Sentiment Analysis

on

Amazon Customer Reviews

Harshul Gupta1

, Utkarsh Gupta2

Ritik3

1 Department, CEIT, ABESEC, India

Email: harshul.18bci1055@abes.ac.in

2 Department, CEIT, ABESEC, India

Email: utkarsh.18bci1056@abes.ac.in

3 Department, CEIT, ABESEC, India

Email: ritik.18bci1045@abes.ac.in

**Abstract**

Sentiment Analysis is one the most important analysis that needs to be done while determining the popularity of any website or any product. The reviews of the customers determine whether that product or website will see further growth and demand in the market or not. It is also important for the company to keep track of the reviews for their products continuously so that they can enhance and improve the quality of their product as per the needs of the customers.

But analysing all these reviews and determining whether they talk in favour or complains about the product manually can be a really tedious task especially when the most important thing is to know the negative reviews to be able to better the quality of the product but, without analysing the reviews one cannot determine whether a review is positive or negative. This is where Natural Language Processing (NLP) comes into the picture. NLP can separate the negative reviews from the positive ones and thus one further doesn’t need to read the positive reviews and can only focus on the negative reviews. This can also help determine the popularity of the product by taking the ratio of negative to positive reviews.

Using an NLP model created to test the positive and negative sentiment from within a written statement can help us to give a tag of positive or negative to a review and thus can help us to separate the two types of reviews. For this NLP uses tokenization, word embeddings and is trained using Bidirectional LSTM neural network to get as much accuracy as possible. LSTM uses a cell space which runs through the whole network and thus helps to give more appropriate weights to any words in a sequence of words.

**1. INTRODUCTION**

Understanding the negative or positive sentiment from within a review can help to separate the

positive reviews from the negative reviews and thus can help us to determine the popularity of any

product.[1] Determining the popularity can not only help in determining which product is most popular among the consumers but, can also help to identify which product needs more improvement and which product is good as it is and further innovation can still wait.[2]

Determining the correct sentiment from a review by simply comparing some particular words such as “not working well”, “didn’t liked the product”, etc., can be successful up to some extent but reviews are not that simple. People don’t really type their reviews in such simple manner like “good” for actual positive review and “bad” for actual negative review. Many a times people tend to write some reviews such as “As if it really works as well as the company states”, trying to give a sentiment to this review by simply using a comparison with “works as well” will give it a positive tag but, in-fact this review is a negative review[2]. To be able to understand the correct sentiment for such a review the only correct way is to use an NLP (Natural Language Processing) model which is trained to correctly identify the sentiment for even such reviews as mentioned above.[3]

**2. LITERATURE SURVEY**

A review for any product or website can mainly consist of two types of sentiments, first is a positive sentiment, which favours the quality of the product and may convey that the product is as good and useful as the seller claims or might be even better.[3] The second sentiment that a review may contain is negative, which may signify that the quality or usability of the product is not as good as the seller claims or is even worse.

Positive review can build can heighten the popularity of the product and increase its demand in the market, whereas negative reviews can result in total drop in the demand of that product and can also affect the reputation of the seller or the manufacturer. Thus, improving and innovating the product as per the negative reviews can help the product to maintain its demand in the market.

Recognising the sentiment from a sequence of words or a sentence requires to first read the whole sentence and then figure out the positive or negative sentiment expressed in that sentence[4].

This can be done by training a model on a dataset which contains of mainly two attributes, review and the sentiment expressed in that review.

Training a robust model for the purpose can be achieved by using Deep Neural Network Techniques such as Embeddings and Bidirectional LSTMs.[5-20]

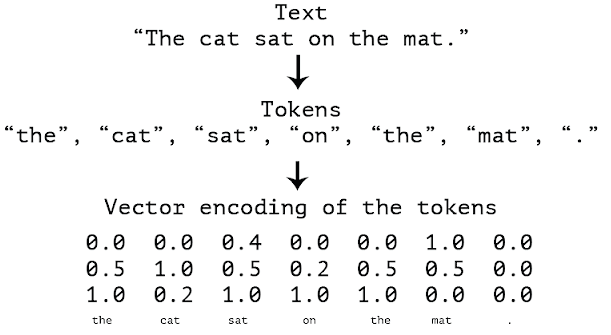
**3. METHODOLOGY**

To make a model that can identify the negative or positive sentiment from within a review we firstly trained our NLP model on a dataset containing reviews and the sentiments associated to them.

**3.1 Preparing the Model**

**3.1.1 Tokenization**

Tokenization is the process of breaking down the sentences into tokens (such as words) and then each token is assigned with a specific value which can be represented in the form of a vector which can be easily feed into a neural network.[6]

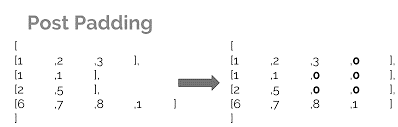


This creates a set of words (also called as vocabulary) along with their vector encoding. We can choose the max size for this vocabulary and for those word which are not specified in our vocabulary we can assign a special character for the unknown words such as “<OOV>” or any other character as per one’s choice.

For this model we had taken the vocab size as 10000 and the for unknown word the token is “<OOV>”.[7]

**3.1.2 Padding**

All the reviews may or may not be of the same size so to make all the reviews of similar size we use padding. Padding can help fill the missing values within a vector for the sentences which makes the calculation accurate. We added padding at the end of the sentences and kept the max size of padding as 120.[7]



**3.1.3 Neural Network Layers Used**

In the proposed work, the tokenized data created is then feed to a model containing various neural network layers.

Three layers has been used for this model:

**3.1.3.1 Embedding Layer:**

Embedding layer trains its weights in such a way so that they can use to represent all the words form the tokenized data in a 2-D space in such a manner so that each word will occupy a close position to those words which have a close relation to them.[7-8]

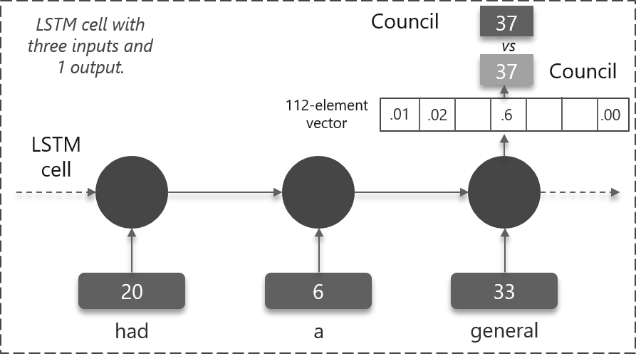
For example, words like better, best or bad, worse, worst, which imply same meaning or similar possibility of occurrence in a sentence are close to each other in the 2-D representation of the tokens.

The embedding layer has been given three arguments, first is the size of the vocabulary, second the dimension of the embeddings required and third is the max length of each input or the sentence whose sentiment has to be figured out.

**3.1.3.2 Bidirectional LSTM:**

Long Short-Term Memory (LSTM) is a recurrent neural network which is capable of processing the whole sequence of words in one go as it contains the ability to carry the information of the weights for the previous words even for very large sequences unlike a standard RNN layer which losses the information over long sequences.

LSTM contains a cell state that stores this information and carries it along the whole network and tunes the weights while considering the information contained within that cell-state.[11]



In case of Bidirectional-LSTM the training of the LSTM layer is not done only for the sequence in the given order but also for the reverse order also i.e., for a sentence “This is bad” the training will be done for two sentences, first for “this is bad” and second for the sentence “bad is this”. [12]

The Bidirectional-LSTM layer is also feed with one argument that is the max size of each input.

**3.1.3.3 Dense Layer:**

Dense layer is a fully connected layer of neurons in which each input is connected to an output. Two parameters are given to thus layer, first the max length of inputs and second is the activation function for the classification. This layer has been used as the final output layer for this model and the activation function used in this layer is SoftMax.[13-14]

**3.1.3.3.1. SoftMax:**

This is an activation function which is used for the problems where we need the neural network to classify more than one class. Since, this sentiment analysis is only using two classes namely, positive and negative, thus, the softmax function in here will work as a sigmoid function only. [18-19]



sj 🡪 scores for each class given by the network

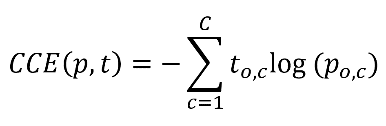
c 🡪 classes

**3.1.4. Model Compilation:**

Model is then compiled using the Sparse- categorical-cross entropy loss as the loss function and the Adam optimizer as the optimization function with a learning rate of 0.001.[6]

**3.1.4.1. Sparse-Categorical-Cross entropy:**

It is used when the target values are of integer type



CCE() 🡪 sparse categorical cross entropy loss function

p 🡪 probability for each class

t 🡪 target vector

c🡪 class

**3.1.4.2. Adam Optimizer:**

The optimizers are used to update the values of the weights at each neuron to get as close to the most optimised weights as possible. Adam optimizer uses RMS-Prop and the Adaptive Gradient Algorithm.[20-21]

**3.2 Live Data Extraction**

Working on live data is the key to the optimal insight for the latest products. For this live review data is extracted using web-scrapping for the product which the user wants to buy.

**3.3 Integrating all onto a Web Page**

All the work done up till this point has been integrated in a web-page which makes the product reviewing easy for the user.

**4. EXPERIMENTAL DESIGN**

**4.1 DATA SET COLLECTION AND ANALYSIS**

In this analysis the Amazon review dataset has been used which consisted of only one attribute, the label for sentiment concatenated with the review as a string.[27]

**4.2 PERFORMANCE PARAMETERS**

In the research work 4 performance parameters are used to evaluate the performance of the NLP model.[25]

**4.2.1. Training Accuracy:**

Training Accuracy talks about how many correct predictions did our model made for the training dataset.[26]

**4.2.2. Validation Accuracy:**

Validation Accuracy talks about how many correct predictions did our model made for the validation dataset.[28]

**4.2.3. Training Loss:**

It is the difference between the actual and the predicted value at each epoch for the training dataset.[26]

**4.2.4. Validation Loss:**

It is the difference between the actual and the predicted value at each epoch for the validation dataset.[28]

**5. RESULTS**

The result gained after training and validating the compiled neural network model on the training and validation datasets are:

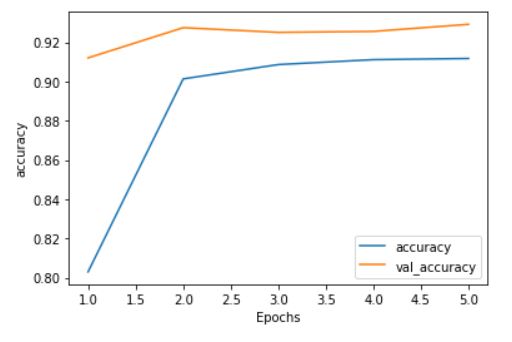


Fig 1: Training and Validation Accuracy

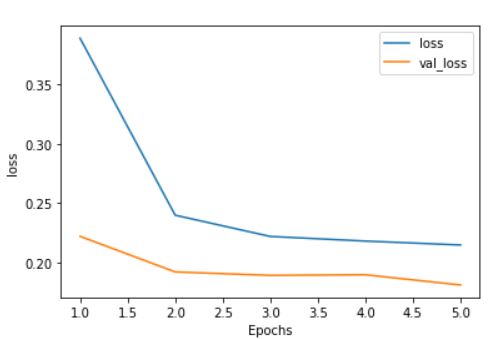


Fig2: Training and Validation Loss

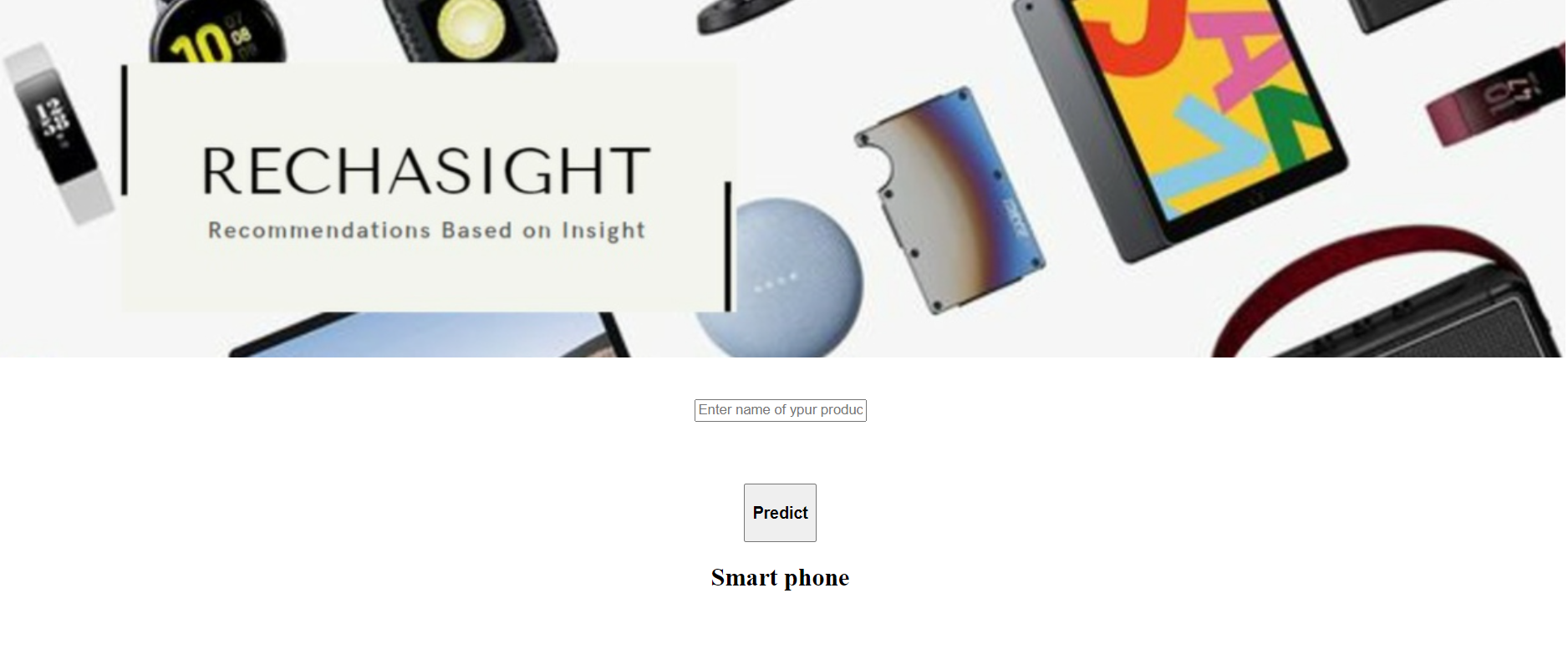


Fig 3: Website

**6. CONCLUSION**

Sentiment analysis is an important feature that enables to automate the process of labelling the reviews on the basis of the sentiments they contain. Thus, this neural network model will help by predicting what type of sentiment does a review conveys.

REFERENCES

[1] Dos Santos, C. and Gatti, M., 2014, August. Deep convolutional neural networks for sentiment analysis of short texts. In *Proceedings of COLING 2014, the 25th International Conference on Computational Linguistics: Technical Papers* (pp. 69-78).

[2] Severyn, A. and Moschitti, A., 2015, August. Twitter sentiment analysis with deep convolutional neural networks. In *Proceedings of the 38th International ACM SIGIR Conference on Research and Development in Information Retrieval* (pp. 959-962).

[3] Rain, C., 2013. Sentiment analysis in amazon reviews using probabilistic machine learning. *Swarthmore College*.

[4] Haque, T.U., Saber, N.N. and Shah, F.M., 2018, May. Sentiment analysis on large scale Amazon product reviews. In *2018 IEEE International Conference on Innovative Research and Development (ICIRD)* (pp. 1-6). IEEE.

[5] Tang, D., Qin, B. and Liu, T., 2015, September. Document modeling with gated recurrent neural network for sentiment classification. In *Proceedings of the 2015 conference on empirical methods in natural language processing* (pp. 1422-1432).

[6] Mikolov, T., Kombrink, S., Burget, L., Černocký, J. and Khudanpur, S., 2011, May. Extensions of recurrent neural network language model. In *2011 IEEE international conference on acoustics, speech and signal processing (ICASSP)* (pp. 5528-5531). IEEE.

[7] Webster, J.J. and Kit, C., 1992. Tokenization as the initial phase in NLP. In *COLING 1992 Volume 4: The 15th International Conference on Computational Linguistics*.

[8] Nasr, A., Béchet, F., Rey, J.F., Favre, B. and Le Roux, J., 2011. Macaon: An nlp tool suite for processing word lattices.

[9] Chen, Y., Perozzi, B., Al-Rfou, R. and Skiena, S., 2013. The expressive power of word embeddings. *arXiv preprint arXiv:1301.3226*.

[10] Li, J., Chen, X., Hovy, E. and Jurafsky, D., 2015. Visualizing and understanding neural models in nlp. *arXiv preprint arXiv:1506.01066*.

[11] Wang, J., Yu, L.C., Lai, K.R. and Zhang, X., 2016, August. Dimensional sentiment analysis using a regional CNN-LSTM model. In *Proceedings of the 54th Annual Meeting of the Association for Computational Linguistics (Volume 2: Short Papers)* (pp. 225-230).

[12] Wen, S., Wei, H., Yang, Y., Guo, Z., Zeng, Z., Huang, T. and Chen, Y., 2019. Memristive LSTM network for sentiment analysis. *IEEE Transactions on Systems, Man, and Cybernetics: Systems*.

[13] Zhou, P., Qi, Z., Zheng, S., Xu, J., Bao, H. and Xu, B., 2016. Text classification improved by integrating bidirectional LSTM with two-dimensional max pooling. *arXiv preprint arXiv:1611.06639*.

[14] Ding, Z., Xia, R., Yu, J., Li, X. and Yang, J., 2018, August. Densely connected bidirectional lstm with applications to sentence classification. In *CCF International Conference on Natural Language Processing and Chinese Computing* (pp. 278-287). Springer, Cham.

[15] Alhagry, S., Fahmy, A.A. and El-Khoribi, R.A., 2017. Emotion recognition based on EEG using LSTM recurrent neural network. *Emotion*, *8*(10), pp.355-358.

[16] Pelt, D.M. and Sethian, J.A., 2018. A mixed-scale dense convolutional neural network for image analysis. *Proceedings of the National Academy of Sciences*, *115*(2), pp.254-259.

[17] Bouchard, G., 2007. Efficient bounds for the softmax function, applications to inference in hybrid models. In *Presentation at the Workshop for Approximate Bayesian Inference in Continuous/Hybrid Systems at NIPS-07*.

[18] Gao, B. and Pavel, L., 2017. On the properties of the softmax function with application in game theory and reinforcement learning. *arXiv preprint arXiv:1704.00805*.

[19] Zhao, Y., Gao, X., Bates, D., Mullins, R. and Xu, C.Z., 2019. Focused quantization for sparse CNNs. In *Advances in Neural Information Processing Systems* (pp. 5584-5593).

[20] Cosentino, J., Zaiter, F., Pei, D. and Zhu, J., 2019. The Search for Sparse, Robust Neural Networks. *arXiv preprint arXiv:1912.02386*.

[21] Zhang, Z., 2018, June. Improved adam optimizer for deep neural networks. In *2018 IEEE/ACM 26th International Symposium on Quality of Service (IWQoS)* (pp. 1-2). IEEE.

[22] Bock, S., Goppold, J. and Weiß, M., 2018. An improvement of the convergence proof of the ADAM-Optimizer. *arXiv preprint arXiv:1804.10587*.

[23] Huang, J., Rathod, V., Sun, C., Zhu, M., Korattikara, A., Fathi, A., Fischer, I., Wojna, Z., Song, Y., Guadarrama, S. and Murphy, K., 2017. Speed/accuracy trade-offs for modern convolutional object detectors. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 7310-7311).

[24] Iandola, F.N., Han, S., Moskewicz, M.W., Ashraf, K., Dally, W.J. and Keutzer, K., 2016. SqueezeNet: AlexNet-level accuracy with 50x fewer parameters and< 0.5 MB model size. *arXiv preprint arXiv:1602.07360*.

[25] Khumaidi, A., Yuniarno, E.M. and Purnomo, M.H., 2017, August. Welding defect classification based on convolution neural network (CNN) and Gaussian kernel. In *2017 International Seminar on Intelligent Technology and Its Applications (ISITIA)* (pp. 261-265). IEEE.

[26] Garg, A., Noyola, J. and Bagadia, S., 2016. Lip reading using CNN and LSTM. *Technical report, Stanford University, CS231 n project report*.

[27] [Dataset].Kaggle,

<https://www.kaggle.com/bittlingmayer/amazonreviews> (Updated on November, 2019).

[28] Fang, X. and Zhan, J., 2015. Sentiment analysis using product review data. *Journal of Big Data*, *2*(1), p.5.