

Sentiment Analysis of Students' Comment using Long-Short Term Model

Irfan Ali Kandhro¹, Shaukat Wasi², Kamlesh Kumar¹, Malook Rind¹ and Muhammad Ameen¹

¹Department of Computer Science, Sindh Madressatul Islam University, Karachi, Pakistan; irfan@smiu.edu.pk, kamlesh@smiu.edu.pk, malook.rind@smiu.edu.pk, ameen.chhajro@smiu.edu.pk

²Department of Computer Science, Mohammad Ali Jinnah University, Karachi, Pakistan; shaukat.wasi@jinnah.edu

Abstract

Objectives: Teacher's evaluation in education system is quite important to improve the learning experience in institutions. For this purpose, sentiment analysis model is developed to identify the student sentiments from the piece of text. **Methods/Statistical Analysis:** Long Short-Term Memory Model (LSTM) is used for analyzing the sentiments expressed by students through textual feedback. For this purpose, dataset has been built through student's feedback and then divided into 70% and 30% for training and testing. The proposed model has been trained using softmax and adam along with drop out values 0.1 and 0.2. Obtained results showed that our model provides 99%, and 90% accuracy over training and validation with 0.2 and 0.5 losses respectively. **Findings:** It was found that proposed model provides an efficient way for sentiment analysis for teacher's evaluation. Model used input as word embedding over the LSTM for mapping the words. And moreover, the model is collected significant semantic and syntactic information by implementing pre-trained word vector model. Hence, this model has the prospective to overcome several flaws in traditional methods e.g., bag-of-words, n-gram, Naïve Bayes and SVM models where order and information about word is vanished. The experimental results show that the model can achieve state-of-the-art accuracy on student feedback dataset. **Application/Improvements:** The study helps for improving the quality of teaching in education system. And moreover, it will be upgrade by increasing the data samples of neutral comments in dataset.

Keywords: Course Evaluation, Opinion Mining, Sentiment Analysis, Student's Feedback, LSTM, RNN

1. Introduction

Sentiment Analysis (SA) often known as opinion mining is used for analyzing or classifying user intentions from the word, sentences or document. These sentiments can further be categorized into different contexts based on certain entity such as Positive, Negative, Neutral, happy, unhappy, Good, Bad, angry, and disgusted in order to know the user's comments towards a subject¹. In a nutshell, SA use public views, in order to analyze different tweets and reviews. Recently, sentiment analysis has been a hot

spot area for scholars in various domains namely: computational linguistics, information retrieval, semantics, natural language processing, artificial intelligence and machine learning. The potential reason behind SA popularity is due to increasing interest of mining gigantic data which heavily come from the social media in the shape of unstructured data. The focus in this research work to address data mining problem in education² like Schools, Colleges, Universities; the evaluation process is a quite complex and depends upon different factors in order to maintain quality standards³⁻⁶. For this purpose, different

*Author for correspondence

approaches are being used to collect data i.e., formal and informal methods in order to analyze opinion from students as to improve the learning and way of teaching.

However, in this paper SA has been carried out for university teacher's evaluation through survey questionnaires³. As the objective of Conducting the survey is to get textual feedback which may contain required information about teachers as well as methodology of teaching. In this manner, we can get knowledge about the course content, punctuality, regularity and the presentation skills. Moreover, it can use for evaluating the particular course in a semester or judging the overall performance of a program. And eventually it would aid course mentors to enhance teaching pedagogy, and overall student learning. Various Machine learning and Natural Processing Language (NLP) techniques have been used for SA^{6,7}. The NLP is an sub area of the computer science, that provides ability to the machines to comprehend the human language. Later it could be used for SA in order to process text for classification such as for extracting features, like words frequency by implementing machine learning algorithms. However, they are classified into following three families⁸.

Supervised Learning: This type of machine learning utilizes labeled input data for classification and regression outcomes. Initially, a learning model is selected and provided with training on sample. Secondly, testing is performed over subsample for correct decision. Most notable types of supervised machine learning algorithms are Linear Regression, Classification and Regression Trees, K-Nearest Neighbor Classifier, Naive Bayes and Support Vector Machines and Neural Networks⁹.

Unsupervised Learning: These algorithms do not require labeled input data for training and testing. So, they look at intrinsic features in order to mine novel patterns. Popular unsupervised learning methods are clustering and association analysis¹⁰.

Semi-Supervised: Semi-supervised learning falls in-between aforementioned learning algorithms. In various applications, data labeling demands high cost and also need human experts. So, if there are few labels available for majority of the observations. Then in this scenario,

semi-supervised algorithms are the best choice for the model building. In¹¹⁻¹³ proposed these methods works well even though the group memberships of the unlabeled data are unknown, this data capture valuable information about the group parameters.

Deep Learning: Deep learning is an advanced concept in the realm of machine learning and AI which aims to make machines more powerful in decision making as similar to human. However, compared to conventional machine learning techniques they are more intelligent to automate the feature learning process from data such as images, video and text data without relying on strict coding rules or human expertise. Deep Learning architecture provides flexible architectures in the sense that their learning capability is higher for raw data³⁻⁵. Therefore, they are good at improving their predictive accuracy when realized over more data. The main aim of this study is to develop a sentiments analysis system using students' reviews for teacher's evaluation. For this purpose, dataset has been built for checking the polarity of words expressed by students as a feedback such as positive, negative, and neutral. And then Recurrent Neural Network (RNN) model has been implemented for text classification which is based on Long Short Term Memory (LSTM) approach for predictive opinion mining. The advantage of using RNN as compared to other machine learning algorithms is that they are more suitable for sequential data such as time series, text, financial data, and multimedia data and beside other related data. In addition, such kinds of model produce in depth understanding of sequence for a certain context. However, RNN suffers from short term memory i.e., that cannot remember long sequence of characters. To overcome this problem, they are added with long term memory by incorporation of LSTM networks. In this manner, RNN provides better performance for analyzing temporal sequence with length in text classification¹⁴. Rest of this study is arranged as follows. Section 2 gives related work in the area of sentiment analysis. Section 3 explains proposed methodology for sentiment analysis over student comments for evaluation of teacher's performance. Section 4 illustrates experimental results and discussion. Section 5 provides comparative analysis among different

tested models. Section 6 provides conclusion and future work.

In recent past, various machine learning and natural language processing methods have been used by researchers for sentimental analysis in the education field⁶. A clustering based idea was presented⁸ in which sentiment analysis is performed using TF-IDF weighting method, voting scheme and importing term scores. Whereas it provides improved results over existing symbolic and supervised learning techniques. In¹⁵ proposed a new approach that uses combination of machine learning and lexicon method for the opinion mining of student feedback in education system. In the process of course evaluation, typically data is collected at the end of semester. And moreover the model is trained by using lexicon feature and weighting method (Term Frequency-IDF) as to analyze the sentiments expressed by student feedback¹⁵. This idea provided a better way to improve the quality of teaching. Similarly, authors¹⁶ developed student feedback mining system by implementing text analytics and opinion mining approach. This method gives instructors with qualitative feedback from students in order to enhance their teaching practices. In other work¹⁷⁻¹⁹ researchers conducted two experiments for sentiment analysis which uses stemmed and non-stemmed technique for generation of tokens. The recorded accuracies were 79.1%, 74.0% for stemmed and non-stemmed and respectively. The evaluated model performed better over stemmed as compared to non-stemmed.

In²⁰ proposed a novel sentiment analysis method which combines Lexicon-based and Learn-based techniques (CLL) for analyzing the cross-domain sentiment reviews for Chinese products reviews. It extracts basic three lexicons for corpus. These are books, hotels and electronics reviews and also they are divided it into four categories. Further 16 features are obtained from them and then trained on six classifiers. The results showed that CLL performed well for books and hotels but achieved lower rate for electronics.

In⁵ proposed suggested RNN-LSTM for efficient neural language for text classification which depends on unsupervised representation of words as input. Their

model is pre-trained on word vectors for the classification of text sentences. And it captures more accurate relationship between semantics and syntactic words. The experimental outcomes showed that simple RNN-LSTM when integrated with word2vec worked well on IMDB dataset. In⁴ employed Naïve Bayes, ID3 and SVM classifiers for effective teaching sentiment analysis. Simulation results revealed that SVM achieved 97% accuracy among other methods in sentiments classification. In^{21,22} described an unsupervised technique for review classification. Reviews are classified on the basis of average semantic orientation of phrases. These phrases are linked with associations such as good or bad depending upon its semantic orientation. Their algorithm attained average accuracy of 74% over opinions reviews dataset namely: Banks, Automobiles, Travel destinations and 66% accuracy on movies review dataset. Here, deep learning techniques was compared with Support Vector Machines (SVM) where it outperformed because deep learning implements hidden layer architecture in order to filter and process the data. Similarly, author²³ proposed novel approach which was based on LSA for identification of related product features. They used SVM model for prediction of features²⁴.

An improved LSTM method was described³ where it was applied over textual emotion attributes for multi-classification. This model achieved higher accuracy in text emotions identification as compared to conventional RNN. Similarly, in²⁵⁻²⁷ performance of the LSTM was improved through using long term dependencies and task of sentence level presentation, word ordering and inter relations. In¹ presented a Sequential Neural Encoder with Latent Structured Description (SNELSD) for novel sentence encoder model. Here, 2-layer hierarchical chain structure has been used for opinion mining and natural language inference tasks. The Algorithm divides sentences into latent word chunks by end-to-end learning. The obtained results are outperformed on discovering task-dependent chunking patterns during the semantic modeling of sentences. Another author¹⁴ employed deep neural network integrated novel approach of sentiment

analysis. It combines kernels of multiple layers in CNN (Convolutional Neural Network) with LSTM.

The model achieved amazing results on video and image analysis. In addition its performance was well on Internet Movie Database (IMDb) review sentiment dataset.

In¹² presented LSTM networks over domain generated algorithms for malware family classification. The suggested method accurately performs multiclass classification. The simulation results reveals that it achieved better performance as compared to state-of-the-art techniques, yielding 0.9993 areas under the ROC curve for binary classification along with a micro-averaged F1 score of 0.9906 respectively. Furthermore, the LSTM based technique gives 90% detection rate with a 1:10000 False Positive (FP) rate which is big improvement over other methods. Their results are tested publicly over existing datasets which shows that LSTM model is superior to other traditional approaches.

2. Proposed Methodology

In this study, LSTM model has been implemented for sentiments analysis and text classification. For this purpose, an algorithms is devised which explains system architecture. As shown in Figure 1, it is divided into five phases such as: 1) Data preprocessing, 2) Word embedding, 3) Long Short-Term Memory model (LSTM) in order to test the hypothesis for prediction accuracy, 4) Dense layer for increasing the model complexity, and 5) Softmax function which is used for multi-class classification problems.

Preprocessing: Initially data is collected from Student's feedback which is a type of unstructured data received in the form of text. Extracting useful information from such unstructured text, it is necessary to apply preprocessing techniques. This can be performed through removing spelling errors, grammatical mistakes and URL's from the text. However, in this research following preprocessing steps is considered. Filtration: In the process of filtration; punctuations, numbers and other special symbols/characters are removed because these characters do not make sense, also create ambiguity and mislead the

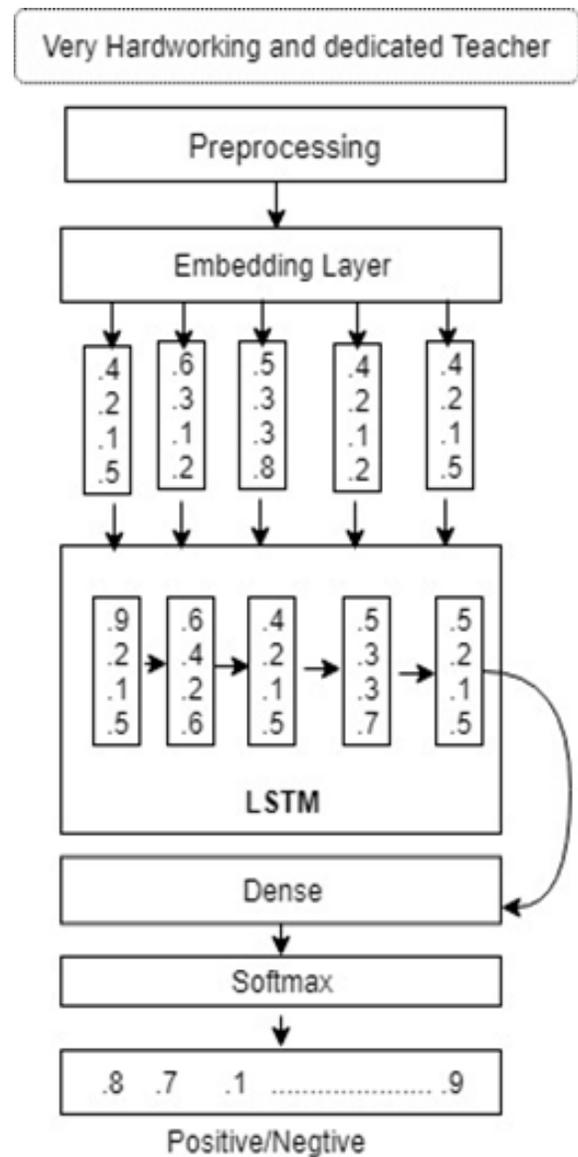


Figure 1. Structure of sentiment analysis model.

context. Tokenization: Tokenization is used for categorization of words from the sentence. It breaks the sentence into words. Case Conversion: After filtering and tokenization process, the tokenized words are transformed into lower case. Stop words: The stopping words are articles used in the sentence. After the filtration, tokenization and case conversion are two most important steps in removing stop-words. These are used to connect the words and make sense to comprehend the sentence. With the help of

stop words we categorize the articles words into meaningful context in order to remove unnecessary information. As results it increases the accuracy of model through assigning the polarity to the words.

Word Embedding: The word embedding is the most significant representation of words in the document. They are mainly for maintaining word relationship, capturing word context in the document for identifying the semantic and syntactic similarity. In our proposed model the word process is a pre-trained word vector which has been inputted to the LSTM network. Firstly these vectors are given to the Word2vec Google's model and secondly they are pre-trained with given model on the Google news updated datasets of 100 billions. The Word2vec model produces 300-dimesional vectors for the 3(million) words, phrases and also support bag-of-words structure. However, Student feedback dataset vocabulary is custom trained vectors. The discovering the CWV, leaning of vectors from the amazon dataset are very efficient, but the output is not very good on the SST model that's the core reason for choosing the Google's word embedding dataset.

LSTM: The representation of sentence in a sequence manner by using the LSTM model. Input of one stage is

word vector that fed to the LSTM layer and the previous hidden state are fed to LSTM for computing the next hidden state. The main pros of implementing the LSTM for sentence vector are that to (Figure 2) out the fixed length sentence vector for any arbitrary variable length sentences. Also, it preserves word order, and it doesn't depend on other linguistic features to capture the semantics¹¹. Prediction is sequentially in RNN. This will assign a memory to the network. Results from previous predictions can improve future predictions. LSTM provides RNN an extra aspect that gives it a fine-grained control over memory. This aspect control how much the current input matters in creating the new memory, and how much the previous memories matters in creating the new memory, and what parts of the memory are important in generating the output, word2vec improved the performance of the model, in the absence of a large supervised training set.

The Equations (1-5), the flow of LSTM model where sigmoid logistic function and „, are the basic gates that control input, forget, memory and output gates. The input gate is regulating that how much new input is added into with model and forget gate shows that the how much old data send by the previous hidden state (both gates are

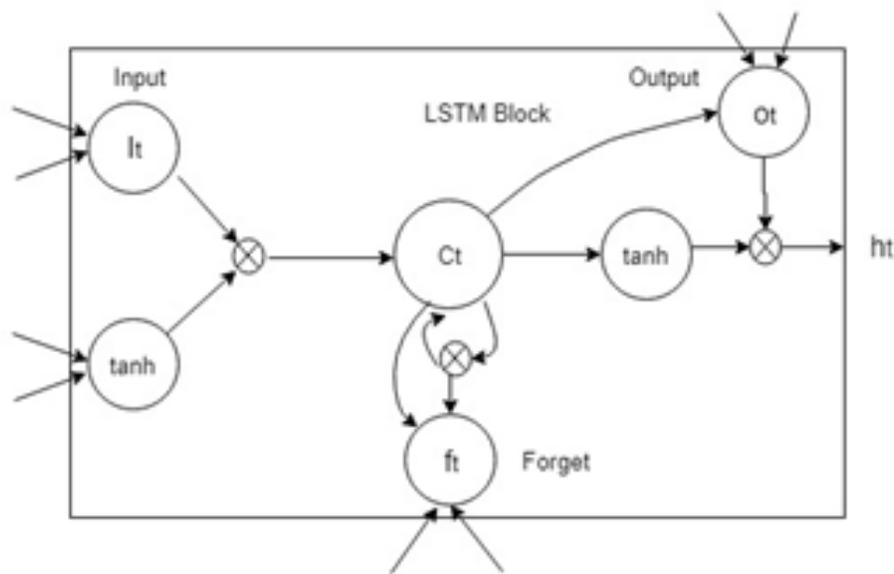


Figure 2. Block diagram of LSTM model.

focus how much work through previous and new state) and output gate determines how much the current node influences the external network. These and provide case-sensitive way to update the information of memory cell²⁴ the equations involved in computing the values of the LSTM are:

Input Activations

$$i_t = \sigma(W^{(xl)}x_t + W^{(hl)}h_{t-1} + W^{(cl)}C_{t-1} + b^{(i)})$$

Input Gate:

$$f_t = \sigma(W^{(xf)}x_t + W^{(hf)}h_{t-1} + W^{(cf)}C_{t-1} + b^{(f)})$$

Forget gate:

$$C_t = f_t \otimes C_{t-1} + i_t \otimes \tanh(W^{(xc)}x_t + W^{(hc)}h_t + b^{(c)})$$

Output gate:

$$o_t = \sigma(W^{(xo)}x_t + W^{(ho)}h_{t-1} + W^{(co)}C_{t-1} + b^{(o)})$$

$$h_t = o_t \otimes \tanh(C_t)$$

Hyperparameters: LSTM layer has 196 nodes which is the output dimension of word vector or word embedding. Varies parameter are used to train the model. The dropout rate 0.2 and activation function softmax is used. The model is trained with adam optimization function with batch size 64, on the dense layer used softmax activation function for multi-classification. We used dropout regularization for avoid from over fitting. The input which is text feedbacks fed to the embedding layer which converts each word to a 300-dimensional vector. The parameters of word embedding layer are maximum features, embedding dimension and length of input and the vector is fed to the LSTM model. The output from the LSTM layer is forward to the dense layer which predicts the output. The loss function used is categorical cross-entropy for multi-class sentiment classification.

3. Result and Discussion

3.1 Description of Dataset

This section describes the proposed approach on the textual feedback of students which has been obtained by various institutions. This analysis is performed after the

Table 1. Random samples of student comments

S.no	Student Comment	Label
1	The Way he taught us we never forgot and Provides us depth knowledge about course	1
2	Our teacher is very helpful, knowledgeable and he motivated me very much	1
3	I'm happy with my teacher	1
4	No depth knowledge about course	0
5	She must come with their proper topic knowledge.	0
6	Waste of time, boring teacher	0

Table 2. Distribution of dataset

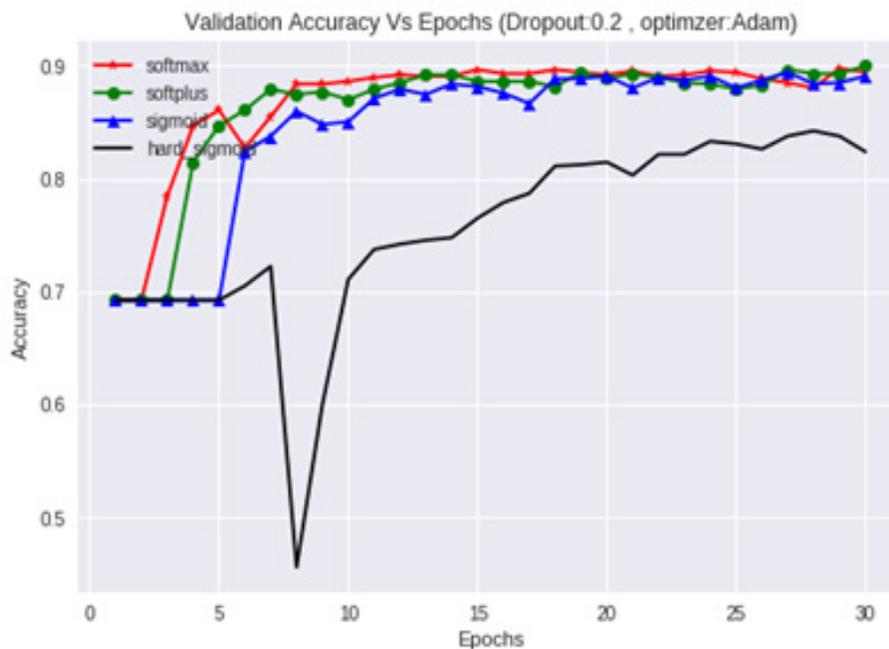
Sentiment Label	Description	Training	Testing
1	Positive	1550	487
0	Negative	750	250
2	Neutral	50	12

partitioned with random sampling criteria. It consists of 30% dataset for testing and remaining 70% for evaluation/testing. The partition of dataset is performed according to the sentiment labels in testing and training. It can be further observed from (Table 2) that the separation of sentiments labels is extremely skewed towards a positive. In addition to aforementioned information; the word cloud visualization method also provides an excellent way for knowing the Metadata associated with sentiments. In this research, we have also visualized the student's feedback dataset in order to better understand students' point of view regarding a course or a teacher. Cloud visualization example of word for positive and negative words has been

shown in Figure 3 and 4, respectively. The most frequent words are good, interesting, excellent, practical, great, helpful, and so forth and they are shown in bold font.

3.2 Comparative Evaluation

This section illustrates experimental results, and these are carried out using multiple and different combination of hyper parameters in order to achieve best results. They are shown in graphically for comparative analysis. It can be observed from Figure 5 that model has been trained using different activations functions such as softmax, softplus, sigmoid, and hard_sigmoid functions with a dropout regularization ratio of 0.2 by optimization parameter

**Figure 5.** Accuracy with dropout 0.2 and adam.

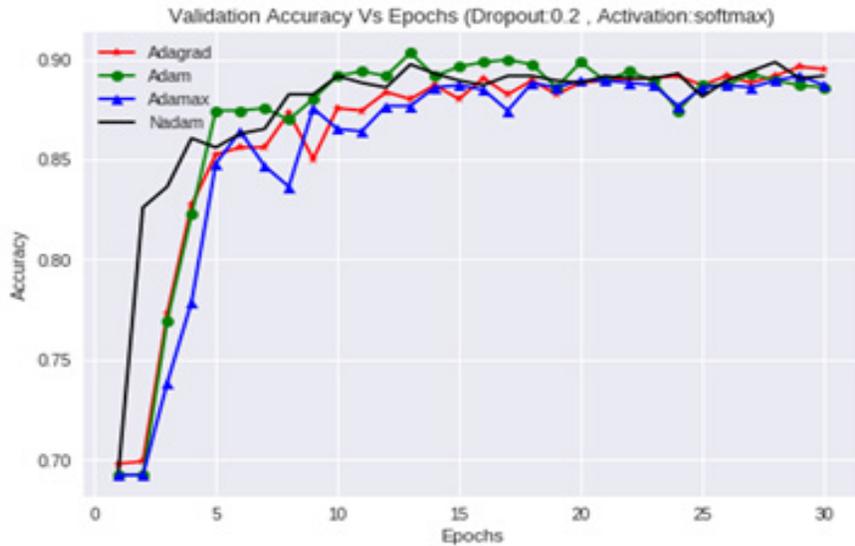


Figure 6. Accuracy with dropout 0.2 and softmax.

adam. In Table 3, results reveal that softmax functions provide better accuracy on certain epochs as compared to other activations functions. Similarly, in Figure 6, another

model is trained using different optimization and activations functions. These optimizers are adagrad, adam, adamax and nadam respectively with dropout of 0.2 ratio

Table 3. Accuracy with dropout 0.2 and adam

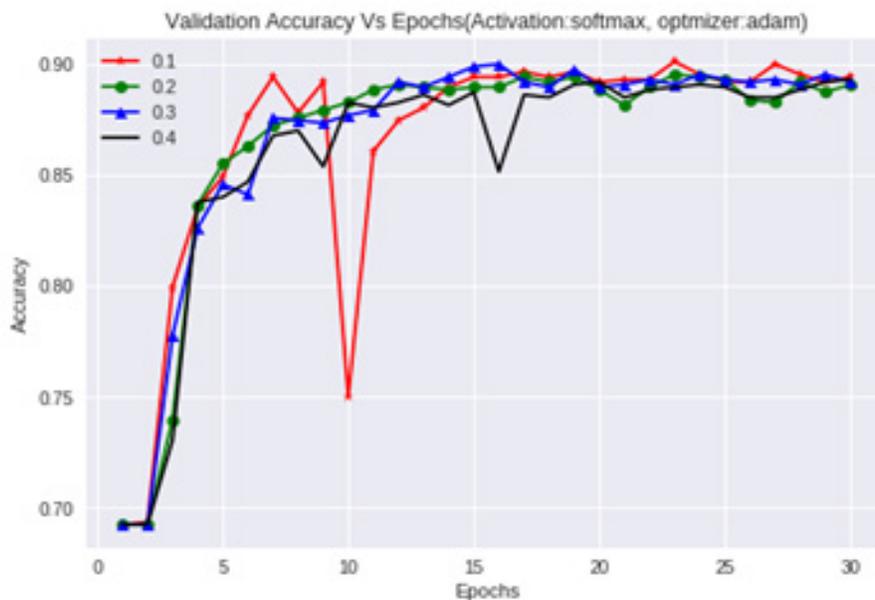
Activation	Optimization	Dropout	Accuracy
Softmax	Adam	0.2	87%
Softplus	Adam	0.2	89%
Sigmoid	Adam	0.2	88%
Relu	Adam	0.2	86%

Table 4. Accuracy with dropout 0.2 and softmax

Activation	Optimization	Dropout	Accuracy
Softmax	Adam	0.2	87%
Softmax	Adagrad	0.2	89%
Softmax	Nadam	0.2	88%
Softmax	Adamax	0.2	88%

Table 5. Accuracy with dropout 0.2 and softmax

Activation	Optimization	Dropout	Accuracy
Softmax	Adam	0.1	89%
Softmax	Adam	0.2	88%
Softmax	Adam	0.3	87%
Softmax	Adam	0.4	86%

**Figure 7.** Accuracy with adam and softmax.

and activation softmax. Table 4 depicts that the adam yields higher accuracy over certain epochs as compared to rest activations functions. Further, Figure 7 shows experimental results which are obtained through different dropout values. These are (0.1, 0.2, 0.3, 0.4) respectively with optimization function adam and activation softmax. However, it can be seen from Table 5 that the model overall achieves better accuracy at dropout ratio 0.1 and 0.2. However, various models are trained using different combination of activations functions and dropout ratios. As shown in Figure 8, an optimal hyper-parameter has been obtained with a dropout rate of 0.2 using activa-

tion function softmax and adam. In this study, we have implemented this model for sentiments analysis because it provides improved results. Moreover, for evaluation purpose accuracy of model has been judged through 99% training and 90 % validation. Similarly, graphical results in Figure 9 represent performance measure using epoch's vs loss. And representation shows that obtained loss for training and validation is about 0.02 and 0.5 respectively.

And these are as minimum losses as shown in Table 6. From these results it can be concluded that performance of our model is quite well in terms of accuracy and f-score for sentiment analysis. Our proposed LSTM model

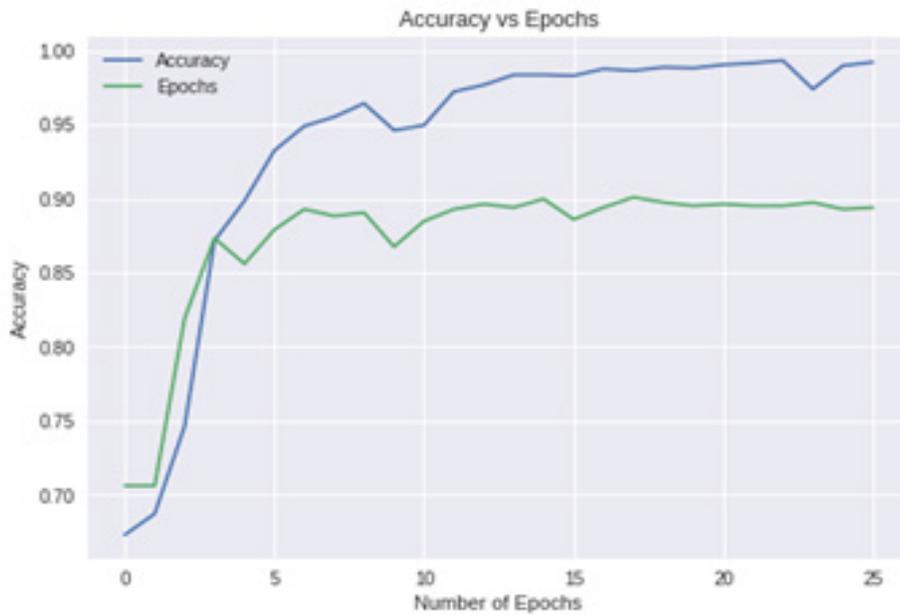


Figure 8. Accuracy with (dropout 0.2, adam and softmax).

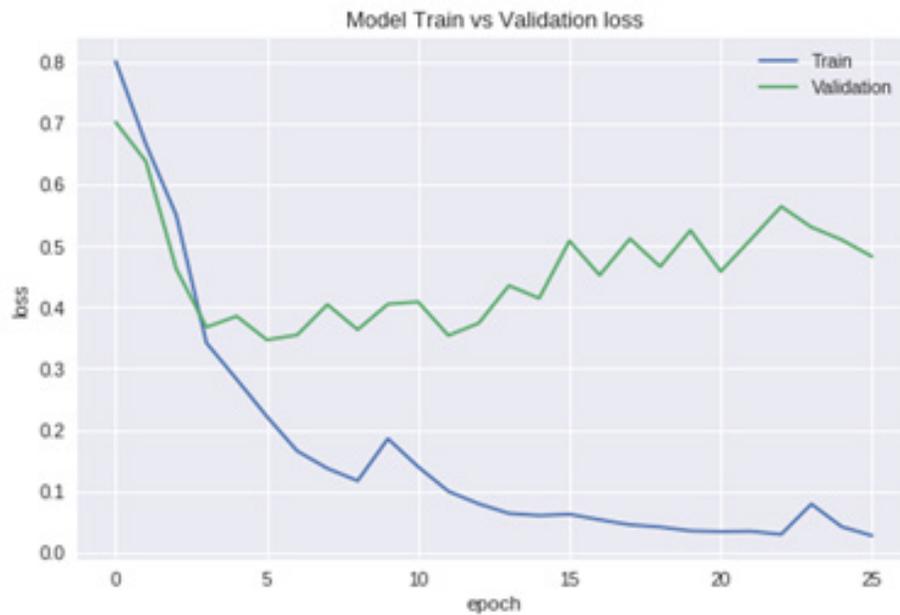


Figure 9. Loss with (dropout 0.2, adam and softmax).

provides improved outcomes as compared to existing models. Because the use of LSTM networks for sentiment analysis over product features requires no prior knowledge, such as syntactic parsing or sentiment lexicon. In

addition LSTM model captures the effect of context on feature emotional polarity, in this manner emotional polarity in product features can be judged automatically. Moreover, due to its long-term memory for the context of

reviews text, which makes up for the disadvantage of the traditional sentiment analysis of neglecting product feature text context, which makes the judgment of emotional tendency on product features more accurate.

$$\text{Accuracy} = \frac{\text{Number of Correctly classified samples Polarity}}{\text{Number of marked samples Polarity}}$$

Recall is depicting the ratio of the number of positive samples which are classified correctly. And that would be though calculated as true positive and false negative sampling. Increased ratio in recall would indicate the correct

class recognition and shows that there is a small number of false-negative sampling in the dataset.

$$\text{Recall} = \frac{\text{True Positive}(TP)}{\text{True Positive}(TP) + \text{False Negative}(FN)}$$

Getting the value for precision, we would divide the number of classified positive samples which are correctly identified from the positive samples which are predicted correctly and positively. Increase in precision would indicate the sample labeled as positive means there is small number of False Positive

$$\text{Precision} = \frac{\text{TruePositive}(TP)}{\text{TruePositive}(TP) + \text{FalsePositive}(FP)}$$

Table 6. Training and validation accuracy vs training and validation loss

Evaluation	Accuracy	Loss	F-score
Training	99%	0.02	0.7547
Validation	90%	0.5	0.6547

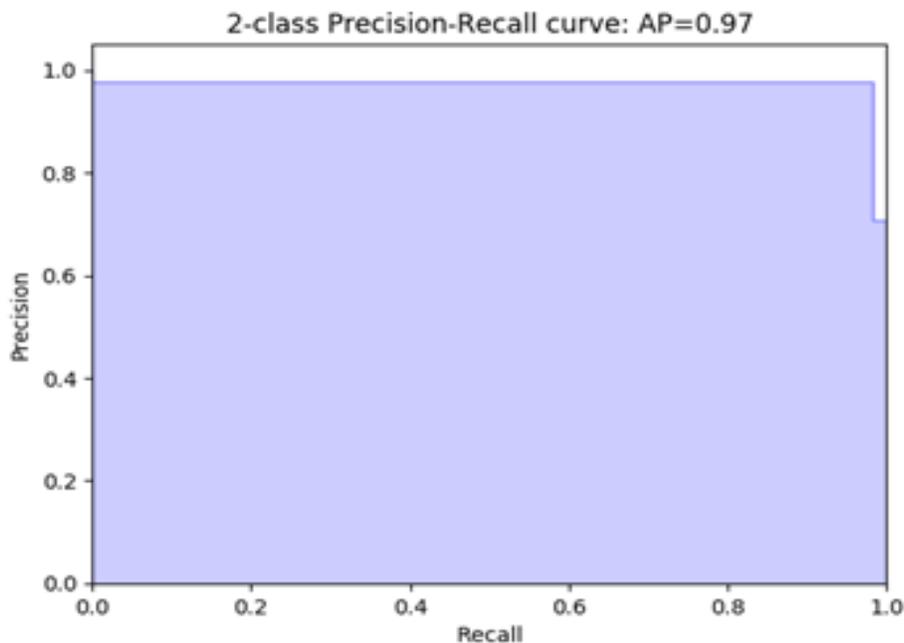


Figure 10. Precision-Recall.

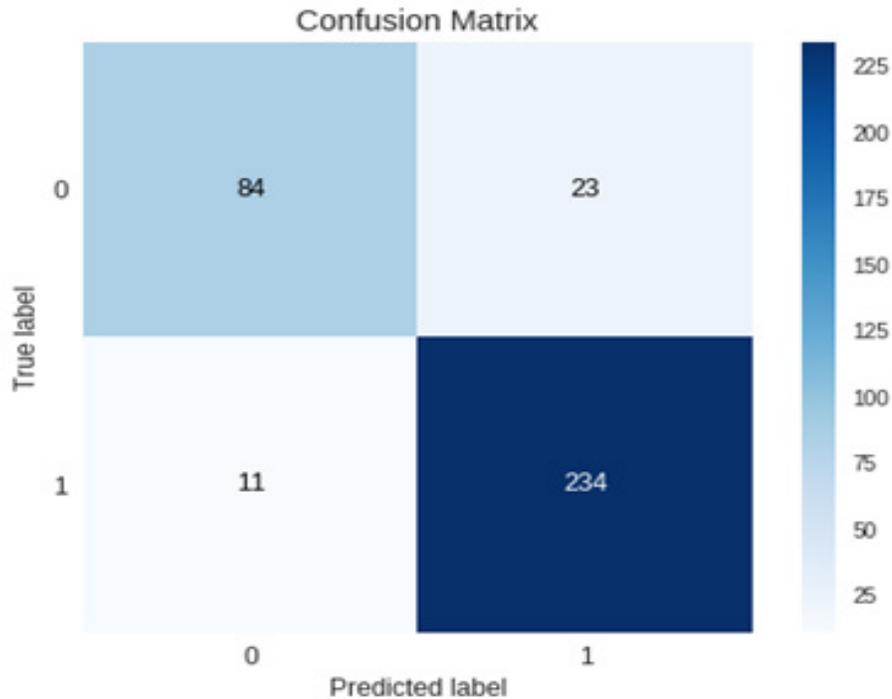


Figure 11. Confusion matrix shows correct and wrong prediction of comments.

Figure 10 shows that the sentiment analysis model provides a better degree of recall and precision, the curve represents the 98 precision and 98.5 recall, it shows that model accuracy is excellent in terms of precision and recall. In addition, the ratio of false negative is greater than false positive. F-measure is a metric which is used commonly in the multi-class classification. It is a geometric mean for precision and the recall. The ratio between the correctly predicted samples and the number of predictions including correct and incorrect predictions which are observed by the system is known as the precision, however; the ratio between the prediction which are correctly predicted and the number of predictions including correct and incorrect predictions observed by the model and the number of sentiments which are truly labeled. This is an effective way or metric for measuring the performance of the model where the data is badly imbalanced²⁰.

$$F1 - score = \frac{2 * Recall * Precision}{Recall + Precision}$$

Confusion matrix is collections of rows and columns that describe the classification performance of model on text data. And model test data which are true basically known values. Confusion matrix terminology is little bit confusing but matrix easy to understand. True Positives: Both cases student's opinion Positive and Negative, Students point of view positive the model predicted positive and similarly negative, model predicted negative. True Negatives: The model predicted negative, student's point of view about the teacher negative. False Positives: The model predicted positive, but the student's point of view about is negative.

False Negatives: System predicted comments negative, but they are positive (Figure 11). The model tested 352 comments, on diagonal side matrix shows true values or correct prediction of model in which 234 comments are True Positive and 84 True Negative and similarly other side 23 False Positive and 11 False Negative (Figure 12). In bar chart demonstrates the testing accuracy and F1 score of the model. Hence the graph represents that the testing

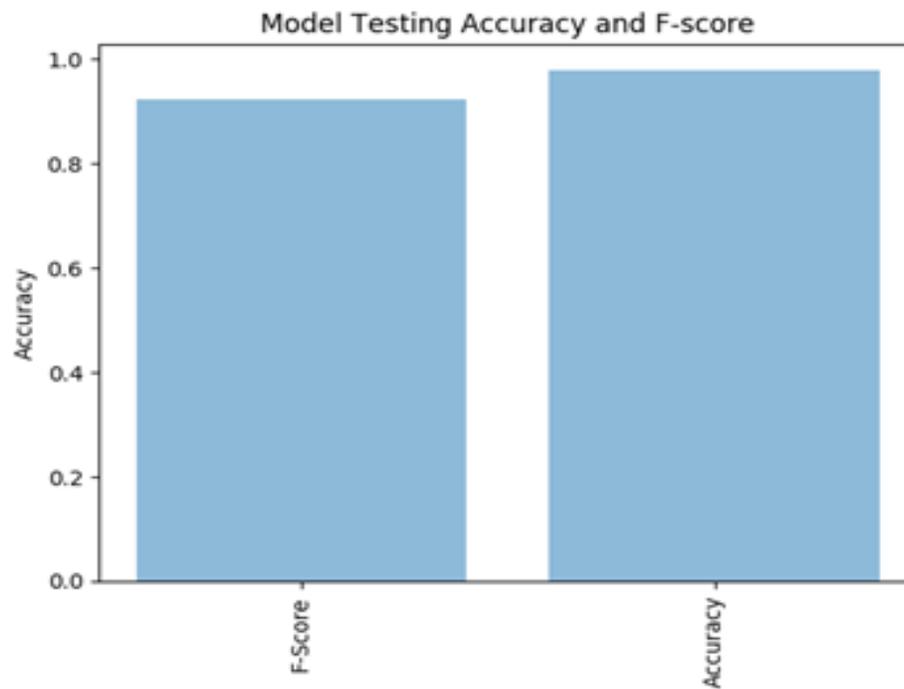


Figure 12. Model testing accuracy and F-score.

Table 7. Positive and negative accuracy

Sentiment Label	Accuracy
Positive	92%
Negative	79%

accuracy is 90% and F1 -Score is 86%. Positive and negative words accuracy are shown in Table 7. The accuracy of our model on positive words are better than negative words in LSTM model on student feedback dataset.

4. Conclusion

In this study, LSTM model has been proposed in order to evaluate the teacher's performance using the student's feedback. For this purpose, textual feedback/comments database was built through Google forms.

Moreover, multiple models are trained for selection of the suitable hyper parameters for our proposed sentiment analysis model.

In addition, these are trained with different combinations of parameters namely: optimization functions; activations; and dropout values. The training and validation accuracy is obtained through various models. Proposed model is used input as word embedding over the LSTM for mapping the words and collected significant semantic and syntactic information by implementing pre-trained word vector model. Hence, this model has the prospective to overcome several flaws in traditional methods e.g., bag-of-words, n-gram, Naive Bayes and SVM models where order and information about word is vanished. The experimental results show

that the model can achieve state-of-the-art accuracy on student feedback dataset.

5. Future Work

In future, this research work shall be extended in order to implement it over multi lingual for sentiments analysis.

6. References

- Ruan YP, Chen Q, Ling ZH. A Sequential neural encoder with latent structured description for modeling sentences. *IEEE/ACM Transactions on Audio, Speech, and Language Processing*. 2018; 26(2):231–42. <https://doi.org/10.1109/TASLP.2017.2773198>.
- Sentiment analysis of student feedback using machine learning and lexicon based approaches. Available from: <https://ieeexplore.ieee.org/document/8002475>
- Esparza GG, de-Luna A, Zezzatti AO, Hernandez A, Ponce J, Álvarez M, Cossio E, de Jesus Nava J. A sentiment analysis model to analyze students reviews of teacher performance using support vector machines. *International Symposium on Distributed Computing and Artificial Intelligence*; 2017. p. 157–64.
- Pong-Inwong C, Rungworawut WS. Teaching senti-lexicon for automated sentiment polarity definition in teaching evaluation. *International Conference on Semantics, Knowledge and Grids*; 2014. p. 84–91. <https://doi.org/10.1109/SKG.2014.25>.
- Sentiment analysis with recurrent neural network and unsupervised neural language model. Available from: https://www.researchgate.net/publication/320086745_SENTIMENT_ANALYSIS_WITH_RECURRENT_NEURAL_NETWORK_AND_UNSUPERVISED_NEURAL_LANGUAGE_MODEL
- Rani S, Kumar P. A sentiment analysis system to improve teaching and learning. *Computer*. 2017; 50(5):36–43. <https://doi.org/10.1109/MC.2017.133>.
- Li D, Qian J. Text sentiment analysis based on long short-term memory. *IEEE International Conference on Computer Communication and the Internet (ICCCI)*; 2016. p. 471–5. <https://doi.org/10.1109/CCI.2016.7778967>.
- Li G, Liu F. A clustering-based approach on sentiment analysis. *IEEE International Conference on Intelligent Systems and Knowledge Engineering*; 2010. p. 331–7. <https://doi.org/10.1109/ISKE.2010.5511190>
- Jagtap B, Dhotre V. SVM and HMM based hybrid approach of sentiment analysis for teacher feedback assessment. *International Journal of Emerging Trends and Technology in Computer Science (IJETTCS)*. 2014; 3(3):229–32.
- Liu CL, Hsaio WH, Lee CH, Lu GC, Jou E. Movie rating and review summarization in mobile environment. *IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews)*. 2012; 42(3):397–407. <https://doi.org/10.1109/TSMCC.2011.2136334>.
- Tang D, Qin B, Liu T. Document modeling with gated recurrent neural network for sentiment classification. *Proceedings of the 2015 Conference on Empirical Methods in Natural Language Processing*; 2015. p. 1422–32. <https://doi.org/10.18653/v1/D15-1167>. PMID:25656747 <https://pubmed.ncbi.nlm.nih.gov/25656747/>
- Predicting domain generation algorithms with long short-term memory networks. Available from: <https://arxiv.org/abs/1611.00791>.
- Sentiment analysis in student experiences of learning. Available from: https://www.researchgate.net/publication/221570557_Sentiment_Analysis_in_Student_Experiences_of_Learning.
- Yenter A, Verma A. Deep CNN-LSTM with combined kernels from multiple branches for IMDb review sentiment analysis. *IEEE 8th Annual Ubiquitous Computing, Electronics and Mobile Communication Conference*; 2017. p. 540–6. <https://doi.org/10.1109/UEMCON.2017.8249013>.
- Baktha K, Tripathy BK. Investigation of recurrent neural networks in the field of sentiment analysis. *International Conference on Communication and Signal Processing (ICCSP)*; 2017. p. 2047–50. <https://doi.org/10.1109/ICCSP.2017.8286763>.
- Nitin GI, Swapna G, Shankararaman V. Analyzing educational comments for topics and sentiments: A text analytics approach. *IEEE Frontiers in Education Conference (FIE)*; 2015. p. 1–9. <https://doi.org/10.1109/FIE.2015.7344296>.
- Im TL, San PW, On CK, Alfred R, Anthony P. Analysing market sentiment in financial news using lexical approach. *IEEE Conference on Open Systems (ICOS)*; 2013. p. 145–9. PMID:23402820.
- Aspect level sentiment classification with deep memory network. Available from: <https://arxiv.org/abs/1605.08900>
- Medhat W, Hassan A, Korashy H. Sentiment analysis algorithms and applications: A survey. *Ain Shams Engineering Journal*. 2014; 5(4):1093–113. <https://doi.org/10.1016/j.asej.2014.04.011>.
- Mao K, Niu J, Wang X, Wang L, Qiu M. Cross-domain sentiment analysis of product reviews by combining lexicon-based and learn-based techniques. Available from: <https://ieeexplore.ieee.org/document/7336187>
- Ghaleb OA, Vijendran AS. Survey and analysis of recent sentiment analysis schemes relating to social media. *Indian Journal of Science and Technology*. 2016; 9(41):1–16.

22. Turney PD. Thumbs up or thumbs down?: semantic orientation applied to unsupervised classification of reviews. *Proceedings of the 40th annual meeting on association for computational linguistics*; 2002. p. 417–24.
23. Ain QT, Ali M, Riaz A, Noureen A, Kamran M, Hayat B, Rehman A. Sentiment analysis using deep learning techniques: a review. *International Journal of Advanced Computer Science and Application*. 2017; 8(6):424–33.
24. Sentiment Analysis on Bangla and Romanized Bangla Text (BRBT) using Deep Recurrent models. Available from: <https://arxiv.org/abs/1610.00369>.
25. Verma S, Saini M, Sharan A. Deep sequential model for review rating prediction. *10th International Conference on Contemporary Computing (IC3)*; 2017. p. 1–6.
26. Leong CK, Lee YH, Mak WK. Mining sentiments in SMS texts for teaching evaluation. *Expert Systems with Applications*. 2012; 39(3):2584–89. <https://doi.org/10.1016/j.eswa.2011.08.113>.
27. Balahur A, Mihalcea R, Montoyo A. Computational approaches to subjectivity and sentiment analysis: Present and envisaged methods and applications. *Computer Speech and Language*. 2014; 28(1):1–6. <https://doi.org/10.1016/j.csl.2013.09.003>.