

Predicting depression using deep learning and ensemble algorithms on raw twitter data

Nisha P. Shetty, Balachandra Muniyal, Arshia Anand, Sushant Kumar, Sushant Prabhu

Department of Information and Communication Technology, Manipal Institute of Technology,
Manipal Academy of Higher Education, India

Article Info

Article history:

Received Sep 11, 2019

Revised Jan 13, 2020

Accepted Feb 2, 2020

Keywords:

Depression

LSTM

Machine learning

Sentiment analysis

Social media

ABSTRACT

Social network and microblogging sites such as Twitter are widespread amongst all generations nowadays where people connect and share their feelings, emotions, pursuits etc. Depression, one of the most common mental disorder, is an acute state of sadness where person loses interest in all activities. If not treated immediately this can result in dire consequences such as death. In this era of virtual world, people are more comfortable in expressing their emotions in such sites as they have become a part and parcel of everyday lives. The research put forth thus, employs machine learning classifiers on the twitter data set to detect if a person's tweet indicates any sign of depression or not.

Copyright © 2020 Institute of Advanced Engineering and Science.
All rights reserved.

Corresponding Author:

Balachandra Muniyal,
Department of Information and Communication Technology,
Manipal Institute of Technology,
Manipal Academy of Higher Education,
Manipal-576104 India.
Email: bala.chandra@manipal.edu

1. INTRODUCTION

Social networking sites have become a habitual component, with sites like Twitter and Facebook being the 7th and 2nd favorite sites having millions of subscribers [1]. Such sites have become an open dais for people to reach out and express their feelings, likes, routines etc. Unlike mood fluctuations, depression is a common mental disorder which brutally affects a person's daily routine life. Any person who has undergone some adverse experiences like sudden death, unemployment etc. is liable to it. According to WHO, more than 300 million people of all age groups suffer from it, with only fewer than 10% receiving suitable treatment, owing to reasons such as lack of suitable health care, social humiliation and timely and right diagnosis [2].

Frances A et al. [3], states that the depression displays the following symptoms in the given order; sad mood, eluding all activities, weight and sleep fluctuations, body agitation, energy loss and tiredness, feeling of triviality, loss of decision-making capacity and finally suicidal tendencies. Emotion and sentiment analysis exercises machine learning algorithms to examine a text with respect to the emotion conveyed. Sentence level analysis is applied in this study to inspect if a tweet is emotionally vulnerable or not. Recently, the death of a 16-year-old Malaysian teen after calling for a poll in her Instagram [4] (where many people voted for "Death") has gathered huge media. Such incidents support the fact that if with proper monitoring of such sites [5] is done timely help and can be provided, hence avoiding such catastrophes.

2. LITERATURE SURVEY

Haque et al. [7] scrutinized 3D facial features and language spoken to gauge the intensity of depression. They compared their sentence level embedded convolutional neural network (CNN) model [8] with the existing works, but the data was collected by human computer interviews which lacked the precision of a formal diagnosis. Furthermore, the authors plan to include more parameters such as depression tallies from interviews taken for different periods of time.

Sharifa Alghowinem et al. [9] extracted and inspected the eye movement features for signs of depression. Their methodology employed support vector machines (SVM) and Gaussian Mixture Models for classification. However, owing to a small amount of data set they did not get a fairly high accuracy. In future the authors plan to integrate face, body and voice features to get a more apt detection.

Quan Hu et al. [10] assembled classification and regression models to inspect behavioral and dialectal features from social media for indications of depression. Their results can increase significantly by taking varied participants and increasing the observation period for improved analysis. Akkapon Wongkoblaph et al. [11] utilized deep learning model (5-fold cross validation [12]) to investigate the posts in social media. The obtained accuracy of 72 percent can be increased by including further features like interactions with friends, comments/replies etc. Mandar Deshpande and Vignesh Rao [13] applied natural language processing for analyzing the sentiments of the tweets. Due to inaccuracy of proper language style in social network the accuracy of the proposed model reduced.

Guntuku et al. [14] appraised the studies that predicted the mental health of people based on the survey responses, posts and groups interacted in social media. Tsugawa et al. [15] considered user activities in social media such as frequencies of words related to melancholy in a tweet, topics tweeted on, posting regularity etc. to check for the manifestation of depression. Nonetheless, by employing techniques like principle component analysis the feature set used could be improved. The methods such as deep learning and ensemble methods are expected to offer better results than SVM. Maryam Mohammed Aldarwish and Hafiz Farooq Ahmed [16] used SVM and Naïve Bayes Models on the preprocessed posts obtained from social network platforms. The accuracy obtained can be increased by training and constructing better models.

3. METHODOLOGY

Figure 1 highlights the methodology followed in the paper. The aim of the proposed work is to predict depression in individuals using their behavior online (on twitter specifically) [17-19]. This is done in two main stages. First being the stage where sentiment analysis [20] is applied on a particular individual's twitter posts to predict binary classes (i.e. depressed/not depressed). The twitter posts were obtained using the twitter API from a developer twitter account. A deep learning module known as long short-term memory (LSTM) [21, 22] is employed. The proposed LSTM model used a Kaggle dataset on twitter tweets related to depression to learn and validate.

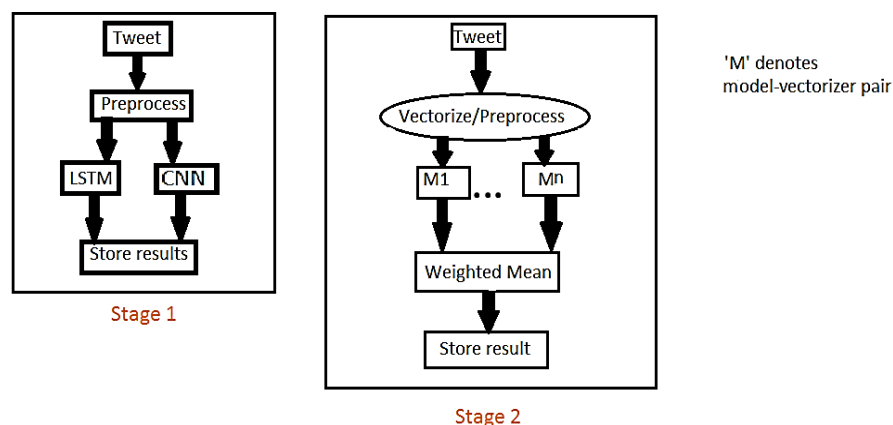


Figure 1. Overall methodology

Preprocessing as depicted in Figure 2 include dropping empty tweets, removing punctuations, creating a dictionary to map words to integer values and finding maximum length of the tweets [23]. Final generated features trimmed to the sequence length are fed into the model. The sequential LSTM model

network architecture encompasses an embedding layer, which has input dimension set to total number of tweets, output dimension set to 200 and input length set to sequence length as stated earlier. The next part includes an LSTM layer with 500 units, dropout 0.2 and recurrent dropout also 0.2. Last part of the model network includes a dense layer with one unit and a sigmoid activation to concise the generated output between zero and one. This completed the proposed model's architecture and the model is compiled using Adam optimizer and loss is dealt by using binary crossentropy with accuracy metric selection to observe and evaluate.

Post compilation, the model is fitted on the earlier generated features and validated on labels specified in the dataset. A validation split of 0.3 (70% training data, 30% test data) is introduced and this model is trained for 5 epochs. Finally this model is stored in a JSON file for future use. Every new tweet obtained from the twitter API goes through the same preprocessing procedure mentioned above before being forwarded to the LSTM model. The result of the model is obtained from the model.predict () function and is rounded to an integer value (0 or 1 in this case).

The obtained accuracy is compared to sequential Convolutional Neural Network (CNN) [24, 25]. The preprocessing steps are exactly the same. The network architecture includes an embedding layer similar to the LSTM embedding. The only difference is a weights argument is given an embedding matrix, with random values in the range of 200 to total number of tweets multiplied by 0.01. This is followed by a dropout of 0.4 which feeds the data to a total of four 1D convolution layers. Each of the convolution layers has kernel size set to 3, padding set to valid, activation is relu and strides is set to 1. Only thing that differed is the filter (dimensionality of the output space). It is decreased by 50% at each layer. First layer had filter set to 600, second had 300, third 150 and fourth had 75.

After this, flatten is included in the model architecture to flatten the input. This is followed by adding a dense layer with 600 units, dropout of 0.5, activation set to relu, a dense layer with one unit, and finally, an activation set to sigmoid. CNN model is compiled using exactly same arguments for loss, metric and optimizer. The second stage includes trying to improve the outcome of the proposed work using basic machine learning classifiers [26, 27] and a few optimized ensembles. Classifiers used are logistic regression, linear support vector classifier (SVC), multinomial naive bayes, bernoulli naive bayes along with ensembles like random forest classifier and gradient boosting classifier.

Preprocessing as depicted in Figure 3 includes splitting of data into train and test set. This is followed by vectorizing the tweets. The mentioned stage provides the classifiers with three different kind of vectorizers namely count vectorizer, TF-IDF and n-grams [28]. For Count Vectorizer, the preprocessing includes fitting the count vectorizer on the training and data then transforming the documents in the training data to document-term matrix.

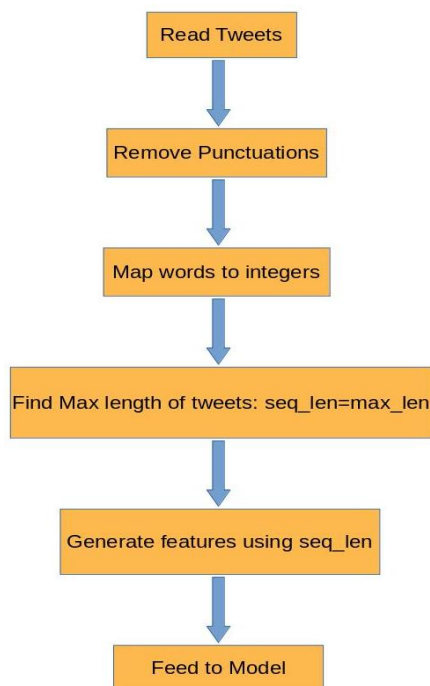


Figure 2. LSTM preprocessing

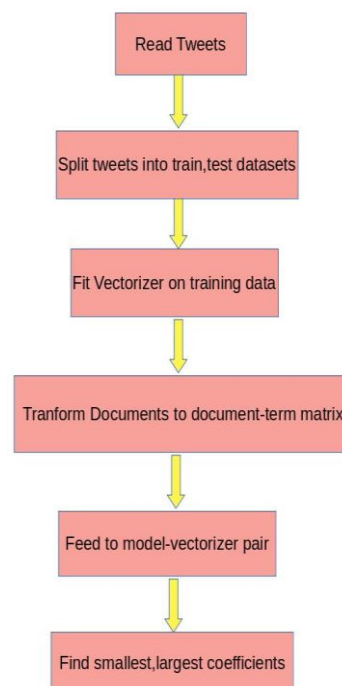


Figure 3. Model vector preprocessing

The obtained training data (features) is fed to the current model selected from the list. A feature names list is initialized and populated using Count Vectorizer's `get_feature_names()` function. Results are predicted using `predict` function. TF-IDF Vectorizer is fit on the training data using `min_df=5`, similar to Count Vectorizer, documents are transformed and the training data is fed to the current model selected from the list. Along with this, feature names and sorted TF-IDF index is also calculated to find the smallest and highest TF-IDF, (least and most important coefficients). For n-grams, similar procedure was followed and is fit on the training data using `min_df=5` and `n-gram_range(1,2)`. Similar to other vectorizers, smallest and highest coefficients are noted and result sentiment of the tweet is predicted.

Each and every classifier and ensemble mentioned above is fit on each of the three vectorizers and results are noted. This stage is largely focused on cross-checking predictions made by the first stage. Every new tweet obtained through the twitter API is sent through first and second stage. In the second stage every possibility of model and vectorizer is executed and results are noted. The final result included the weighted mean of all second stage possibilities. This value is cross-checked with the first stage predictions. The weights are assigned according to accuracy of the model on the data.

4. RESULT

The following tables Tables 1-7 show the results obtained in this research. The mentioned classifiers's results are compared here.

Table 1. LSTM vs CNN

Classifier	Validation Accuracy	Test Accuracy
LSTM	0.7	0.93
CNN	0.68	0.95

Table 2. Logistic regression

Classifier	Accuracy
Count vectorizer	75.81%
TF-IDF	76.22%
n-grams	76.15%

Table 3. Random forest

Classifier	Accuracy
Count vectorizer	70 %
TF-IDF	72%
n-grams	72%

Table 4. Bernoulli naïve bayes (NB)

Classifier	Accuracy
TF-IDF	73.95%
n-grams	75.53%

Table 5. Multinomial naïve bayes (NB)

Classifier	Accuracy
TF-IDF	74.223%
n-grams	76.69%

Table 6. Linear SVC (random_state=100, tol=1e-10)

Classifier	Accuracy
Count vectorizer	73.71%
TF-IDF	75.48%
n-grams	73.53%

Table 6. Gradient Boosting Classifier (subsample=0.8, learning_rate=0.05, n_estimators=250, random_state=5, max_depth=20, max_leaf_nodes=150)

Classifier	Accuracy
n-grams	74.42%

5. CONCLUSION AND FUTURE WORK

Often traditional survey-based questions fail to uncover the extent of users 'mental health depreciation. Nowadays, social media is a common platform, which people use to reach out. Therefore, the above proposed methodology cashes in this popularity of SNS and evaluates the depression levels of the users by employing natural language and machine learning techniques. In future, by collecting more data, more frequently with the aim of improving the accuracy of the work to give a better diagnosis. Such tools which can predict variations in person's mood can be an important method for both clinical observations and self-diagnosing. The method can be time consuming and hence steps must be taken in this regard to improve upon.

REFERENCES

- [1] Statista, "Most popular social networks worldwide as of October," 2019. [Online]. Available: <https://www.statista.com/statistics/272014/global-social-networks-ranked-by-number-of-users/>
- [2] WHO, "WHO Depression," 2019. [Online]. Available: <https://www.who.int/news-room/fact-sheets/detail/depression>
- [3] Frances, A., Pincus, H., & First, M., "Major Depressive Episode. In Diagnostic and statistical manual of mental disorders: DSM-IV," *Washington, DC: American Psychiatric Association*, 1994.
- [4] Mail online, "Malaysian teenager, 16, 'died jumping from the third floor of a shop after conducting an Instagram poll on whether she should kill herself' - and 69 per cent chose 'death'," 2019. [Online]. Available: <https://www.dailymail.co.uk/news/article7031067/Malaysian-teenager-16-killed-conducting-Instagram-poll.html>.
- [5] Serena Gordon, "Facebook Posts May Hint at Depression," 2018. [Online] Available: <https://www.webmd.com/depression/news/20181015/facebook-posts-may-hint-at-depression#1>
- [6] Jack Carfagno, "Social Media Posts Can Help Identify Depression," 2019. [Online]. Available: <https://www.docwirenews.com/docwire-pick/social-media-posts-can-help-identify-depression/>
- [7] Haque, Albert, Michelle Guo, Adam S. Miner and Li Fei-Fei, "Measuring Depression Symptom Severity from Spoken Language and 3D Facial Expressions," 2018. [Online] Available: *arXiv.org arXiv:1811.08592*
- [8] S. Albawi, T. A. Mohammed and S. Al-Zawi, "Understanding of a convolutional neural network," 2017 *International Conference on Engineering and Technology (ICET)*, pp. 1-6, 2017.
- [9] S. Alghowinem, R. Goecke, M. Wagner, G. Parker and M. Breakspear, "Eye movement analysis for depression detection," 2013 *IEEE International Conference on Image Processing, Melbourne*, pp. 4220-4224, 2013.
- [10] Q. Hu, A. Li, F. Heng, J. Li and T. Zhu, "Predicting Depression of Social Media User on Different Observation Windows," 2015 *IEEE/WIC/ACM International Conference on Web Intelligence and Intelligent Agent Technology (WI-IAT), Singapore*, pp. 361-364, 2015. DOI: 10.1109/WI-IAT.2015.166.
- [11] A. Wongkoblap, M. A. Vadillo and V. Curcin, "Classifying Depressed Users with Multiple Instance Learning from Social Network Data," 2018 *IEEE International Conference on Healthcare Informatics (ICHI)*, pp. 436-436, 2018. DOI: 10.1109/ICHI.2018.00088.
- [12] J. D. Rodriguez, A. Perez and J. A. Lozano, "Sensitivity Analysis of k-Fold Cross Validation in Prediction Error Estimation," in *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 32, no. 3, pp. 569-575, March 2010. doi: 10.1109/TPAMI.2009.187.
- [13] M. Deshpande and V. Rao, "Depression detection using emotion artificial intelligence," 2017 *International Conference on Intelligent Sustainable Systems (ICISS), Palladam*, pp. 858-862, 2017. DOI: 10.1109/ISS1.2017.8389299.
- [14] Sharath Chandra Guntuku, David B. Yaden, Margaret L. Kern, Lyle H. Ungar, Johannes C. Eichstaedt, "Detecting depression and mental illness on social media: an integrative review," *Current Opinion in Behavioral Sciences*, vol. 18, pp. 43-49, 2017.
- [15] Sho Tsugawa, Yusuke Kikuchi, Fumio Kishino, Kosuke Nakajima, Yuichi Itoh, and Hiroyuki Ohsaki, "Recognizing Depression from Twitter Activity," In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems (CHI '15)*. ACM, New York, NY, USA, pp. 3187-3196, 2015. DOI: <https://doi.org/10.1145/2702123.2702280>.
- [16] M. M. Aldarwish and H. F. Ahmad, "Predicting Depression Levels Using Social Media Posts," 2017 *IEEE 13th International Symposium on Autonomous Decentralized System (ISADS)*, pp. 277-280, 2017. DOI: 10.1109/ISADS.2017.41.
- [17] Farig Sadeque, Dongfang Xu, and Steven Bethard, "Measuring the Latency of Depression Detection in Social Media," In *Proceedings of the Eleventh ACM International Conference on Web Search and Data Mining (WSDM '18)*, pp. 495-503, 2018.
- [18] Burdisso, Sergio G., Marcelo E., and M. Montes-y-Gómez, "A Text Classification Framework for Simple and Effective Early Depression Detection over Social Media Streams," *arXiv.org arXiv:1905.08772*, pp. 1-20, 2019.
- [19] Viridiana Romero M., "A machine learning approach for the detection of depression and mental illness in Twitter," 2019. [Online]. Available: <https://medium.com/datadriveninvestor/a-machine-learning-approach-for-detection-of-depression-and-mental-illness-in-twitter-3f3a32a4df60>.
- [20] S. A. El Rahman, F. A. AlOtaibi and W. A. AlShehri, "Sentiment Analysis of Twitter Data," 2019 *International Conference on Computer and Information Sciences (ICCIS)*, pp. 1-4, 2019.
- [21] J. Wang and Z. Cao, "Chinese text sentiment analysis using LSTM network based on L2 and Nadam," 2017 *IEEE 17th International Conference on Communication Technology (ICCT)*, Chengdu, pp. 1891-1895, 2017. DOI: 10.1109/ICCT.2017.8359958.
- [22] F. Liu, X. Zhou, J. Cao, Z. Wang, H. Wang and Y. Zhang, "A LSTM and CNN Based Assemble Neural Network Framework for Arrhythmias Classification," *ICASSP 2019-2019 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, Brighton, United Kingdom, pp. 1303-1307, 2019. DOI: 10.1109/ICASSP.2019.8682299.
- [23] D. Zimbra, M. Ghiassi and S. Lee, "Brand-Related Twitter Sentiment Analysis Using Feature Engineering and the Dynamic Architecture for Artificial Neural Networks," 49th *Hawaii International Conference on System Sciences (HICSS)*, pp. 1930-1938, 2016. doi: 10.1109/HICSS.2016.244.
- [24] Abdelghani Dahou, Mohamed Abd Elaziz, Junwei Zhou, and Shengwu Xiong, "Arabic Sentiment Classification Using Convolutional Neural Network and Differential Evolution Algorithm," *Computational Intelligence and Neuroscience*, vol. 2019, pp. 1-16, 2019. DOI: <https://doi.org/10.1155/2019/2537689>.
- [25] Kim H. Jeong, Y.-S., "Sentiment Classification Using Convolutional Neural Networks," *Appl. Sci.*, vol. 9, 2019.

- [26] Hao Guo, Mengna Qin, Junjie Chen, Yong Xu, and Jie Xiang, "Machine-Learning Classifier for Patients with Major Depressive Disorder: Multifeature Approach Based on a High-Order Minimum Spanning Tree Functional Brain Network," *Computational and Mathematical Methods in Medicine*, vol. 2017, pp. 1-14, 2017.
- [27] Ishita Bhakta and Arkaprabha Sau, "Prediction of Depression among Senior Citizens using Machine Learning Classifiers," *International Journal of Computer Applications*, vol. 144, no. 7, pp. 11-16, 2016.
- [28] Ankit Basarkar, "Document Classification using Machine Learning," *San Jose State University*. California, 2017.

BIOGRAPHIES OF AUTHORS



Nisha P. Shetty has published in the areas network security and machine learning. Currently she is working in the area of social network security.



Dr. Balachandra's research area includes Network Security, Algorithms, and Operating systems. He has more than 30 publications in national and international conferences/journals. Currently he is working as the Professor and Head in the Dept. of Information & Communication Technology, Manipal Institute of Technology, Manipal. He has 25 years of teaching experience in various Institutes.



Arshia Anand is currently pursuing her bachelor's degree in Computer and Communication from MIT, Manipal. Her areas of interest include Data Science and Full Stack Development.



Sushant Kumar is currently pursuing his bachelor's degree in Computer and Communication Engineering from MIT Manipal. His areas of interests are Data Science and full-stack development.



Sushant Prabhu is currently pursuing B. Tech in Computer and Communication Technology (CCE) at MIT, Manipal. Currently working in Deep Learning (Graph Embedding Networks). Interested in areas of Deep Learning, Data Science, Analytics and its applications.