JOURNAL OF MECHANICS OF CONTINUA AND MATHEMATICAL SCIENCES
www.journalimcms.org

ISSN (Online) : 2454 -7190 Vol.-15, No.-1, January (2020) pp 108-117 ISSN (Print) 0973-8975

A STUDY ON SENTIMENT POLARITY IDENTIFICATION OF INDIAN MULTILINGUAL TWEETS THROUGH DIFFERENT NEURAL NETWORK MODELS

Koyel Chakraborty[1], Sudeshna Sani[2], Rajib Bag[3]

[1,3]Department of Computer Science & Engineering, Supreme Knowledge Foundation Group of Institutions, Mankundu, West Bengal, India

[2]Technique Polytechnic Institute, Hooghly, West Bengal, India

[1]koyel.chak88@gmail.com, [2]sudeshnasani@gmail.com, [3]rajib.bag@gmail.com

Corresponding Author: Koyel Chakraborty

https://doi.org/10.26782/jmcms.2020.01.00008

Abstract

India is a country of having versatile language and culture. Here, people speak in 22 different languages. With the help of Google Indic keyboard people can express their sentiments about any product, news, incidents, laws, games etc. over the social media in their native languages from individual smart phones, tablets or laptops. Sentiment analysis (SA) itself is a tough job, while multilingual SA is even harder as the style of expression varies in different languages. Among the existing approaches of SA till now the machine learning approach through neural network has overcome the limitations of others. The main aim of this paper is to represent a detailed study of the outputs generated from three different models implemented using Convolution Neural Network(CNN), Simple Recurrent Neural Network(RNN) and an amalgamated model of CNN and Long Short Term Memory (LSTM) without worrying about versatility of multilingualism using 2600 sample reviews in Hindi and Bengali. It is anticipated that the experimental results on these realistic reviews will prove to be effective for further research work.

Keywords : Machine learning, Neural Network, Sentiment Analysis, Multilingual Tweets.

I. Introduction

Today's Internet users like to furnish their opinions as comments, feedbacks, questions, requests etc. in different websites or social networking media. Sentiment Analysis is the task of extracting information by NLP and analyzing them as positive or negative reviews in large number of documents. In general, SA plans to decide the demeanor of a speaker or an author with respect to a subject or the general tonality of a record. From the sentiments hidden in those reviews the researchers can infer great information.

Until now the researchers have found a large number of machine learning models that can be applied in NLP differently. Among them Deep Learning approach have been qualified to obtain very high accuracy across many other different NLP models. These models can frequently be prepared with a single end-to-end model and don't require conventional component explicit characterization. Neural system models have accomplished predominant outcomes on different language-related commercial applications when contrasted with customary machine learning models like SVM or logistic regression or Multinomial Naïve Bayes models. [XVII].

Indian native languages are very scarce resourced language. Proper datasets and sentiment lexicons are not developed well and hence are not available for research work. Thus SA of multilingual tweets or blogs is very difficult job in traditional lexical analysis of supervised models of sentiment polarity analysis. India has a population of 130 crores speaking in about 22 different languages. As per 2011 census in urban India with an estimated population of 4.55 crore, it has been found that already 2.95 crore people are using the internet and from the rural areas of India, with an estimated population of 9.18 crore, about 1.86 crore people are using interne. India is the second largest online market, ranked after China. Hence in social media like Facebook, Twitter, YouTube, Amazon, Flipkart etc. the reviews in the local language of Indian people bears a very significant consequence over international promote. Analyzing the polarity of these multilingual reviews in machine learning method has been propped easier than traditional lexical analysis models as it is independent of the grammar of different languages.

As English is the most widely spoken language in the world, the researchers have done SA mostly on English. But work in native or regional languages is limited. In this paper, we aim to present a details study of the features of different Neural Network algorithms and analysis of their outcome in achieving best accuracy in minimum time. The datasets for two popular Indian languages, Bangla and Hindi have been combined after downloading from https://doi.org/10.3390/data3020015[II] and also from Twitter API . CNN has a convolutional layer to obtain information by a bigger bit of content, so we work for sentiment analysis with convolutional neural system, and after designing a simple convolutional neural network model we test it on the Bengali and Hindi datasets of the reviews of Restaurants and Cricket matches. We further have tested the same dataset to a second model created by Simple RNN and successively a third model of LSTM followed by CNN layers. And the result shows that it achieves better accuracy performance in twitter sentiment classification than some of traditional method such as the SVM and Naive Bayes methods. The results and comparison charts are included in this paper.

This paper provides a description of related work on multilingual text analysis and details methodology and comparison of CNN, RNN and LSTM.. A later part of the paper explains background discussion about application of Convolutional Neural Network in NLP and also Recurrent Neural network with help of Long Short term

Memory model. In the next part we furnish flow-diagram and table of results from different models with different number of sample tweet texts and different size of training batches.

II. Related Work

Till now various machine-learning techniques are used for document level sentiment polarity detection [XIV-XX], while more linguistically motivated approaches are used for more fine grained analysis [IX-XIX]. Researchers have also studied sentiment analysis of review documents in specific genres: blogs [XVIII], discussion boards or forums [XIII], user reviews [VI] and expert reviews [X].

The authors have used Multinomial Naïve Bayes algorithm trained using n-gram and SentiWordnet features[III] with an accuracy of 45% using the data sets collected from SAIL MIKE2015[IV].

Chowdhury and Chowdhury did a work sentiment analysis of Bengali tweets [XXV] using Support Vector Machines (SVM) and Maximum Entropy (MaxEnt) on a combination of various features set and later on compared the performance of these two machine learning algorithms. As a result, they secured a high accuracy of 93% using SVM against a feature Unigram+Emoticons (Emotion Symbols). But, their work for only some specific features (like emotion symbols) performs well while in real life these features don't appear in every sentence. Their proposed system is also not suitable for analyzing complex sentences.

Hassan et.al., worked on not just standardized Bangla, but Banglish (Bengali words added with English words) and Romanized Bangla using Deep Recurrent Model more specifically Long Short Term Memory [II] with an accuracy of 78%.

The author in [XXIX], [XVI] implemented SA of comments written in Bengali language using Deep Convolution Neural Network. A sentence is represented as n x k matrix where n is the number of words in the tweet and k is the dimension of the word vector. Their model is supposed to be better than previous model of Deep Belief Network.

Socher et al. [XXIV] has brought in Sentiment Treebank that includes fine-grained sentiment labels for 215,154 phrases in the parse trees of about 12thousand sentences. They used a recursive neural tensor network (RNTN) that predicts fine-grained sentiment for all nodes in the parse trees of sentences. Sentiment label assigned at the root of the parse tree of a sentence is considered as the label of the sentence. They showed that the RNTN performs better than futuristic sentiment classification approaches when it is tested on a test set of positive and negative sentences.

To build the entire system in ASIC we are making the system binary number compatible. For that we are selecting the word length and all co-efficient as 16 bit binary number.

III Background Discussion

Before we proceed, let us discuss about the features of some neural network algorithms that we have used for comparison in this paper. The researchers have considered Deep learning as a part of machine learning for the last few years. Through deep learning various crucial human life application is implemented like - Automatic Speech recognition, Face recognition, Human Language processing [VII]. Two prime neural networks that are used to implement deep learning features are Convolutional Neural Networks and Recurrent neural networks.

CNN which is a feed forward neural network has proved to be the best in case of image processing still now as it contains 4 layers: Convolution layer, (Rectified Linear Unit) layer, Pooling and Fully Connected Layer. Every layer has its own functionality and performs feature extractions and finds out hidden patterns. CNN performs text classifications using word embeddings. The convolution process is applied in 2D matrix where each row represents a word. The author applied filters that slide over rows of the matrix that covers full word [XXVIII]. Later on, a pooling technique is applied on the generated feature maps with an activation function. Pooling will result in a feature vector for individual feature map. Then all the feature vectors will be concatenated into one a big feature vector. At the end, the big feature vector is fed into a fully-connected layer for regularization and classification.

While the RNN or Recurrent Neural Network takes two inputs, the present data and the previous state of output for which the next output depends on. In most of real life situation of human being is like that. The researchers have found that CNN works on the spatial data (images) and RNN works well on sequential data. Recurrent neural networks were usually difficult to train. The Long Short Term Memory, or LSTM Network maybe the best RNN on the grounds that it defeats the issues of preparing a repetitive system and thus has been utilized on a wide scope of uses [XII][XI]. LSTMs help preserve the error that can be back-propagated through time and layers.

IV Proposed Algorithm and Flowchart

We have followed the following common approach for conducting the experiments and result analysis. Fig.1 depicts the flow diagram of the implementation process.

Step-1: Preparation of DATASET

The dataset for our work is collected from of Bengali and Hindi tweets downloaded from Twitter API. The language filtering of the received posts is supported by Twitter API. The optional language parameter in the Twitter Search URL was set to ‘bn’ to extract only Bangla tweets and to search Hindi tweets we set it to ‘hi’. We also collected some datasets from [II] and combined it with twitter data. After collection we split 80% of these dataset into training set and 20% as test set.

Step-2: Preprocessing of the Texts

a) Tokenization - Tokenization is a process of extracting tokens (terms / words) from a corpus. Python’s library Keras has inbuilt model for tokenization which can

be used to obtain the tokens and their index in the corpus.
b) Normalization - Identifying punctuation marks, elongated words of both Bengali and English words.
c) Labeling - We label the positive tweets as 1 and negative tweets as 0 and then combine it into single set.

Step-3: Building up three similar Neural Network Models.
Step-4: Train the Model with training data and label it.
Step-5: Compile the model.
Step-6: Test the model with test data to show accuracy.
We have implemented three models using Neural Network and tested the models with datasets of different number of samples of Bengali and Hindi randomly and furnished the results in the Table 1 and Table 2.
Model-1: Using Traditional Convolutional Neural Network. [XXVIII].
Model-2: Using Recurrent Neural Network [VIII].
Model-3: Using a convolutional layer before LSTM network to reduce the filtering time. [XI]

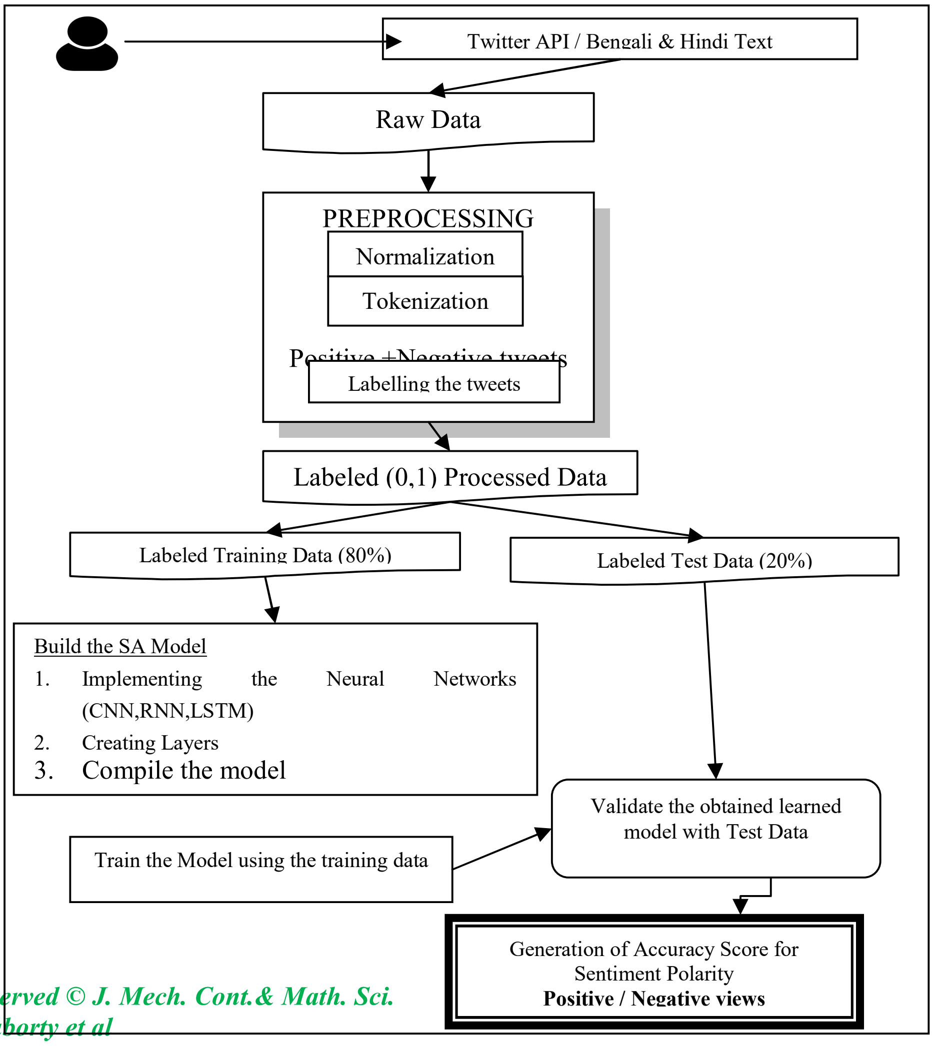

Fig.1 - Flow Diagram of the Sentiment Analysis Model.

V. Simulation Result and Analysis

We have found the following result after analyzing and testing the models with 256 batch size of training data. Table 1 shows the accuracy score and the execution time (in seconds) from creating the model up to validation of test data.

Table 1 - The variation of accuracy in different models

Dataset	**Accuracy Obtained w.r.t Execution Time**					
	Model-1(CNN)	Execution Time (seconds)	Model-2 (RNN)	Execution Time (seconds)	Model-3 (CNN+LSTM)	Execution Time (seconds)
Hindi -1 (750 samples)	76.55%	20.91	73.79%	16.46	76.55%	**13.87**
Bengali-1 (500 samples)	77.11%	1.655	46.99%	2.775	77.11%	3.670
Bengali-1 (1800 samples)	75.35%	6.079	75.62%	7.34	75.35%	6.92
Bengali-2 (400 samples)	68.57%	1.218	68.57%	2.099	68.57%	3.25
Bengali-2 (2500 samples)	77.80%	4.54	67.16%	9.40	70.09%	9.131

We also observed the good training accuracy in Table 2, which is obtained from the Model -3 irrespective of the amount of sample data sets and Training Batch sizes.

Table 2 – Exhibiting the best training accuracy in Model-3

Dataset No	**Model-3**								
	Accuracy	Exec time	Batch size	Testing on (samples)	Validate on (samples)	Training Accuracy	Testing Accuracy	Training Loss	Validation Loss
Hindi-1 (Total 727)	76.55%	14.90	128	580	146	**0.7724**	0.7655	0.3823	0.5882
	76.55%	13.87	256	580	146	0.7724	0.7655	0.5013	0.5486

Bengali-1(1800 samples)	68.42%	7.328	128	1446	362	**0.9703**	0.6842	0.0784	1.5357
	75.35%	6.92	256	1446	362	0.7386	0.7535	0.3872	0.5669
Bengali-2(2600 samples)	68.99%	9.69	128	2181	546	**0.9853**	0.6899	0.0445	1.5861
	70.09%	9.131	256	2181	546	0.9757	0.7009	0.0804	1.1731
	77.80%	9.181	512	2181	546	0.7955	0.7780	0.3690	0.5359

The following graph in Fig. 2 & Fig. 3 shows the progress level of testing and training accuracy and losses. The horizontal-axis represents number of epochs and the vertical-axis represents the amount of accuracy score per epoch.

The details are as :

Accuracy: 76.55%
Training Accuracy: 0.7724
Training loss: 0.5013
Testing Accuracy: 0.7655
Validation loss: 0.5486
The execution time : 13.871559381484985

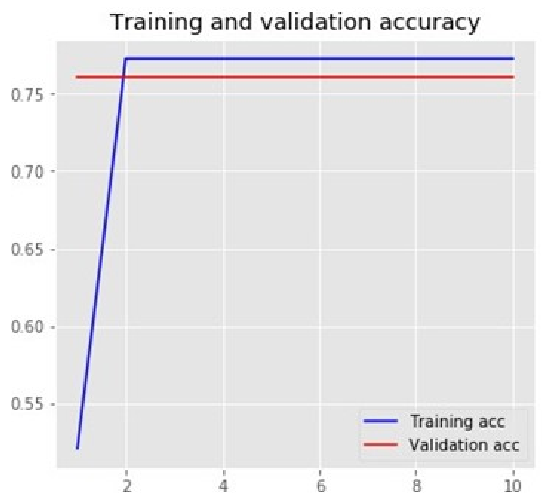

Fig.2 – Graph representing Accuracy

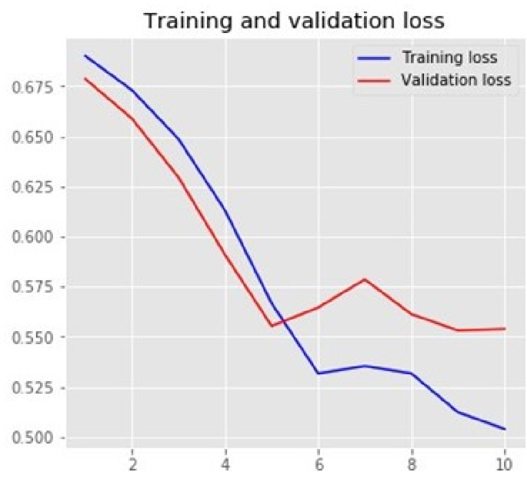

Fig.3 – Graph representing Loss

V. Conclusion

As depicted in the result tables, the enhancement in the loss of validation is observed for Model-2(RNN) and Model-3(CNN+LSTM). Amongst all the models, the best accuracy is being achieved through Model-1(CNN). Though the best accuracy is received for Model-3(CNN+LSTM), but it happens in case when the amount of data is less. In today's world, where the amount of data is a big issue and is constantly increasing, it is anticipated that better models be designed in the future that gives better accuracy and yet requires minimal execution time. Handling emotions of humans is one of the toughest challenges in the world today, and hybrid models should be created that can handle the perception capability of humans efficiently. The Long Short-Term Memory, or LSTM, network is found to be a better RNN because it defeats the problems of training a recurrent network and so may be used on a wide range of applications. Research may also be trodden into latest naïve fields like emotional intelligence, stance analysis, sentiment identification from individual progress [XXI] etc and special attention should be paid to analysis on real-time data from various social media sites.

References

I. A. B. Goldberg, and X. Zhu, "Seeing stars when there aren't many stars: graph-based semi- supervised learning for sentiment categorization" In Proceedings of the First Workshop on Graph Based Methods for Natural Language Processing, Association for Computational Linguistics, pp. 45-52,2006.

A. Hassan, N. Mohammed, and A. K. A. Azad, "Sentiment analysis on bangle and romanized bangla text (BRBT) using deep recurrent models," Computing Research Repository (CoRR), vol. bs/1610.00369, 2016. [Online]. Available: http://arxiv.org/abs/1610.00369.

A. Trivedi, A. Srivastava, I. Singh, K. Singh, and S. K. Gupta, "Literature Survey on Design and Implementation of Processing Model for Polarity Identification on Textual Data of English." IJCSI,2011.

B. G. Patra, D. Das,, A. Das, and R. Prasath, "Shared task on sentiment analysis in indian languages (sail) tweets-an overview", In International Conference on Mining Intelligence and Knowledge xploration, pp. 650-655,Springer International Publishing, December, 2015.

B. Pang, and L. Lee, "Opinion Mining and sentiment analysis,", Foundations and trends in information retrieval, 2(1-2), 1-135, 2008

B. Pang, L. Lee, "A sentimental education: Sentiment analysis using subjectivity summarization based on minimum cuts" In Proceedings of the 42nd annual meeting on Association for Computational Linguistics (p. 271),2004.

B. Pang, L. Lee, and S. Vaithyanathan, “Thumbs up?: sentiment classification using machine learning techniques”. In Proceedings of the ACL-02 conference on Empirical methods in natural language processing, Association for Computational Linguistics, Vol. 10, pp. 79- 86,2002

Bullinaria, J.A., “Recurrent neural networks”. Neural Computation: Lecture, 2013.12.

E. Riloff, and J. Wiebe “Learning extraction patterns for subjective expressions” In Proceedings of the 2003 conference on Empirical methods in natural language processing, Association for Computational Linguistics, pp. 105-112,2003.

F. Zhu, and X. Zhang, “Impact of online consumer reviews on sales: The moderating role of product and consumer characteristics” Journal of Marketing, 74(2), 133-148,2010.

Gers, F.A., J. Schmidhuber, and F. Cummins, “Learning to forget: Continual prediction with LSTM”. Neural computation, 2000.12(10): p.2451-2471.

Hochreiter, S. and J. Schmidhuber, “Long short-term memory”. Neural computation,1997. 9(8): p. 735-1780.22.

J. Kim, G. Chern, D. Feng, E. Shaw, and E Hovy, “Mining and assessing discussions on the web through speech act analysis” In Proceedings of the Workshop on Web Content Mining with Human Language Technologies at the 5th International Semantic Web Conference,2006.

J. Zhao, K. Liu, and G. Wang “Adding redundant features for CRFs based sentence sentiment classification” In Proceedings of the conference on empirical methods in natural language processing Association for Computational Linguistics, pp. 117-126,2008.

K. Dave, S. Lawrence, and D. M. Pennock, “Mining the peanut gallery: Opinion extraction and semantic classification of product reviews” In Proceedings of the 12th international conference on World Wide Web, ACM, pp. 519-528,2003.

Kamal Sarkar, “Sentiment Polarity Detection in Bengali Tweets Using Deep Convolutional Neural Networks”, https://doi.org/10.1515/jisys-2017-0418.

K Sarkar, M Bhowmick, “Sentiment polarity detection in bengali tweets using multinomial Naïve Bayes and support vector machines”, - 2017 IEEE Calcutta Conference (CALCON), 2017.

L-W Ku, Y.T. Liang, and H-H Chen, “Opinion Extraction,Summarization and Tracking in News and Blog Corpora” In AAAI Spring Symposium: Computational Approaches to AnalyzingWeblogs,2006.

M. Wiegand, A. Balahur, B. Roth, D. Klakow, and A. Montoyo, “A survey on the role of negation in sentiment analysis” In Proceedings of the workshop on negation and speculation in natural language processing, Association for Computational Linguistics, pp. 60-68, 2010.

Q. Miao, Q. Li, and D. Zeng, "Fine grained opinion mining by integrating multiple review sources" Journal of the American society for information science and technology, 61(11), 2288-2299,2010.

Rajaram, Santhoshkumar & Geetha, M.. (2019). "Deep Learning Approach for Emotion Recognition from Human Body Movements with Feed forward Deep Convolution Neural Networks". Procedia Computer Science. 152. 158-165. 10.1016/j.procs.2019.05.038.

R. Narayanan, B. Liu, and A. Choudhary, "Sentiment analysis of conditional sentences" In Proceedings of the 2009 Conference on Empirical Methods in Natural Language Processing, Association for Computational Linguistics, Vol. 1, pp. 180-189,2009.

R. Prabowo, & M. Thelwall, "Sentiment analysis: A combined approach" Journal of Informetrics, 3(2), 143-157, 2009.

R. Socher, A. Perelygin, J. Y. Wu, J. Chuang, C. D. Manning, A. Y. Ng and C. Potts, "Recursive deep models for semantic compositionality over a sentiment Treebank", in: Proceedings of the Conference on Empirical Methods in Natural LanguageProcessing (EMNLP), vol. 1631, p. 1642, ACL, Seattle, USA,2013.

S. Chowdhury and W. Chowdhury, "Performing sentiment analysis in bangla microblog posts," 2014 International Conference on Informatics, Electronics & Vision (ICIEV), vol. 00, pp. 1–6,2014.

T. Joachims, "Making large scale SVM learning practical",1999.

T. Mullen, and N. Collier, "Sentiment Analysis using Support Vector Machines with Diverse Information Sources" In EMNLP, Vol. 4, pp.412-418, 2004.

Y. Kim, "Convolutional neural networks for sentence classification," in the Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing, pp. 1746– 1751,2014.

Y. Zhang and B. C. Wallace, "A sensitivity analysis of (and practitioners' guide to) convolutional neural networks for sentence classification," Computing Research Repository (CoRR), vol. abs/1510.03820, 2015. [Online]. Available: http://arxiv.org/abs/1510.03820.