Proceedings of the International Conference on Intelligent Sustainable Systems (ICISS 2017)
IEEE Xplore Compliant - Part Number:CFP17M19-ART, ISBN:978-1-5386-1959-9

Depression Detection using Emotion Artificial Intelligence

Mandar Deshpande
Electrical and Electronics Engineering Department
Visvesvaraya National Institute of Technology
Nagpur, India
Email: mandar061095@gmail.com

Vignesh Rao
Department of Computer Science and Engineering
Visvesvaraya National Institute of Technology
Nagpur, India
Email: raovignesh0210@gmail.com

***Abstract*—Depression is a leading cause of mental ill health, which has been found to increase risk of early death. Moreover it is a major cause of suicidal ideation and leads to significant impairment in daily life. Emotion artificial intelligence is a field of ongoing research in emotion detection, specifically in the field of text mining. The advent of internet based media sources has resulted in significant user data being available for sentiment analysis of text and images. This paper aims to apply natural language processing on Twitter feeds for conducting emotion analysis focusing on depression. Individual tweets are classified as neutral or negative, based on a curated word-list to detect depression tendencies. In the process of class prediction, support vector machine and Naive-Bayes classifier have been used. The results have been presented using the primary classification metrics including F1-score, accuracy and confusion matrix.**

***Keywords*—Emotion Artificial Intelligence, Support Vector Machine, Naive Bayes, Depression Detection, Machine Learning, Natural Language Processing**

I. Introduction

Depression is a mental disorder which can impair many facets of human life. Though not easily detected it has profound and varied impacts [8]. In today's world, the stresses of daily life events may increase chances of depression. It's diagnosis is made if at least five of the below symptoms occur almost every day for at least 2 weeks:

1. Depressed Mood
2. Loss on interest in activities
3. Suicidal thoughts
4. Feeling of worthlessness or hopelessness
5. Worsened ability to think and concentrate

There might be other reasons like genes and family history which might also lead to depression.

Nowadays people tend to express their emotions, opinions and disclose their daily lives [9] through a variety of social media platforms like Twitter, Facebook and Instagram. This expression can be through images, videos and mainly through text. Due to the widespread presence and reach of these social media platforms, there is a plethora of user data available for undertaking explorative analysis. Textual data being the most widely used form of communication offers a bunch of characteristics which makes it the best choice for doing data analysis, for emotion AI.

Text data has the following benefits:
1. Easy to handle
2. Simple and quick to pre-process
3. Quantitative and qualitative availability
4. Significantly smaller memory storage size compared to image and video data

Twitter, which has a fixed limit on the amount of characters allowed in a single Tweet [7], proves to provide the best platform to apply emotion artificial intelligence for depression detection.

Emotion AI is an upcoming field of research in sentiment analysis, which aims to utilize machine learning techniques and algorithms for emotion detection. Successive work in the domain of emotion artificial intelligence will ultimately lead to breakthroughs in large scale opinion mining [13], market research and in diagnosing medical conditions [10]. Emotion AI is not limited to textual data, but can have wide applications in computer vision through image and video data, for facial expression detection. Furthermore advancements in recurrent neural network based models have led to state of art results in speech emotion artificial intelligence.

Determining the sentiment of an entire document is referred to as coarse level and fine level deals with attribute level sentiment analysis [6]. Sentence level emotion AI comes in between these two. Twitter being the data source of choice in this paper, mostly deals with Tweets which are short message which are bound by a 140 character limit [2]. In this concise format, users express their emotions and feelings about ongoing happenings in their life and the world around them.

Emotion AI has been applied on the collected and pre-processed Tweet data, which is classified into potential categorization of negative or neutral emotion state. Supervised learning is the machine learning task which involves providing the algorithm with labelled dataset, which is then used to learn model parameters (weight, bias). This paper implements Naive-Bayes and Support Vector Machines classifier for detecting Tweets which demonstrate signs of depression and emotional ill-health.

The structure of the remaining part of the research paper is as follows. A brief description of the classifiers used in the implementation has been presented in the Section II. Section III deals with the methodology of the research paper. Experimental results and discussion is presented in Section 4. A conclusion is provided in the final section 5.

II. Background

A. Dataset

The dataset comprises of tweets collected using the Twitter API. A total of 10,000 Tweets were collected for generating the training and test dataset. A ratio of 80:20 has been adopted for splitting the data collected into training and test dataset.

Two word-list were compiled for the training and test datasets. The training word-list comprised of curated list of words suggesting depression tendencies like 'depressed', 'hopeless', 'suicide'. For the test dataset, tweets were collected at random which included neutral as well as negative components.

B. Naive Bayes Classifier

Naive Bayes Classifier is a classifier which implements Bayes theorem with a solid (naive) independence assumptions[1], particularly, independent feature model. Bayes Theorem works on conditional probability which finds out the probability of an event given that some other event has already occurred. It predicts the conditional probability of a class given the set of evidences and finds the most likely class based on the highest one. A naive Bayes classifier is a famous and popular technique because it is very fast approach and gives a high accuracy[3]. The equation of the conditional probability is defined as follows:

$$P(H|M) = \frac{P(E1|H) * P(E2|H) * P(En|H) * P(H)}{P(M)}$$

Here,
H is the probability of a classification
E1 to En are the Evidence variables
M is the Set of all evidences

There are three types of Naive Bayes Classifier:
1. Gaussian Naive Bayes
2. Multinomial Naive Bayes
3. Bernoulli Naive Bayes.

In this paper, Multinomial Naive Bayes has been used as classifier. The Multinomial Naive Bayes works well on multinomially distributed data and is widely used in text classification.

C. Support Vector Machine

Support Vector Machine(SVM) is a supervised learning algorithm that analyzes the data and recognizing patters used for classification[4]. Given an input set, SVM classifies them as one or the other of two categories. SVM can deal both with both linear and non-linear classification.

With kernel trick, it can efficiently perform non-linear classification. It does so by mapping the input set to high dimensional feature space. The types of kernel includes polynomial, Gaussian radial basis function (RBF) ,Laplace RBF kernel, Hyperbolic tangent kernel and Sigmoidal kernel. Construction of hyper plane is employed by SVM for the classification.

In the context of this paper, the goal of a text classification system is to determine whether a given tweet belongs to a set predefined categories [12]. An optimal SVM algorithm for text classification does this via multiple optimal strategies.

III. Methodology

The approach being taken up in this paper is modular in its' organization. Individual component of the work flow have been segregated into stand-alone steps, to improve quality of implementation.

The work flow starts with data collection step, which utilizes Twitter API for generation of dataset. Natural language processing facilitates much better than average classification for sentimental analysis done by the human [14]. Following the creation of datasets, the data preprocessing module which systematically churns the data through tokenization, stemming and stop words removal. POS tagger then identifies essential pieces of the text to be utilized. After this the text classifier is trained on the processed text data from Twitter, in the training phase. In the testing phase, class prediction is made on the test dataset to identify potential Tweets demonstrating depression tendencies.

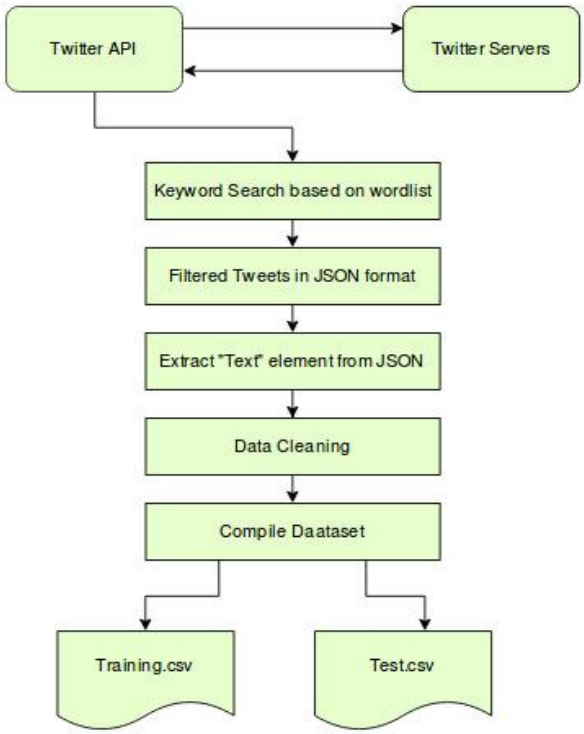

Fig. 1: Data Collection

A. Data Collection

Every machine learning or sentiment analysis task starts with collection of relevant data from various sources. In this paper Twitter is considered as the data source for analysis, in the form of User Tweets. This portion covers tasks from streaming the data from the Twitter servers, to compile the training and test datasets.

1. App Authentication: Application-only authentication involves communication between the application request and Twitter API, without a user context. This requires the creation of a Twitter app which is assigned a set of unique consumer key and secret key. Further to access the twitter data from

incoming streams, an unique access token and secret token needs to be supplied. This two way communication is handled using the Twitter API.

2. Keyword list: Using a pre-created wordlist for detecting trigger words symbolizing poor mental well-being, Tweets from all over the world are collected at random. These keyword specific tweets are mixed with a general batch of non-weighted Tweets, in form of JSON objects.

3. Extracting text from JSON: The collected tweets in the JSON objects are parsed to extract only the text field of the Tweets. Other meta-data related to any particular Tweet is removed.

4. Data Cleaning: To avoid errors in encoding textual data, the Text is purged for links(http) and non-ASCII characters like emoticons. Result is a clean dataset, rid of non-conformative character types.

5. Generate csv file for train and test set: The cleaned text data from individual Tweets is added to the training and test dataset, in a vectorized format. Classification labels for the training and test datasets are manually added, to create a csv file using comma as the delimiter.

B. Data Preprocessing

The csv file is read and several data preprocessing steps are performed on it. Natural language processing [11] has been utilized for preprocessing methods applied on the extracted data:

1. Tokenization: Tokenization is a process of dividing a string into several meaningfull substring,such as units of words,sentences, or themes[5]. In this case, the first column of the csv file containing the tweet is extracted and is converted into individual tokens.

2. Stemming: Stemming involves reducing the words to their root form. This would help us to group similar words together. For implementation,Porter Stemmer is used.

3. Stop Words Removal: The commonly used words ,known as stop words need to removed since they are of no use in the training and could also lead to erratic results if not ignored. Nltk library has a set of stop-words which can be used as a reference to remove stop-words from the tweet.

4. POS Tagger: To improve the quality of the training data, the tokenized text is assigned the respective parts of speech by using POS Tagger. This would be used to extract only the adjectives ,nouns and adverbs since other parts of speech are not of much significance. Example: 'I love coding' - 'love' being a noun is extracted, rest are removed.

After all these pre-processing steps, a bag of words is formed. Bag of words calculates the number of occurrence of each word, which is then used as a feature to train a classifier.

C. Training

The classifier requires two parameters: training set and label. The training set in this case is the set of tweets which needs to be further processed in order to feed into a classifier. The set of tweets need to converted into vector format for further processing. The set of labels corresponding to each tweet is also fed into the classifier in the form a vector.

Saving the Classifier and the Count Vectorizer Object: Since training needs to be done once, the trained classifier object needs to be loaded into a pickle file. Same is applicable with the Count Vectorizer object. Thus both these objects are dumped into a pickle file for further use.

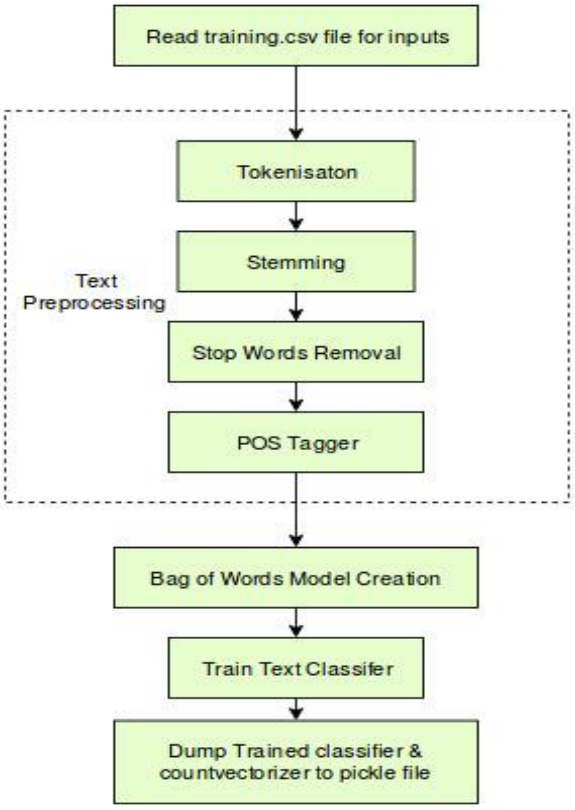

Fig. 2: Training phase

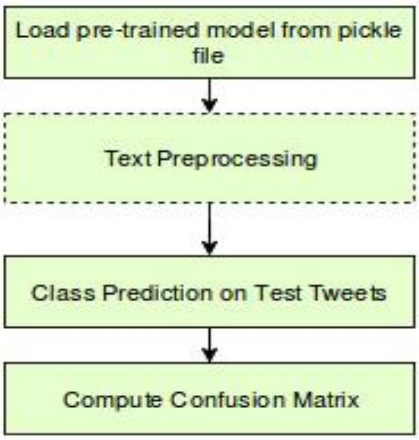

Fig. 3: Testing phase

D. Testing

Testing the classifier involves following steps:

1. Loading saved models: The trained classification models are loaded from the pickle file, to be used for prediction on test dataset.

2. Data Preprocessing: The test dataset is preprocessed in a manner similar to the training data.

3. Class Prediction on Test Tweets: Each tweet is classified into a depressed or neutral class.

4. Computation of Confusion Matrix: Based on the values of true or false positives and negatives we compute the confusion matrix, for the evaluation of classification performance.

IV. Results

The results are evaluated on the basis of F1 score and accuracy. The F1 score is the primary performance measure and accuracy is the secondary measure. F1 score is calculated based on the precision and recall.

$$F1\ score = 2 * \frac{P * R}{P + R}$$

. Here, P stands for Precision and R is the Recall.

It can be noticed from the results that Multinomial Naive Bayes has performed the best with the F1 score of 83.29 whereas SVM has achieved a lower F1 score of 79.73. The Precision and Recall follow the same trend with Multinomial Naive Bayes outperforming SVM. Fig.4 and Fig.5 shows the normalized confusion matrix. It consists of two rows and two columns in which various parameters like false positives,false negatives,true positives and true negatives can be analyzed. The accuracy of the Mutinomial Naive Bayes is 83% and is 79% in case of SVM.

The accuracy of the above classifiers is slightly less due to the fact that the tweets contain text which is not in standard format. For example, people write ty instead of Thankyou. Thus it is a bit challenging task to train the classifier and achieve significant results. This calls for further research on this area to improve the accuracy of the model.

TABLE I: Results

Comparison of performance metrics				
Name	Precision	Recall	F1 score	Accuracy
Multinomial Naive Bayes	0.836	0.83	0.8329	83%
Support Vector Machine	0.804	0.79	0.7973	79%

V. Conclusion

Text based emotion AI has successfully been applied to the task of depression detection using Twitter data. The results delivered in this paper are at par with the previous results achieved in this domain. Supervised learning classification have a limitation and cannot grant a human level accuracy in prediction of depression through text data. Moreover there is significant noise in the Tweets collected before pre-processing,

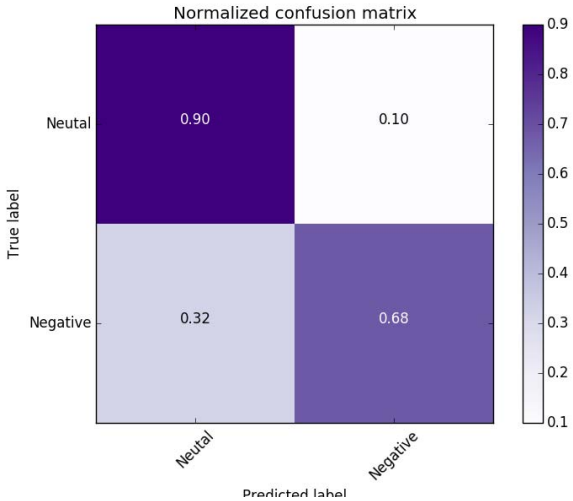

Fig. 4: Confusion Matrix for SVM

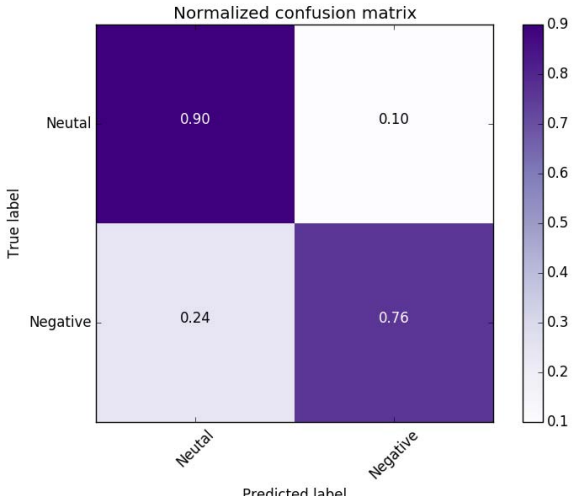

Fig. 5: Confusion Matrix for Naive Bayes

which eliminates about a third of the data due to third-person and news references. In future, a layer of expert-based suggestion can be added to the model to reduce number of false positives. This would increase the precision of sentiment analysis for depression detection.

References

[1] A. N Hasan, B. Twala, and T. Marwala,*Moving Towards Accurate Monitoring and Prediction of Gold Mine Underground Dam Levels*, IEEE IJCNN (WCCI) proceedings, Beijing, China, 2014.

[2] A.K.Jose, N.Bhatia, and S.Krishna, *Twitter Sentiment Analysis*,National Institute of Technology,Calicut, 2010.

[3] T.Mitchell, H.McGraw,*"Machine Learning,Second Edition, Chapter One"*,January 2010.

[4] C.D.Manning, P.Raghavan, H.Schutze,*Introduction to Information Retrieval*,Cambridge UP, 2008

[5] P.Taylor, Text-to-Speech Synthesis, Cambridge, U.K.:Cambridge University Press, 2009

[6] Y. Mejova, *Sentiment analysis: An overview*,http://www. cs. uiowa. edu/ymejova/publications/CompsYelenaMejova.pdf, 2009.

[7] M.S.Neethu, Rajsree, *"Sentiment analysis in twitter using machine learning techniques"*,Fourth International Conference on Computing, Communications and Networking Technologies (ICCCNT), 2013

[8] Benjamin L. Cook,Ana M. Progovac, Pei Chen, Brian Mullin, Sherry Hou and Enrique Baca-Garcia, *"Novel Use of Natural Language Processing (NLP) to Predict Suicidal Ideation and Psychiatric Symptoms in a Text-Based Mental Health Intervention in Madrid"*, Computational and Mathematical Methods in Medicine Volume,2016

[9] M.Rambocas, and J. Gama, *Marketing Research:The Role of Sentiment Analysis*, The 5th SNA-KDD Workshop11. University of Porto, 2013

[10] Stphane Meystre,Peter J.Haug, *"Natural language processing to extract medical problems from electronic clinical documents: Performance evaluation"*, Journal of Biomedical Informatics,Volume 39, Issue 6, December 2006

[11] Chowdhury, G., *"Natural language processing"*, Annual Review of Information Science and Technology, 2003

[12] Zi-qiang Wang,Xia Sun, De-xian Zhang, Xin Li, *"An Optimal SVM-Based Text Classification Algorithm"*, International Conference on MachineLearning and Cybernetics, 2006

[13] Mitali Desai, Mayuri A. Mehta,*"Techniques for sentiment analysis of Twitter data: A comprehensive survey"*, International Conference on Computing, Communication and Automation (ICCCA), 2016

[14] Krystian Horecki and Jacek Mazurkiewicz, *"Natural Language Processing Methods Used for Automatic Prediction Mechanism of Related Phenomenon"*, Springer International Publishing Switzerland ICAISC, 2015