ISSN 2347 - 3983

Volume 8. No. 8, August 2020

International Journal of Emerging Trends in Engineering Research

Available Online at http://www.warse.org/IJETER/static/pdf/file/ijeter60882020.pdf

https://doi.org/10.30534/ijeter/2020/60882020

Machine Learning Based Sentiment Analysis on Twitter Data

Khushboo H. Saglani[1], Dr. Nitin J. Janwe[2]

[1]Rajiv Gandhi College of Engineering ,Research and Technology,Chandrapur, India, khushboosaglani1996@gmail.com

[2]Rajiv Gandhi College of Engineering ,Research and Technology,Chandrapur, India, nitinj_janwe@yahoo.com

ABSTRACT

This paper addresses the problem of sentiment analysis in twitter; that is characterizing tweets as per the feeling communicated in them: positive, negative, or impartial. Twitter is online small-scale blogging and person to person communication stage which permits clients to compose short notices of greatest length 140 characters. It is a quickly extending administration with more than 200 million enlisted clients - out of which 100 million are dynamic clients and half of them sign on twitter day by day - creating almost 250 million tweets for every day. Because of this huge measure of use, we would like to accomplish an impression of open estimation by dissecting the notions communicated in the tweets. Tearing down the open evaluation is critical for certain applications, for instance, firms endeavouring to find the response of their things in the market, predicting political choices and envisioning monetary wonders like the stock exchange. The point of this task is to build up a practical classifier for exact and programmed feeling grouping of an obscure tweet stream by utilizing the convolutional neural system.

Key words: Convolutional Neural Network (CNN), Deep Learning Machine Learning, Sentiment Analysis, Twitter Sentiment Analysis.

1. INTRODUCTION

We have decided to work with twitter since we feel it is a superior estimate of open conclusion rather than regular web articles and web websites. The explanation is that the measure of pertinent information is a lot bigger for twitter when contrasted with conventional blogging locales. Besides, the reaction on twitter is a guide and increasingly broad (since the quantity of clients who tweet is considerably more than the individuals who compose web writes day by day). Notion examination of open is profoundly basic in large scale financial marvels like anticipating the securities exchange pace of a firm. This should be possible by dissecting by and large open assumption towards that firm concerning time and utilizing financial aspects apparatuses for finding the connection betweenness open assessment and the company's securities exchange esteem. Firms can likewise gauge how well their item is reacting in the market, which regions of the reaction (since twitter permits us to download stream of geo-labelled tweets for specific areas. In case organizations can get this data, they can dissect the purposes for can get this data, they can dissect the purposes for topographically separated reaction, thus they can advertise their item in a more streamlined way by searching for fitting arrangements like making appropriate market portions. Foreseeing the aftereffects of well-known political decisions and surveys is additionally a rising application to assumption examination. One such investigation was led by Tumasjan et al. in Germany for foreseeing the result of government decisions in which reasoned that twitter is a decent impression of disconnected assessment.

of dissecting assumptions of tweets goes under the area of “Data Mining” & “Pattern Classification”. The two terms are firmly related and interlaced, and they can be officially characterized as the way toward finding "helpful" designs in a huge arrangement of information, either naturally (unsupervised) or semi-consequently (supervised). The paper would heavily rely on techniques of “Convolutional Neural Network” in extracting significant patterns and features from the large data set of tweets and on “Machine Learning” techniques for accurately classifying individual unlabeled data samples (tweets) as indicated by whichever design model best describes them possible.

2. LITERATURE REVIEW

Authors give an examination and an overall examination of existing strategies for supposition mining like AI and jargon-based procedures, together with appraisal estimations. Using distinctive AI computations like Max Entropy, Naive Bayes, & Support Vector Machine, Authors gave a study trendy twitter data creeks overall difficulties and usages of Sentiment Analysis on Twitter are also analyzed. [2]

Authors research the utility of etymological highlights for distinguishing the supposition of Twitter messages. We assess the convenience of existing lexical assets just as highlights that catch data about the casual and innovative language utilized in microblogging. Creators adopt a regulated strategy to the issue, yet influence existing hashtags in the Twitter information for building preparing information. [3]

Authors centre around issues and present a bound together structure for grouping tweets utilizing a half breed order conspire. The proposed strategy targets improving the presentation of Twitter-based opinion examination frameworks by consolidating an improved domain-specific classifier. In the wake of applying the preprocessing steps, the information content is gone through the emoji and slang classifiers. In the following stage, SentiWordNet-based and domain-specific classifiers are applied to characterize the content all the more precisely. At long last, feeling order is performed at sentence and record levels. The discoveries uncovered that the proposed technique beats the impediments of past strategies by thinking about slang, emojis, and domain-specific terms. [4]

Authors attempt to examine the twitter posts about electronic items like mobiles, workstations and so on utilizing Machine Learning approach. By doing an assessment examination in a particular space, it is conceivable to distinguish the impact of area data in conclusion characterization. Authors present another component vector for ordering the tweets as constructive, pessimistic and remove people groups' conclusion about items. [6]

Another half breed supposition arrangement for Twitter is proposed by installing a component choice technique. A correlation of the precision of the order by the PCA, LSA, and RP include determination techniques are introduced in this paper. Besides, the half breed assumption arrangement was approved utilizing Twitter datasets to speak to various areas, and the assessment with various order calculations additionally exhibited that the new crossover approach created significant outcomes. The usage indicated that the new half and half estimation order had the option to improve the precision execution from the current gauge slant characterization strategies by 76.55, 71.62 and 74.24%, individually. [7]

Notwithstanding the feeling terms utilized by and large reason assumption investigation, Authors coordinate emojis, modifiers and space explicit terms to dissect the audits posted in online networks. To test the adequacy of the proposed strategy, Authors considered clients audits in three areas. The outcomes got from various examinations show that the proposed technique conquers impediments of past strategies and the exhibition of the opinion investigation is improved in the wake of thinking about emojis, modifiers, invalidations, and space explicit terms when contrasted with benchmark strategies. [9][16]

The Naive Bayes and Neural Network classifier are consolidated for slant grouping. In Experimental outcomes, the film survey is grouped into positive or negative polarities of feeling utilizing classifiers. The precision of slant investigation is expanded up to 80.65% by joining Naive Bayes classifier with Neural Network for unigram include on film audit dataset. Catchphrases: Sentiment Analysis, Sentiment Classification, Naive Bayes, Neural Network Classifier. [8][10]

Twitter is an utmost well-known long-range social communication sites where individuals are communicating their perspectives, supposition and feelings generously. These tweets are recorded and broke down to mine feelings of individuals identified with a psychological oppressor assault (Uri assault). Present examination recovers tweets about Uri assault and discovers feelings and extremity of tweets. To mine feelings and extremity in tweets, content mining systems are utilized. Roughly 5000 tweets are recoded and pre-handled to make a dataset of now and again showing up words. R is utilized for mining feelings and extremity. The exploratory outcome demonstrated that 94.3% of individuals were disturbed by Uri assault.[12][20]

Premise prerequisites of sentiment mining to investigate the current strategies used to build up a full-fledged framework. Is features the chances or organization and research of such frameworks. The accessible apparatuses utilized for building such applications have even given their benefits and constraints. [13][17]

To deal with foresee slant for three diverse Twitter datasets. Our outcomes show a normal increment of F consonant precision score for recognizing both negative and positive supposition of around 6.5% and 4.8% over the baselines of unigrams and the grammatical feature includes separately. Likewise, look at against a methodology dependent on assumption bearing subject examination, and locate that semantic highlights produce better Recall and F score when grouping negative estimation and better Precision with lower Recall and F score in positive assumption arrangement. [14]

TwiSent, a notion examination framework for Twitter. In light of the subject looked, TwiSent gathers tweets relating to it and classifies them into the diverse extremity classes positive, negative and target. Be that as it may, investigating miniaturized scale blog entries have numerous intrinsic difficulties contrasted with the other content genres. The framework execution is assessed on physically explained highest quality level information and on a naturally commented on tweet set dependent on hashtags. It is a typical practice to show the viability of an administered framework on a naturally clarified dataset. Nonetheless, we show that such a framework accomplishes lesser arrangement accuracy when tried on nonexclusive twitter dataset. We likewise show that our framework performs far superior to a current framework. [15][21]

There are various online classification tools for the sentiment analysis. Authors used the Arabic opinion collections and classify them using the most efficient online sentiment analysis tools ie, Paralleldots and Repustate. The results shows that the Paralleldots works better than the Repustate for Arabic sentiment analysis.[22]

In data analysis, its crucial that the data to be analyzed is in offline mode or live streaming(real-time). Kyeongjoo Kim et.al[23] greatly discusses about streaming data. They have worked and analyzed on streaming data(real-time data analysis).

3. METHODOLOGY

3.1 Dataset

We are using Sentiment140 dataset for this experiment. Sentiment140 isn't open-source, yet there are assets with open source code with a comparative execution. The methodology was special because our preparation information was consequently made, instead of having people manual clarify tweets. In our methodology, we accept that any tweet with positive emojis, like :), was certain, and tweets with negative emojis, as :(, were negative. We utilized the Twitter Search API to gather these tweets by utilizing watchword search.

The information is a CSV with emojis expelled. The information record design has 6 fields:

0 – polarity
1 – id
2 – date
3 – query
4 – user
5 – text

3.2 Word2Vec

Word2Vec is a model to shape/make word Embeddings. It's a cutting-edge method for speaking to Word, wherein each word is spoken to by a vector (Array of numbers dependent on Embedding Size). Vectors are only loading of neurons, so if we set neurons of size 100, at that point we will have 100 loads and those loads are our Word Embeddings or just Dense vector.

Word embedding is solitary of the utmost notable portrayal of archive jargon. It is equipped for catching setting of a word in a report, semantic and syntactic comparability, connection with different words, and so forth.

Word2Vec is a technique to develop such an inserting. It may be superbly acquired utilizing two techniques: Common Bag of Words & Skip Gram.

3.2.1. Common Bag of Words Model

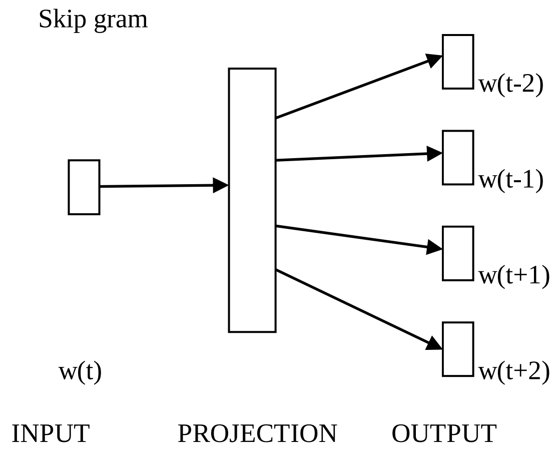
This technique takes the setting of each word as the info and attempts to anticipate the word relating to the unique circumstance. Think about our model: Have an extraordinary day.

Leave the contribution to the Neural Network alone the word, amazing. Notice that here we are attempting to anticipate an objective word utilizing a solitary setting input word incredible. All the more explicitly, we utilize the one-hot encoding of the information word and measure the yield blunder contrasted with the one-hot encoding of the objective word. During the time spent anticipating the objective word, we get familiar with the vector portrayal of the objective word.

Let us look further into the visual representation.

- We can see that W(t) is being anticipated given that the setting/neighbouring words are taken care of as info. P-esteem is assessed for target word dependent on the info that is taken care of to the system.
- We've got a shallow neural system here. Size of each word inserting or individual thick vector relies upon the number of neurons we've at removal.
- These vectors are loads learnt by every neuron.
- Despite how enormous or little the word is, it will be spoken to by the size of implanting we set.

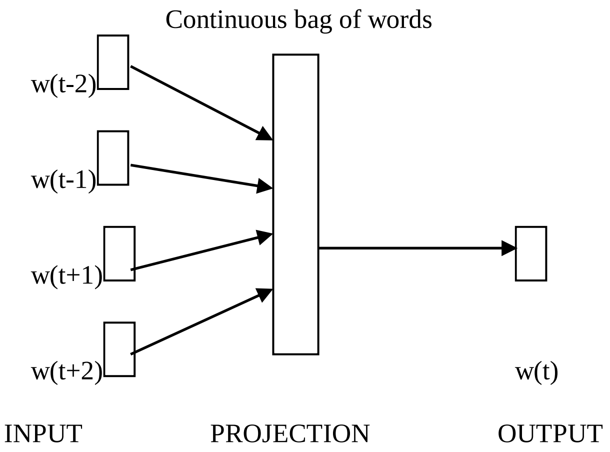
3.2.2. Skip Gram



This resembles the different setting CBOW model just got flipped. Somewhat that is valid. We input the objective word into the system. The model yields C likelihood circulations. For every setting position, we get C likelihood dissemination of V probabilities, one for each word.

Figure 2: Skip Gram

- A Target word is taken care of to our shallow neural system, loads are found out from concealed layer. A single word is picked indiscriminately from the neighbouring words. Moreover, P-value is being determined for setting words being near the info word we took care of to the system.



3.3 Convolutional Neural Network (ConvNet/CNN)

It is a Deep Learning calculation which can take in an information picture, dole out significance to different viewpoints/protests in the picture and have the option to separate one from the other. The pre-handling required in a ConvNet is a lot of lower when contrasted with other arrangement calculations. While in crude strategies channels are hand-built, with enough preparing, ConvNets can get familiar with these channels/attributes.

The engineering of a CNN is closely resembling that of the system case of neurons in the human mind and was moved by the relationship of the visual cortex. Singular neurons react to improvements just in a confined district of the visual field known as the Receptive Field. An assortment of such fields covers to cover the whole visual zone.

CNN models were created for picture characterization, in which the model acknowledges two-dimensional info speaking to a picture's pixels and shading channels, in a procedure called include learning.

This equivalent procedure can be applied to one-dimensional groupings of information. The model concentrates highlights from successions information and maps the inner highlights of the grouping. A 1D CNN is viable for getting highlights from a fixed-length fragment of the general dataset, where it isn't so significant where the component is situated in the section.

1D Convolutional Neural Networks function admirably for:

- Examination of a period arrangement of sensor information.
- Examination of sign information over a fixed-length period, for instance, a sound account.
- NLP although RNNs which LSTM cells are more encouraging than CNN as they consider the vicinity of words to make trainable examples.

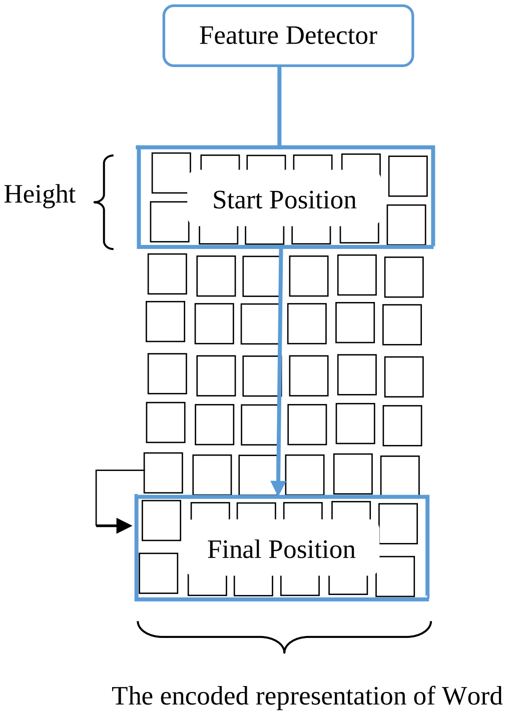


Figure 3: Example of Convolution-1D

4. IMPLEMENTATION

The experiment is done using the Keras framework in the TensorFlow using TensorFlow as backend and gensim to implement word2vec model.

4.1 Word2Vec:

Preparing Word2Vec the model is genuinely direct. Simply start up Word2Vec and pass the surveys that we read in the past advance. Along these lines, we are passing on a rundown of records. Where each rundown inside the fundamental rundown contains a lot of tokens from a client audit. Word2Vec utilizes every one of these tokens to inside making a jargon. Also, jargon means a lot of one of a word.

To prepare the model prior, we needed to set a few parameters. Presently, how about we attempt to comprehend what some of them mean. For reference, this is the order that we used to prepare the model.

(i) size:

It is a thick vector to speak to every token or word (for example the unique circumstance or neighbouring words). On the off fortuitous, you have controlled information, at that point, the size ought to be a lot littler incentive since you would just have such a large number of one of a kind neighbours for a given word. On the off fortuitous, you have bunches of information, it's acceptable to try different things with different sizes.

An estimation of 100–150 has functioned admirably for me for comparability queries.

(ii) window

The most extreme separation between the objective world and its neighbouring word. If your neighbour's position is more noteworthy than the greatest window width to one side or the right, at that point, a few neighbours would not be considered as being identified with the objective word. In principle, a petite window would give you terms that are increasingly related. Once more, on the off fortuitous, your information isn't scanty, at that point the window size ought not to make any difference to an extreme, as long as it's not excessively restricted or excessively expansive. If you are not very secure with this, simply utilize the default esteem.

(iii) min_count

Minimum recurrence check of words. The model would overlook words that don't fulfil the min_count. Incredibly inconsistent words are normally irrelevant, so it’s best to dispose of those. Except if your dataset is extremely little, this doesn't generally influence the model regarding your conclusive outcomes. The settings here most likely has a greater amount of an impact on memory utilization and capacity necessities of the model documents.

(iv) workers

Number of strings to use in the background.

(v) iter

The number of cycles over the corpus. 5 is a decent beginning stage.

4.2 Conv1D - 1D convolution layer:

This layer forms a convolution core that is convolved with the tier input over a sole spatial feature to yield a tensor of yields Lastly if activation stays None, it is pragmatic to the outputs equally fit.

After using this layer in place of the initial layer in the system, offer an input shape argument, e.g. input shape=(10, 128) designed for time series systems of 10-time stages with 128 features per stage in data_format="channels_last", or (None, 128) designed for variable-length systems through 128 features per phase.

(i) arguments:

a. filters
   A number, the dimensionality of the yield space.
b. kernel_size
   A number or tuple/list of a solitary number, determining the length of the window.
c. strides
   A number or tuple/ list of a solitary number, determining the step length.
d. padding
   Helpful when displaying fleeting information where the model ought not to disregard the worldly request.
e. activation
   To use an Activation function.
f. use_bias
   Boolean, regardless of whether the layer utilizes a bias vector.

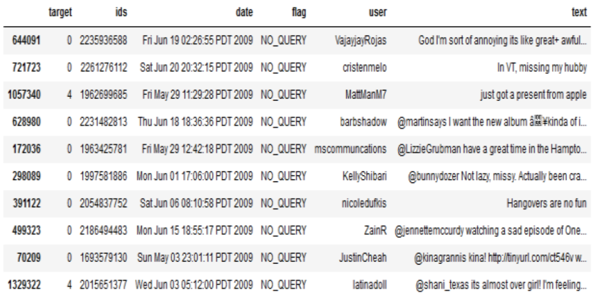
(ii) Input shape

3D tensor with input shape is as follows:
(batch, steps, channels)

(iii) Output shape

3D tensor with output shape is as follows:
(batch, new_steps, filters)

Following are the figures which show the dataset, label distribution and model structure used in the paper:



	target	ids	date	flag	user	text
644091	0	2235936588	Fri Jun 19 02:26:55 PDT 2009	NO_QUERY	VajayjayRojas	God I'm sort of annoying its like great+ awful...
721723	0	2261276112	Sat Jun 20 20:32:15 PDT 2009	NO_QUERY	cristenmelo	In VT, missing my hubby
1057340	4	1962699685	Fri May 29 11:29:28 PDT 2009	NO_QUERY	MattManM7	just got a present from apple
628980	0	2231482813	Thu Jun 18 18:36:36 PDT 2009	NO_QUERY	barbshadow	@martinsays I want the new album â¥kinda of i...
172036	0	1963425781	Fri May 29 12:42:18 PDT 2009	NO_QUERY	mscommuncations	@LizzieGrubman have a great time in the Hampto...
298089	0	1997581886	Mon Jun 01 17:06:00 PDT 2009	NO_QUERY	KellyShibari	@bunnydozer Not lazy, missy. Actually been cra...
391122	0	2054837752	Sat Jun 06 08:10:58 PDT 2009	NO_QUERY	nicoledufkis	Hangovers are no fun
499323	0	2186494483	Mon Jun 15 18:55:17 PDT 2009	NO_QUERY	ZainR	@jennettemccurdy watching a sad episode of One...
70209	0	1693579130	Sun May 03 23:01:11 PDT 2009	NO_QUERY	JustinCheah	@kinagrannis kina! http://tinyurl.com/ct546v w...
1329322	4	2015651377	Wed Jun 03 05:12:00 PDT 2009	NO_QUERY	latinadoll	@shani_texas its almost over girl! I'm feeling...

Figure 4: Data present in the dataset.

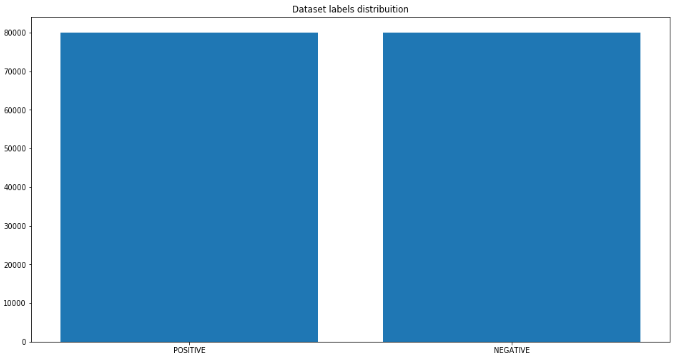
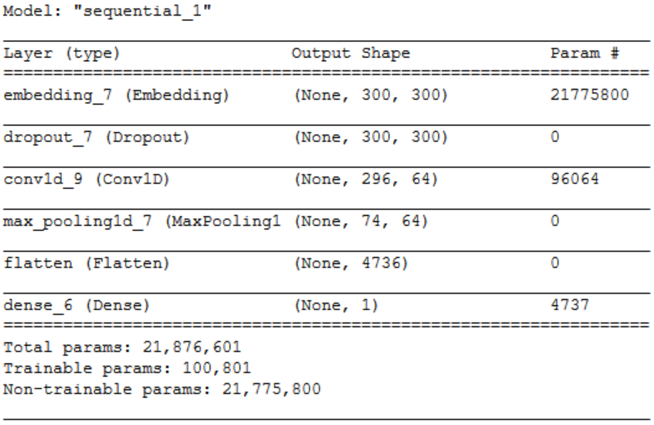


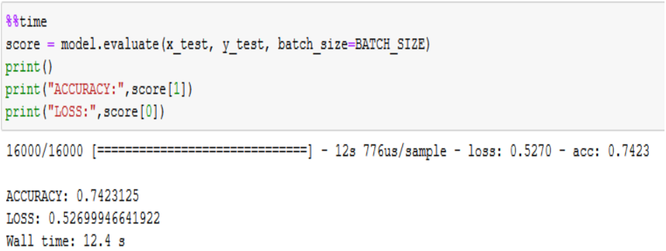
Figure 5: Dataset labels distributions



```
Model: "sequential_1"
_________________________________________________________________
Layer (type)                 Output Shape              Param #
=================================================================
embedding_7 (Embedding)      (None, 300, 300)          21775800
_________________________________________________________________
dropout_7 (Dropout)          (None, 300, 300)          0
_________________________________________________________________
conv1d_9 (Conv1D)            (None, 296, 64)           96064
_________________________________________________________________
max_pooling1d_7 (MaxPooling1 (None, 74, 64)            0
_________________________________________________________________
flatten (Flatten)            (None, 4736)              0
_________________________________________________________________
dense_6 (Dense)              (None, 1)                 4737
=================================================================
Total params: 21,876,601
Trainable params: 100,801
Non-trainable params: 21,775,800
_________________________________________________________________
```

Figure 6: Model Structure

While experimenting we observed some results that are shown in the figure 7, figure 8.a, figure 8.b,figure 9, figure 10 ,figure 11 and table 1 below and table 2 shows the comparison of the different models along with their accuracy.

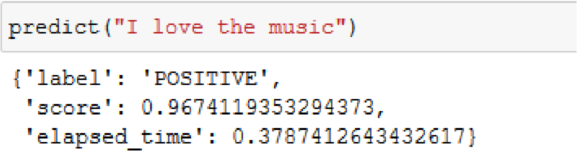


```
%%time
score = model.evaluate(x_test, y_test, batch_size=BATCH_SIZE)
print()
print("ACCURACY:",score[1])
print("LOSS:",score[0])
```

```
16000/16000 [==============================] - 12s 776us/sample - loss: 0.5270 - acc: 0.7423

ACCURACY: 0.7423125
LOSS: 0.52699946641922
Wall time: 12.4 s
```

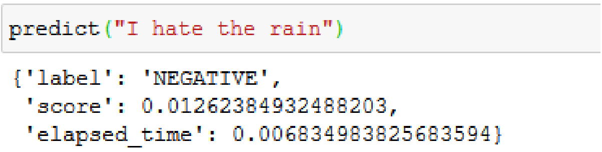
Figure 7: Evaluation of Model



```
predict("I love the music")
```

```
{'label': 'POSITIVE',
 'score': 0.9674119353294373,
 'elapsed_time': 0.3787412643432617}
```

Figure 8.a: Testing Model for runtime query



```
predict("I hate the rain")

{'label': 'NEGATIVE',
 'score': 0.01262384932488203,
 'elapsed_time': 0.006834983825683594}
```

Figure 8.b: Testing Model for runtime query

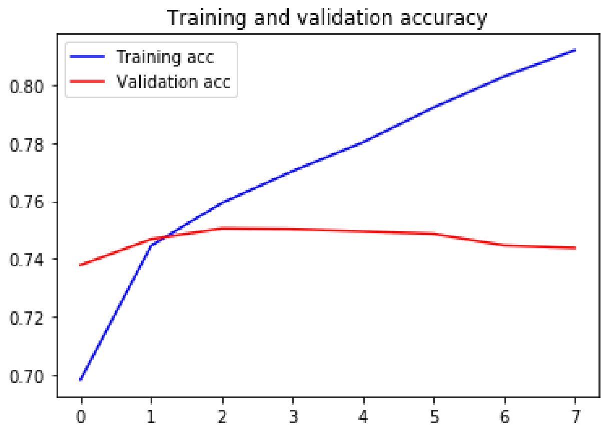


Figure 9: Training and validation accuracy

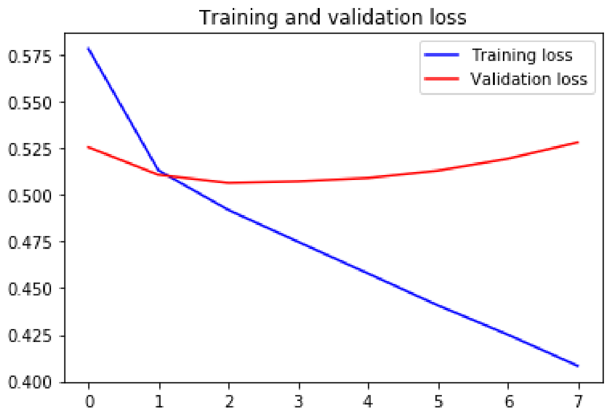


Figure 10: Training and validation loss

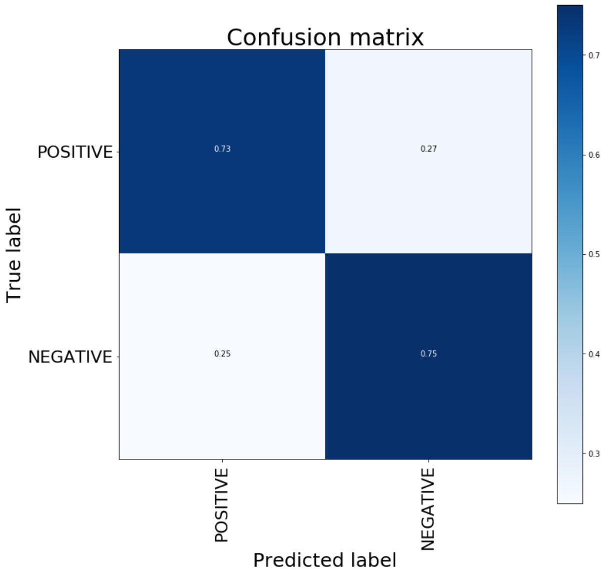


Figure 11: Confusion Matrix

Table 1: Custom Matrix

	Precision	Recall	F1-Score	Support
NEGATIVE	0.75	0.73	0.74	8048
POSITIVE	0.74	0.75	0.74	7952
Accuracy	-	-	0.74	16000
Macro Avg	0.74	0.74	0.74	16000
Weighted Avg	0.74	0.74	0.74	16000

Table 2: Comparisons of Model

Model	Accuracy
Decision Tree	59.64%
Naïve Bayes	56.40%
SVM	54.95%
CNN(My Method)	**74.23%**

5. CONCLUSION

Unarguably, sentimental analysis methods are among the most extreme noteworthy bases in the dynamic procedure. Many individuals rely upon sentimental analysis for accomplishing effective significances of services or products. We began with a model that was good in delivering slant investigation. In this way, applying a pre-prepared language model to beat the forefront look into the scholarly world too. Human dialects are moderately intricate to be comprehended by the machine, which prompts conditions where a contrarily said word has a positive affiliation and the other way around. In this way, a wistful study of Twitter information was a difficult assignment. In this analysis, we utilized the convolution neural network, which drives us to accomplish the assignment of sentiments mining from the tweets, which was prepared on "Sentiment140" Dataset. The prepared framework achieved to achieve an at last extraordinary final precision. For the Twitter slant examination, a convolution neural system model has been made utilizing the neural system having six layers (4 hidden, 1 input, & 1 output). The accuracy of this model has reached 74.23%.

REFERENCES

1. Twitter, **"Second Quarter 2016 Report,"** ed, 2016.
2. V. Kharde and P. Sonawane, **"Sentiment analysis of twitter data: A survey of techniques,"** arXiv preprint arXiv:1601.06971, 2016.
3. E. Kouloumpis, T. Wilson, and J. D. Moore, **"Twitter sentiment analysis: The good the bad and the omg!"** *Icwsm*, vol. 11, no. 538-541, p. 164, 2011.
4. M. Z. Asghar, F. M. Kundi, S. Ahmad, A. Khan, and F. Khan, **"T-SAF: Twitter sentiment analysis framework using a hybrid classification scheme,"** Expert Systems, vol. 35, no. 1, 2018.
5. M.Lovelin Ponn, Felciah, R.Anbuselvi**, "A Study on Sentiment Analysis of Social Media Reviews,"** *IEEE Second Conference on Innovations in Information Embedded & Communication Systems*, 2015.
6. Neethu M.S and Rajasree R., **"Sentiment Analysis in Twitter using Machine Learning Techniques",** *Fourth International Conference on Computing,*

Communications and Networking Technologies (ICCCNT), 2013.

7. N. Zainuddin, A. Selamat, and R. Ibrahim, **"Hybrid sentiment classification on twitter aspect-based sentiment analysis,"** *Applied Intelligence,* pp. 1-15, 2017.
8. M. Z. Asghar, F. M. Kundi, S. Ahmad, A. Khan, and F. Khan, **"T-SAF: Twitter sentiment analysis framework using a hybrid classification scheme,"** Expert Systems, vol. 35, no. 1, 2018.
9. M. Z. Asghar, A. Khan, S. Ahmad, M. Qasim, and I. A. Khan, **"Lexicon enhanced sentiment analysis framework using rule-based classification scheme,"** PloS one, vol. 12, no. 2, p. e0171649, 2017.
10. Lina L. Dhande, DR. Girish K. Patnaik **"Review of Sentiment Analysis using Naive Bayes and Neural Network Classifier".**
11. Prof. Ronen Feldman Hebrew University, **"JERUSALEM Digital Trowel, Empire State Building SENTIMENT ANALYSIS TUTORIAL ".**
12. Dipak R. Kawade, Dr.Kavita S. Oza **" Sentiment Analysis: Machine Learning Approach".**
13. Asmita Dhokrat, Sunil Khillare, C. Namrata Mahender, **"Review on Techniques and Tools used for Opinion Mining,"** *IJCAT,* 2015.
14. H. Saif, Y. He, and H. Alani, Proceeding of the Workshop on Information Extraction**"Semantic Sentiment Analysis of Twitter,"** and *Entity Analytics on Social Media Data. United Kingdom: Knowledge Media Institute,* 2011.
15. Subhabrata Mukherjee, Akshat Malu**, "TwiSent: A MultiStage System For analyzing Sentiment in Twitter,"** *Proceedings of the 21st ACM international conference on Information and knowledge management,* 2012.
    https://doi.org/10.1145/2396761.2398684
16. Bing Liu**., "Sentiment Analysis and Opinion Mining"**, Morgan & Claypool Publishers, May 2012.
17. G.Vinodhini and RM.Chandrasekaran, "**Sentiment Analysis and Opinion Mining: A Survey"**, Volume 2, Issue 6, June 2012 ISSN: 2277 128X *International Journal of Advanced Research in Computer Science and Software Engineering.*
18. Scott, J. (2011**). Social network analysis: developments, advances, and prospects. Social network analysis and mining**, 1(1), 21-26.
19. Ortigosa, A., Martín, J. M., & Carro, R. M. (2014). **Sentiment analysis in Facebook and its application to e-learning.** *Computers in Human Behavior*, 31, 527-541.
20. Poria, S., Cambria, E., & Gelbukh, A. (2015). **Deep convolutional neural network textual features and multiple kernels learning for utterance-level multimodal sentiment analysis.** *In Proceedings of the 2015 Conference on Empirical Methods in Natural Language Processing* (pp. 2539-2544).
21. Dos Santos, C. N., & Gatti, M. (2014, August**). Deep Convolutional Neural Networks for Sentiment Analysis of Short Texts.**
22. Hayel Khafajeh **"Opinion Mining: How efficient are Online classification tools?"***IJETER,* Vol 8. No.2 February 2020.
    https://doi.org/10.30534/ijeter/2020/46822020
23. Kyeongjoo Kim, Jihyun Song, Minsoo Lee " **Real-time Streaming Data Analysis using Spark"** *IJETER*, Vol 6, No.1 January 2018.
    https://doi.org/10.30534/ijeter/2018/01612018