and CNN. Finally, the entire system is described in terms of algorithm steps.

Further, for system implementation the required tools and techniques have been discussed. Additionally, by using the amazon product review dataset is used for conducting the experiments. A total of 30000 instances of the review has been considered, which is belongs from the toy and games category of amazon product. Finally, the performance of the implemented system by using the LSTM and CNN has been evaluated in terms of different performance indicators namely precision, recall, f-score, accuracy and loss. Based on the experiments the LSTM model is able to deliver only 89.31% correct classification while the CNN based model is providing 90% accurate classification results. In addition, it is observed that the LSTM model is utilizing more time for training as compared to CNN. Therefore, CNN algorithm is more efficient and accurate than LSTM model for spam classification in e-commerce platform.

4.2 Future Work

The main aim of the proposed work is to accomplish a machine learning model for classifying the e-commerce reviews into fake and/or legitimate text. The proposed model has been successfully implemented and performance has been evaluated. In near future the following task is proposed for more improvement:

1. The experimental results demonstrate the CNN based classification technique is efficient and accurate thus in near future the CNN based classification technique is recommended for implementation in text classification task.

2. The proposed spam detection approach is a text classification-based technique. That only considers the review text for classification. In near future, the spam review classification system is needed to enhance by incorporating the reviewer's profile attributes also.

5. REFERENCES

[1] Palanivinayagam, C. Z. El-Bayeh, R. Damaševicius, "Twenty Years of Machine-Learning-Based Text Classification: A Systematic Review", Algorithms 2023, 16, 236.

[2] S. Ray, "A Quick Review of Machine Learning Algorithms", 2019 International Conference on Machine Learning, Big Data, Cloud and Parallel Computing (Com-IT-Con), India, 14th -16th Feb 2019, IEEE

[3] J. W. Bi, Y. Liu, Z. P. Fan, E. Cambria, "Modelling customer satisfaction from online reviews using ensemble neural network and effect-based Kano model", International Journal of Production Research, 2019, Vol. 57, No. 22, 7068–7088

[4] H. Jeong, Y. Yi, D. Kim, "An Innovative E-Commerce Platform Incorporating Metaverse to Live E Commerce", International Journal of Innovative Computing, Information and Control ICIC, Volume 18, Number 1, February 2022

[5] T. Chen, P. Samaranayake, X. Y. Cen, M. Qi, Y. C. Lan, "The Impact of Online Reviews on Consumers' Purchasing Decisions: Evidence From an Eye-Tracking Study", Front. Psychol, Vol 13, paper id 865702

[6] A. Onan, "Sentiment analysis on product reviews based on weighted word embeddings and deep neural networks", Concurrency Computat Pract Exper. 2020;e5909

[7] H. Tufail, M. U. Ashraf, K. Alsubhi, H. M. Aljahdali, "The Effect of Fake Reviews on e-Commerce During and After Covid-19 Pandemic: SKL-Based Fake Reviews Detection", VOLUME 10, 2022

[8] N. Hussain, H. T. Mirza, I. Hussain, F. Iqbal, I. Memon, "Spam Review Detection using the Linguistic and Spammer Behavioral Methods," IEEE Access (2020), United States, pp: 53801-53816, Vol: 8, Standard: 2169-3536..