ASSESSMENT OF THE PREVALENCE OF SUICIDE AMONG YOUNG ADULTS USING MACHINE LEARNING

IBODE, R.T., TUNDE-FRANCIS, A. A., AFEYE, A. F., ANIFOWOSE, O.T., OWOLOLA, O. I. AND OGIDAN, O. A.

ABSTRACT

Due to the high rate of suicide all over the world resulting in about 800,000 people dying by suicide each year. The instances where suicide victims constantly publish suicide messages deliberately to express their feelings on social media, there is need to address suicide issues, and how suicide can be prevented. Therefore, as a solution to this, there is need to create a model that classifies these users" social media posts and identify users with suicidal ideations, so as to prevent future suicide cases by getting the identified users the necessary help needed. The study adopted a binary classification of a suicide-related tweet with respect to age 15 up till 29 years, on a document-level basis. A machine learning approach was employed to solve the problem of tweet classification and predictions. The dataset was generated from a Twitter API.

It was observed that suicidal issues are rampant among the young adult, which need urgent attention. The paper recommended that timely intervention should be provided so as to reduce suicidal victims and preserve the future of young adults.

KEYWORDS: Suicide, Tweet, Machine-learning, Social-media.

INTRODUCTION

Machine Learning (ML), has been applied to various areas of research and the health sector is no exception. With machine learning, it is easier to analyse large datasets and make some inferences.

Suicide is a health problem that affects communities, families and the entire country at large. Suicidal rates have increased drastically over the years with an alarming rate among youths and adults. Every year close to 800,000 people take their own life and there are many more people who attempt suicide (WHO, 2019). Suicide occurs throughout the lifespan and was the second leading cause of death among young adults globally in 2016 (WHO, 2019). Being suicide does not just occur in high-income countries, but it's a global phenomenon in all regions of the world. In fact, over 79% of global suicides occurred in low- and middle-income countries in 2016 (WH0, 2019), Suicidal is a serious public health problem; however, suicides are preventable with timely, evidence- based and often low-cost interventions. For national responses to be effective, a comprehensive multi-sectorial suicide prevention strategy is needed (WHO, 2019).

Suicidal ideation detection in online social networks is an emerging research area with challenges. Recent research has shown that the publicly available information spread across social media platforms holds valuable indicators to effectively detecting persons with suicidal intentions. The key challenge of suicide prevention is to understand and detect the complex risk factors and warming signs that may precipitate the event (Mia Johnson et al, 2017). Identifying these risk factors is the first step in suicide prevention.

There are even instances of suicide victims writing their final thoughts on Twitter, Facebook, and other online communities (Vioules et al, 2017). (Aladag et al, 2016). Research has shown that suicide plans were more prevalent among young and its aged between 18 and 25 years (Aladag et al; 2016). Considering the high social media penetration rates of this group, many suicide attempts could possibly be prevented via social media surveillance, this is where this research aims to harness the use of ML to detect if an individual is having suicidal ideation.

Related Work

(Choudhury, 2013) emphasized the usefulness of identifying mental illness such as depressive disorder

Ibode, R.T.; Federal College of Forestry, Jericho, Ibadan, Oyo State, Department of Basic and General Studies

Tunde-Francis, A. A.; Federal College of Forestry, Jericho, Ibadan, Oyo State, Department of Basic and General Studies

Afeye, A. F.; Forestry Research Institute of Nigeria (FRIN) Jericho, Ibadan, Oyo State.

Anifowose, O. T.; Federal College of Forestry, Jericho, Ibadan, Oyo State, Department of Basic and General Studies

Owolola,O. I., Federal College of Forestry, Jericho, Ibadan, Oyo State, Department of Basic and General Studies

Ogidan, O. A.,

through social media posts and building tools that can be used by health care authorities and also by the user to take precautionary measures and to obtain the necessary treatments to avoid further impairment.

Moreover, they raise awareness of the ethical implications of probing social media content to identify indicators of mental health disorders with the intent of informing health care authorities or relevant parties such as family members. De Choudhury et al. (2013) noted that the use of language and emotion embedded within postings could indicate depression. e.g. postings that represent worthlessness, guilt, helplessness, and self-hatred can be considered as indicators of depression. Additionally, increased use of first-person singular pronouns in comparison to second and third person pronouns is also an indicator of depression. In contrast to linguistic features, withdrawal from social activities and changes within the social network relationships are also highlighted as possible indicators of depression. They also note that multimedia content can be used as an indicator. Similar to De Choudhury's work, (Reece et al, 2016) discovered a relation between Instagram filters and depressed users. They show that "inkwell" filter, which turns coloured photos to black and white, was the most commonly used filter among depressed participants. Jashinsky et al, (2014) filtered at-risk tweets from Twitter stream using keywords and phrases created from suicide risk factors, and then compared values for suicide tweeters against national data of actual suicide rates from the Centers for Disease Control and Prevention. They found a strong correlation between state Twitter-derived data and actual state age-adjusted data.

The North American Chapter of the Association for Computational Linguistics (NAACL) organized a Clinical Psycchology Shared Task2015 (CLPsych, 2015) organized by (Coppersmith et al, 2015a) to detect and diagnose mental health problems from Twitter data. The challenge involved; predicting depression users from the control group (PTSD) and differentiating between depression and PISD users, Johnson et al. (2017) presented a new approach to automatically identify sudden changes in user's online behaviour. They used research from the field of psychology, to design and develop behavioural features that quantify the level of risk for an individual according to his online behaviour on Twitter. They created a feature for text analysis called the Suicide Prevention Assistant (SPA) text score. The research focused on detecting distress- and suicide-related content and it developed two approaches to score a tweet: an NLP-based approach and a more traditional machine learning text classifier. The experiments showed that the NLP text-scoring approach successfully separated out tweets exhibiting distress-related content and acts as a powerful input into the martingale framework.

This work focuses on the use of Random Forest and Support Vector Machine (SVM) machine learning techniques to detect and identity of suicidal tweets among users within the age of 15 to 29 years.

METHODOLOGY

The model diagrams in Figure 1 and Figure 2 show how data flow from the different levels and processes involved in classifying and predicting individual user tweet. After data had been gathered it underwent cleaning and pre-processing, after which feature extraction were performed on the data to extract the needed features from the data. The resultant dataset was then normalized.

The goal of normalization is to change the values of numeric values in the dataset to use a common scale and then fed to the Random Forest and VM algorithms for both the age prediction and suicide detection models respectively, to classify and make suicide tendency prediction as output.
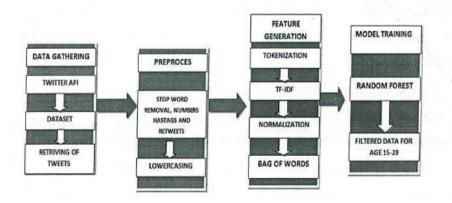

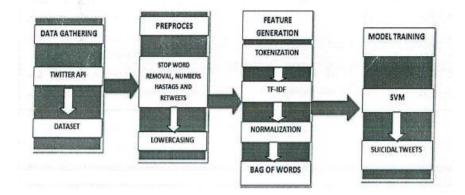**Figure 1: The age classification model diagram**

Figure 2: The suicide defection mode diagram

Datasets

The dataset was generated from a Twitter API. Birthday announcement tweets were collected from the Twitter Search API (https:/api.twitter.com/1.1/search/tweets json) using the search parameters such as "Happy nth Birthday, where n ranges 13 to 40'. This was used to capture self reported birthday tweets. Birthday tweets from the age ranges of under 15, 15 to 29 and above 29 were collected between April 22 and April 25, 2019. These tweets were used to capture usernames so as to acquire previous tweets of the users between the desired age group.

```
In [0]  : #create birthday query keywords
birthday = sorted (list(get_keywords_from_file('.. /data/birthday-keywords. txt')))
birthdays
Out [0] : [ "happy 13rd birthday to me",
"happy 14th birthday to me",
"happy 15th birthday to me",
"happy 16th birthday to me",
"happy 17th birthday to me",
"happy 18th birthday to me",
"happy 19th birthday to me",
"happy 20th birthday to me",
"happy 21st birthday to me",
"happy 22nd birthday to me",
"happy 23rd birthday to me",
"happy 24th birthday to me",
"happy 25th birthday to me",
"happy 26th birthday to me",
"happy 27th birthday to me",
"happy 28th birthday to me",
"happy 29th birthday to me",
"happy 30th birthday to me",
"happy 31st birthday to me",
"happy 32nd birthday to me",
"happy 33rd birthday to me",
"happy 34th birthday to me",
"happy 35th birthday to me",
"happy 36th birthday to me",
"happy 37th birthday to me",
"happy 38th birthday to me",
"happy 39th birthday to me",
"happy 40th birthday to me",
In [0] :    # divide birthdays into groups of <15, 15-29, >29
groups = birthday [ :2 ], birthdays [2:17], birthdays [17: ]
```

Figure 3: Data gathering for age classification

Figure 3 shows a snippet of code in python used to fetch tweets from twitter API using the keywords above for the age classification model. At the end of the data gathering stage a total of 2000 tweets were crawled.

The dataset was generated from a Twitter APL. Tweets containing depressive and suicidal words were crawled from the Twitter Search API (https:/api.twiter.com/1.1 /search/tweets.json) using the search parameters such as suicidal; suicide; kill myself; my suicide note; my suicide letter, and my life; never wake up; can't go on; not worth living: ready to jump; sleep forever, want to die; be dead; better off without me; better off dead; suicide plan; suicide pact; tired of living: don't want to be

here; die alone; go to sleep forever; wanna die; wanna suicide; commit suicide; die now; slit my wrist; cut my wrist; slash my wrist; do not want to be here; want it to be over, want to be dead; nothing to live tor; ready to

die; not worth living: why should I continue living; take my own life; thoughts of suicide; to take my own life; suicide ideation; depressed; I wish I were dead; kill me now to capture depressive and suicidal tweets.(O'Dea *et al*, 2015).

```
In [0] :   keywords = get_keywords_from_file (".. /data/suicide-keywords.txt")
Out [0] : { "be dead",
"better off dead",
"better off without me",
"can't go on",
"commit suicide",
"cut my wrist",
"depressed",
"die alone",
"die now",
"do not want to be here",
"don't want to be here",
"end my life",
"go to sleep forever",
"I wish I were dead",
"kill me now",
"kill myself",
"my suicide letter",
"my suicide note",
"never wake up",
"not worth living",
"nothing to live for",
"ready to die",
"ready to jump",
"slash my wrist",
"sleep forever",
"slit my wrist",
"suicidal",
"suicide",
"suicide ideation",
"suicide pact",
"suicide pact",
"suicide plan",
"take my own life",
"thoughts of suicide",
"tired of living",
```

Figure 4: Data gathering for suicide detection model
Figure shows a snippet of code in python used to fetch tweets from twitter API using the keywords for the suicide detection model.

Data Annotation
For this research our focus was on supervised learning techniques. Hence, we annotated the dataset. The age classification dataset was grouped into 3, group 1

contained tweets of users below 15, group 2 contained tweets of users between 15 and 29 and group 3 contained tweets of users above 29 years old.
The suicide detection dataset was labelled manually to help identify tweets that contain suicide ideation, suicidal tweets were labelled as 1's and non suicidal tweets were labelled as 0's respectively. Figure 6 shows a snippet of labelled tweets.

	tweet_text	label
0	Iâ€™m the real 77 Toast my psn name is slit wr...	0
1	@Allhailsierra @sierraawolf Slit my wrist and ...	0
2	Twinkle twinkle in the sky\nI just wanna fucki...	1
3	see im a waste of life i should just kill myse...	1
4	If one more coworker says "Ugh who left these ...	0

Figure 6: Dataset annotated into 2 groups

Data Pre-processing
After the data gathering stage, the datasets were then preprocessed by removing stop words, hash tags, retweets, and changing the entire dataset into lower cases.
- Stop-word removal: This involves removing frequently used non-informative words, e.g. a, an', "the, and "is, etc

- **Symbols**: such as (HTIML tags S, commas, full-stop and other punctuation symbols) were removed in this step.
- **Lower casing**: the entire dataset was reduced into lower case for simplicity, this helps in the consistency of our expected output.

```
In [13] : import n1tk
n1tk.download ("punkt")
n1tk.download ("stopwords")
[n1tk_data] Downloading package punkt to /root/n1tk_data...
[n1tk_data] Unzipping tokenizers/punkt.zip.
[n1tk_data] Downloading package stopwords to /root/n1tk_data...
[n1tk_data] Unzipping corpora/stopwords.zip.
Out [13] : True
```

Figure 7: Data pre-processing

Figure 7 shows a snippet of code in python used to import the natural language toolkit for natural language processing. It contains text pre-processing libraries for tokenization, classification, stemming etc. The punkt sentence tokenizer divides a text into a list of sentences, while the nltk.download ('stop words') removes stop words from the dataset.

Feature generation
In Machine Learning, feature generation is referred to as the process of using domain knowledge of the data to create features that can be used by machine learning algorithms to find a pattern (Techopedia, 2018). A feature is a piece of information that might be useful for making predictions.

Attribute, variable are some other terms used synonymously to features when defining a Machine Learning task. A dataset given to ML algorithm consists of dependent and independent variables.

The dependent variable is usually a nominal value i.e., the target outcome of a prediction. A variable is nominal when it has a fixed number of categories. Independent variables, on the other hand, can be of type numeric (such as integer or real) or nominal. Features can be provided as part of the dataset by an expert or derived

from data in case of text data. An appropriate dataset will then be generated after this phase.

Tokenization
Our dataset was tokenized by splitting paragraphs into sentences, and sentences into individual words. This relies on pre-trained, language specific algorithms like the punkt models from NLTK.

Normalization
Normalization and weighing was performed on the dataset, in this process normalization of the vectors (datasets cleaned rows and column) into a unit length which was extracted from the original text length of the unclean or raw datasets. The Count Vectorizer library from the Scikitlearns's package of the Python 3 programming language was used to normalize the dataset.

Bag of words
A bag of words model is used to extract features from the dataset for use in the model training phase. The bag of words was used to represent text describing the occurrence of words within a document. The Tf-Idf Transformer library from the Scikitlearns's package of the Python 3 programming language was used to create a bag of words model for the dataset.

```
In [0] : # convert tweets to bag-of-words per groups
# unigram to trigram
# do same for test data
from skeleton.features_extraction.text import CountVectorizer, TfidfTransformer
bow_transformer = CountVectorizer [tokenizers = None, max_features =32191.fit (create_corpus(training_data)]
tweet_box = bow_transformer.transform|create_corpus (training_data)
tf_idf_transformer = TfidfTransformer (). fit | tweet_box|
tweet_tf_idf = tf_idf_transformer.transform |tweet_box|
X_train = pd. DataFrame (tweet_tf_idf.toarray)
#vectorizer.get_stop_words ()
```

Figure 8: Bag of words model

Figure 8 shows a snippet of code in python used to import the count vectonizer, Tf-Idt transformer, BOW transformer from the Scikitlearns's feature extraction library.

Training and Test Split
The dataset was divided to keep 80% tweets for the training set and 20% for the test set. The split was done randomly, and the classifiers were called to feed on the training and testing data respectively. This was done using the train/test split function library available in the Python 3 programming language.

```
In [30] : print (X_train.shape)
print (X_test.shape)
print (Y_train.shape)
print (Y_test.shape)
(1680, 1 )
(420, 1 )
(1680, )
(420, )
```

Figure 9: Training and splitting set for the age classification model

Figure 9 shows the snippet of code in python splitting of the data into training and test data. With 80% of the dataset used for training 20% for testing, which resulted in 1680 tweets for training and 420 tweets for testing of the age classification model.

```
In [45] : print (X_train.shape)
print (X_test.shape)
print (Y_train.shape)
print (Y_test.shape)
(3035, 1)
(759, 1)
(3035, )
(759, )
```

Figure 10: Training and splitting set for the suicide detection model.

Figure 10 shows the snippet of code in python splitting of the data into training and test data. With 80% of the dataset used for training 20% for testing, which resulted in 3035 tweets for training and 759 tweets for testing of the suicide detection model.

RESULTS
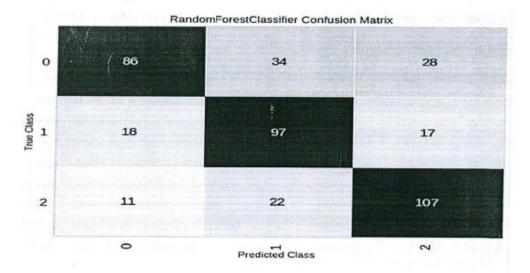
Confusion Matrix of Results Obtained

Figure11: Confusion Matrix for age classification model

Figure 11 is a confusion matrix which depicts the classification correctness of the model. It shows the predicted class and the true class and shows where a model classification lies for every input.

We have 3 age groups (under 15, 15-29, above 29) therefore a tweet in any of the 3 age groups should be detected by the model correctly classifying it in either one of the3 age groups, the under 15 age group is labelled with index 0,15-29 age group is labelled as index 1 and the above 29 age group is labelled as index 2.We tested 420 tweets,l15 tweets belong to the age group 0,158 tweets belong to the age group I and 152 tweets belong to the age group 2.

This confusion matrix image shows that for tweets in age group 0. 86 tweets were classified correctly out of 115 tweets, 97 tweets were classified correctly out of 158 tweets for tweets in the age group 1 and 107 tweets were classified correctly out of 152 tweets in the age group 2
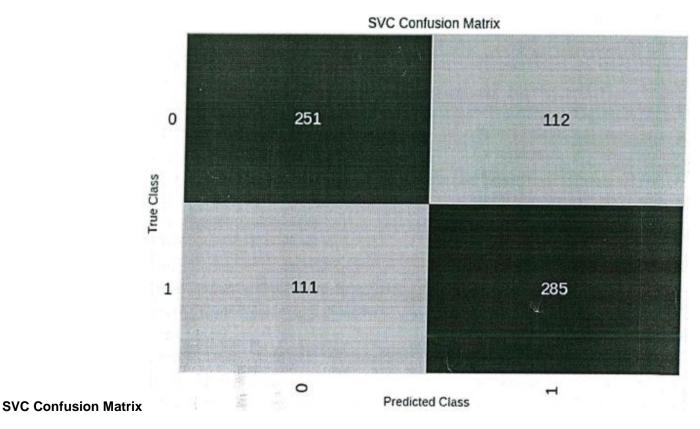

SVC Confusion Matrix

Figure12: Confusion Matrix for suicide detection model

Figure 12 is a confusion matrix which depicts the classification correctness of the suicide detection model. It shows the predicted class and the true class and shows where a model classification lies for every input. We have 2 classes of data in this dataset, suicidal tweets were labelled with the index 1, while non suicidal tweets were labelled with the index 0.The model was tested with 759

tweets with 362 tweets belonging to the class 0 and 397 tweets belonging to the class 1. This confusion matrix image shows that for tweets in class 0, 25l tweets were classified correctly out of 362 tweets, and for tweets in the class l, 285 tweets were classified correctly out of 397 tweets.

CONCLUSION

A classification accuracy of 67%, was achieved for the age classification model, and a classification accuracy of 70%, was achieved for the suicide detection model and the result of the analysis was presented using precision, recall and macro Fl-score with the proposed model showing significant performance.

From this research, it can be concluded that detection of suicidal tweets among young adults using machine learning algorithms yields significantly relevant results when classifying tweets of young adults as either suicidal or non suicidal.

RECOMMENDATION

The developed model performed quite well with 67% and a 70% accuracy on quite a small number of datasets. Major challenge was in the volume of data that could be gathered for this work. The model will perform excellently well if more volume of data can be gathered. It is therefore recommended that machine learning approaches should be used to predict suicidal tweets and timely intervention should be provided so as to reduce suicidal victims and preserve the future of young adults.

REFERENCES

Aladag, A.,Muderrisoglu, S., Akbas, N., Kahvecioglu, O, and Bingol, H., 2016. Fatal injury reports, national, regional and state. Centers for Disease Control and Prevention: WISQARS.

Choudhury, M., 2013. Role of social media in tackling challenges in mental health. Proceedings of the 2nd international workshop on Socially- aware multimedia.

Pirkis, J. and Blood, R., 2001. Suicide and the medía. Crisis: The Journal of Crisis Intervention and Suicide Prevention, 22(4), 155-162.

World Health organization, 2019. Global World Estimates

Johnson V., Moulahi B., Aze J, Bringay S., 2017. Detection of Suicide-Related post in Twitter Data Stream , IBM Journal of Research and Development 62 (1)

O'Dea B., Wan S., Batterham J.,CalearA., Paris C., Christensen HG., 2015. Detecting Suicidality on Twitter, Internet Intervention 2, 2015. 183-188

Jashinsky J., Burton S., Hanson C., Josh W., 2014. Tracking Suicide Risk Factors through Twitter in the US, Journal of Crisis Intervention and Suicide Prevention 35 (1): 1-9

Coppersmith G., Dredze M., Harman C., Hollingshead K., Mitchell M., 2015.CLPsych 2015 shared Task, depression and PTSD on Twitter