Detection of Depression among Nigerians using Machine Learning Techniques

MSc Research Project
Data Analytics

ADEDEJI ADEGOKE
x18174582

School of Computing
National College of Ireland

Supervisor: Dr Catherine Mulwa

National College of Ireland

MSc Project Submission Sheet

School of Computing

Student Name:	ADEDEJI ADEGOKE
Student ID:	X18174582
Programme:	MSc DATA ANALYTICS **Year:** 2019
Module:	RESEARCH PROJECT
Lecturer:	DR CATHERINE MULWA
Submission Due Date:	12TH of December 2019
Project Title:	DETECTION OF DEPRESSION AMONG NIGERIANS USING MACHINE LEARNING TECHNIQUES
Word Count:	8033 **Page Count:** 28

I hereby certify that the information contained in this (my submission) is information pertaining to research I conducted for this project. All information other than my own contribution will be fully referenced and listed in the relevant bibliography section at the rear of the project.
ALL internet material must be referenced in the bibliography section. Students are required to use the Referencing Standard specified in the report template. To use other author's written or electronic work is illegal (plagiarism) and may result in disciplinary action.

Signature: ………………………………………………………………………………………………………

Date: ………………………………………………………………………………………………………

PLEASE READ THE FOLLOWING INSTRUCTIONS AND CHECKLIST

Attach a completed copy of this sheet to each project (including multiple copies)	□
Attach a Moodle submission receipt of the online project submission, to each project (including multiple copies).	□
You must ensure that you retain a HARD COPY of the project, both for your own reference and in case a project is lost or mislaid. It is not sufficient to keep a copy on computer.	□

Assignments that are submitted to the Programme Coordinator Office must be placed into the assignment box located outside the office.

Office Use Only	
Signature:	
Date:	
Penalty Applied (if applicable):	

Detection of Depression among Nigerians using Machine Learning Techniques

ADEDEJI ADEGOKE

x18174582

Abstract

Depression has always been a cause for concern in Nigeria and a dreadful entity that contributes to suicide rate among the young and the old in the country. In other to reduce the suicide rate and observed the depressive state of Nigerians on Twitter, this research project was tailored towards help highly skilled psychologists in the Nigeria health sector to detect depression among Nigerians. In this research project, various features were used to train different machine learning models in order to detect depression among Nigerians. The best model was a regularized generalized linear model using term frequency inverse document frequency and normalized data with a precision of 0.89, recall of 0.91 and f-measure of 0.90 which will be used by highly skilled psychologist in this domain to identify and detect depressive Nigerians on twitter. From the literature, class imbalance of dataset was dealt with synthetic sample which diminished the recall of their models, but the gap was breached by introducing tweets with positive polarity from a different dataset because positivity in a text has no effect other than positivity. Classification method was used to classify the depressive and non-depressive tweets using various classifiers and regularized generalized linear model with term frequency inverse document frequency and normalized data performed brilliantly. This will help Nigerians a lot in other to curb sudden death and suicide rate among its citizen and reduce high death rate statistics due to depression according to the World Health Organization.

1 Introduction

Depression is defined as a psychiatric condition in a person's knowledge, emotional regulation or actions, which is the evolutionary mechanism of cognitive functioning, mental, biological or behavioural disorder (American Psychiatric Association, 2013). According to World Health Organization (WHO,2004), mental health is an important element of stability and the well-being of every person, depression may impact anybody of any age, personal background, affluent, poor, male or female. Practice in mental health is often described as complicated, particularly if it is associated with demeaning harm and misunderstandings because mental illness is difficult to diagnose. There is no credible laboratory test for most types of mental health, and the diagnosis is typically focused on the nature of the person and the mental stability review.

As far back as the early 1960s, violence in social groups, families and the society at large have been attributed to depression which affects an individual's psychological and emotional ability to effectively make rational decisions. Today, depression has gotten the attention of everyone as it is not seen today as an old people sickness or associated with people considered to be lower class citizens but in today's world finding, it's place is found most especially in young people and this modern-day epidemic is now seen as a global problem.

1.1 Background and Motivation

Depressive symptoms, decreased encouragement, sick-mindedness and psychological deterioration in individuals as often as possible were associated with the usage of social networking sites (Healthline, 2018). The liberty of humans in Nigeria, particularly teenagers to use social media has evolved substantially recently and the majority of the Nigerians are extremely intelligent with the innovation and make it an important way of interacting and expressing their feelings and emotions, the use of social media between Nigerians needs to be examined and monitored because it has a profound effect on almost every relational context and emotional development in their lives. Many scientists and analysts have used shared opinion of people on social media to understand these people's understanding, attitude, thoughts, perceptions of a concept and online concerns (Ashiekpe et.al, 2017).

Natural Language Processing is the human thought interaction research with a computer which practically enables the understanding of human thoughts or language, it is an arm of computer science which deals with artificial intelligence. Natural Language Processing involve processes such as text extraction, text summarization, topic segmentation, stemming and lemmatization, POS tagging, word corrections in which some of them were used in this research project. According to (MonkeyLearn, 2019), Aspect Based Sentiment Analysis is an analysis which classify thoughts identified in a text section to establish whether a writer's opinion on a subject is bad, good or neutral.

Machine Learning is an application of Artificial Intelligence, which offers machines with the ability to instantly learn about knowledge without express programming, to ensure a good understanding and to provide machine-learning techniques with certain distinctive features that can assist in examining and revealing the distinctive rhythm of social interaction.

1.2 Research Question

This research project aimed at solving depression issues among Nigerians using their twitter data in order to classify their depressive interaction using machine learning techniques to detect tweets that are depressive and non-depressive which will help highly skilled psychologist in Nigerian Health Sector to know the set of Nigerian users that are prone to depression based on their tweets and emotional expression on twitter. Classification model was adopted based on literature review because it has worked effectively in classifying and detecting depressive traits which was used to answer the research question:

RQ ; *"To what extent can detection models developed using machine learning techniques (linear discriminant analysis, classification and regression tree, random forest, extreme gradient boosting, adaptive boosting, regularized generalized linear model and C5.0 Decision Tree) enhance and help highly skilled psychologist in Nigeria improve their health sector in order to classify depression (anxiety, suicidal thoughts, suicide, anger, mental health, rage, mental disorder) among Nigerians using Nigerians Twitter data?"*

Depression is a significant state of mind that affect the mental ability of a human being and must be monitored properly in order to save lives. This information and communication technology solution provided by this research fulfils a niche for skill personnel in Nigerian health sector to tackle the issue of depression among its citizens.

To solve the research question, the following objectives were implemented as presented in subsection 1.3.

1.3 Research Objectives and Contributions

Objective one involved a critical literature review on detection of depression on social media was done with time scope (2007 – 2019) in order to understand and extract necessary insights and information which was used to achieve this research project and based on the research question, the following key objectives were set to solve the research questions as presented in the Table 1.

Table 1. Research Project Objectives

S/N	Objectives	Evaluation Metrics
Objective 1	Dataset creation using twitter API and features extraction process	
Objective 1.1	Extraction of data (tweets) from twitter using twitter API and using location filtering process to get Nigerian user tweets on depression.	
Objective 1.2	Extraction of features (n-grams, sentiment-polarity, hash) using term frequency-inverse document frequency (TF-IDF) and tokenization on sparse and tokenized data.	
Objective 2	Implementation of depression detection models using sparse data and tokenized data	
Objective 2.1	Implementation, evaluation and results of linear discriminant	

	analysis (LDA) using sparse data	
Objective 2.2	Implementation, evaluation and results of classification and regression tree (CART) using sparse data	
Objective 2.3	Implementation, evaluation and results of extreme gradient boosting (XGBoost) using sparse data	
Objective 2.4	Implementation, evaluation and results of C5.0 decision tree using sparse data	Precision, Recall and F-Measure
Objective 2.5	Implementation, evaluation and results of adaptive boosting (AdaBoost) using sparse data	
Objective 2.6	Implementation, evaluation and results of random forest using sparse data	
Objective 2.7	Implementation, evaluation and results of regularized generalized linear model (GLMNET) using sparse data	
Objective 2.8	Implementation, evaluation and results of regularized generalized linear model (GLMNET) using tokenized data and normalized data	
Objective 3	Comparison of the developed algorithms (objective 2)	
Objective 4	Comparison of the developed algorithms (objective 3) with the existing ones in the literature.	

Major Contribution: The major contribution resulting from this research project is the depression detection models with effective evaluation metrics for supporting highly skilled psychologists in Nigerian health sector to detect depressive tweets among Nigerians.

Minor Contributions

- A new dataset for Nigerian depression was created using Twitter Application Programming Interface.
- The results of the explanatory data analysis using word cloud
- The results of the reviewed literature on detection of depression and identified gaps from the literature.

The rest of the research project are as follow: Chapter 2 shows the critical review of the literature and past research works on this research topic, Chapter 3 shows the depression methodology approach, Chapter 4 shows the implementation process of the research, evaluation and results, Chapter 5 shows the conclusion and recommended future work.

2. Literature Review on Depression Detection (2007 – 2019)

2.1 Introduction

According to world health organization Statistic (WHO, 2004), it is construed that depression is a growing factor among individuals and to avoid loss of life, keen attention must be paid to depression. In this research work section, various research relating to classification of depression were examined. Analysis were based on 3 keen areas ranging from the background of the problem solved with the research project, how researchers did their data collection, data pre-processing, feature extraction and identifying effective data mining techniques towards classification of depression.

2.2 A Critical Review on Depression as a problem in Nigeria

Depression is a significant state of mind that affect the mental ability of a human being and must be monitored properly in order to save lives. The liberty of humans in Nigeria, particularly teenagers to use social media has evolved substantially recently and the majority of the Nigerians are extremely intelligent with the innovation and make it an important way of interacting and expressing their feelings and emotions, the use of social media between Nigerians needs to be examined and monitored because it has a profound effect on almost every relational context and emotional development in their lives. Many scientists and analysts have used shared opinion of people on social media to understand these people’s understanding, attitude, thoughts, perceptions of a concept and online concerns (Ashiekpe et.al, 2017).

The most recent statistics published by the world health organization (WHO, 2019), show that Nigeria has 7,079,815 victims of depression which indicates 3.9 percentage of the country’s population while 4,894,557 Nigerians (2.7 percentage of the population), experiencing anxiety disorders. Suicide rate over the years in Nigeria has severely risen, which can be tracked to mental illness mainly depression or post-traumatic stress disorder which are mostly gone unnoticed. A review of suicide rates in Nigeria reveals that somewhere between April 8, 2017, and May 12, 2018, there have been 79 recorded suicide incidents with a WHO study claiming Nigeria as the world's 30th highly susceptible nation to suicide.

Nigeria's high ranking was largely attributed to the scarcity of psychologists and psychiatrists, a lack of clinical mental health services, the expense of seeking medical care for mental health and to perceptions of being evaluated and the severity of mental illnesses, and to the apprehension of putting trust in an outsider . These are some of the main reasons why people refuse to seek medical care regarding their depressive situation which is really a big issue that needs to be tackle hands down with computing innovations currently available.

2.3 A Review of Methodologies used in Data Collection

To justify performing the use of classification in terms of detecting depression on social media, a state-of-the-art review was carried out identifying, comparing and critiquing how other researchers collected their dataset for the purpose of their research. Tsugawa et.al (2015), collected data from twitter on depression level by using a website to provide a questionnaire and spread information over twitter about the website, the answers to the questionnaire were collected and utilized to evaluate the level of depression of twitter users who partook in the procedure. twitter posts were collected using hashtags relating to depression, thus the dataset was extracted utilizing twitter application programming interface and an advanced symbolic team was incorporated to compute the data according to Jamil (2007).s Park et.al (2012), utilized another method of data collection where the twitter application programming interface was utilized to find depression related tweets with keyword "depression" and data about the keyword was generated over a period of two months from June to July 2019. During Park et.al (2012) process of data collection, it could be understood that their data sample in comparison to the English written and gathered tweets were limited from the united states of America and the geographical location of the tweets were identified using the twitter geo-location application programming interface.

Self-report survey which is another method for data collection utilizes interviews in epidemiological and psychological in its research which constitutes the gathering of data by recording responses on occasions of depression occurrence and also outcomes on the Beck's Depression Inventory and Revised Depression Scale Centre of Epidemiologic Studies Several categories of depression were reviewed in this study ranging from negative emotion, more self-focus, diurnal cycles, less social interaction and talking about key terms relating to depression throughout the time of study starting prior to the beginning of the depression. Recce et.al. (2016) identified post-traumatic stress disorder (PTSD) and depression by collecting the data for the research by getting various individuals to participate by utilizing the individual's data as well as Mechanical Turk Crowd work platform by running a survey on the latter and making available to various individuals the twitter history which were pointed out to be in connection to depression and PTSD data sample. Another method by which Recce et.al (2016) collected data was by utilizing a basic sample survey questionnaire where participating individuals were required to answer series of questions on the questionnaire making available answers regarding to the history of their mental health, demographic history and making available their social media data. and Revised Depression Scale Centre of Epidemiologic Studies was implemented to sieve out depression level of individuals participating in the exercise and qualified individuals participating in the exercise were prompted to divulge their tweet history and usernames, and in logging into their twitter account, a secure application was utilized after a predominant agreement to divulge their data. Furthermore, 279,951 tweets that concern depression were gathered from 204 twitter users and 243,755 tweets were gathered from 174 users to carry out PTSD analysis. De Choudhry et.al (2013), in carrying out prediction and analysis of salient depression disorder employed the process of collecting data and also its presence in individual's posts on social media platforms by utilizing crowd-sourcing and its effective manner of pointing out behavioral pattern in a dataset across various population of individuals. Like in the case of Recce et.al. (2016), survey was carried out utilizing screening and questionnaire to further test for depression among participating individuals was found using Beck's Depression Inventory and the consent for the utilization of twitter accounts and feeds was given by participating individuals. Between September 15 – October 31, 2012, a total 1,583 participating individuals filled and completed the questionnaire and a consent of agreement to access their account was provided by 637 participating individuals while 476 participating individuals ranging from 243 males and 233 females were with depression.

Prieto et.al (2014), acquired tweet data for their research work using twitter search application programming interface (API) with geocoding information contained in the tweet data metadata by sorting and filtering the data to obtain only tweets that originates from Spain and Portugal. The tweets were acquired for 30 days (from October 30, 2012 to November 30, 2012) which produced tweet data approximately 5.8 million and 4.5 million tweets in Spanish and Portuguese respectively.

In conclusion, survey responses and questionnaire responses by participants provide reliable data because of consent given by participant to researchers to access their accounts and feeds on twitter for proper selection of data for analysis and predictive modelling. However, the cost required for the method is lucrative and time consuming unlike the use of publicly accessible application programming interface (API) assessment criteria using keywords and hashtags relating to depression according to (Park et.al, 2012) and (Jamil, 2007)respectively, which is a self-declared method by user which was adopted for the purpose of this research considering the 3 months' time scope of actualizing the research work.

2.4 A Critical Review on Data Pre-Processing, Feature Selection and Data Mining Techniques

Data pre-processing is an important aspect of data analytics and is necessary to be performed on any form of data in order to achieve proper cleaning and transformation of the data as well as generating relevant features to build a model. Pre-processing of text data takes different turn unlike other structured data and in this case, pre-processing of a tweet data took a route of text pre-processing. According to Pak et.al (2010), in order to extract relevant features such as sentiments, n-gram as binary features from a twitter data which involves processes such as filtering, tokenization, removal of stop words, construction of n-gram, text conversion to lower case, removal of punctuation, elimination of whitespace, removal of URLs, stemming and lemmatization and these processes were adopted (Naf'tan et.al, 2019), (Karanasou et.al, 2016), (Islam et.al, 2018) and (Pietro et.al, 2015). Researchers have done many analyses and implemented many models and algorithm in order to detect depression in individuals posts on social media and in course of achieving this, proper literature review of related works show that classification model and its several algorithms is the most applicable model that perform effectively in the process of detecting depression on social media. According to Tsugawa et.al (2015), support vector machine (SVM) has proved a great deal in the process of detecting depression on social media when classifying with features such as sentiments, topics using linguistic word count (LIWC) and has proved relevant with 69% accuracy. Hassan et.al (2017) modelled SVM with features such as n-gram, n-gram negation and sentiment using Part of Speech Tagging with an accuracy of 91% likewise Naïve Bayes proved a great deal in depression detection which was adopted by Nadeem (2016) using n-gram as feature for the classifier with an AUC of 0.70. Benton et.al (2017), classified tweet data using n-gram for a Neural network in detecting depression, bipolar and Suicide with an AUC of 0.76, 0.75 and 0.83 respectively. Data were scrapped on SINA microblog in other to detect depression among active users of the microblog, Wang et.al (2013) used sentiment as feature using linguistic rules and vocabulary construction on a naïve bayes, Tree j48 and rules decision and evaluating the outcome with f-measure of 0.85, 0.76 and 0.81 respectively. logistic regression prove effective using n-grams, sentiment and topics as its features and according to Pietro et.al (2015), AUC of 0.91 and 0.85 were achieved on PTSD and depression, and term frequency- inverse document frequency was adopted by Naf'tan et.al (2019) in order to extract n-gram as a feature for detecting cyberbullying on instagram comments with naïve bayes classifying with 84% accuracy, Sau et.al (2018), used random forest with recursive feature elimination with an accuracy of 81.2%. Islam et.al (2018), Karanasou et.al (2016) both adopted support vector machine with sentiment as the feature being extracted from facebook and twitter data respectively which produced a recall of 0.80 and 0.85 respectively. In the same context, Islam et.al (2018) tried other classification algorithm with the same feature where k-nearest neighbor (KNN) produce recall of 0.59 and decision tree performed better with a recall of 0.98, Logistic Regression with sentiment as its feature using LIWC for detection of post-partum depression produced pseudo R^2 of 0.36 according to De Choudhry et.al (2014).

In context of the reviewed literature above from various research work on detection of depression by researchers, it is evident that support vector machine according to De Choudhry (2013), Islam et.al (2018), Karanasou et.al (2016), Hassan et.al (2017) provide effective results with sentiment as its feature using LIWC. Random Forest proved its effectiveness according to Sau et.al (2018), Recce et.al (2016), logistic regression did a great deal according to Pietro et.al (2015), De Choudhry et.al (2014). decision tree outperformed many algorithms according to Islam et.al (2018) with a recall of 0.98 which bring initiative toward a c5.0 decision tree as form of novelty to explore its predictive power as a classifier. From the review, evaluation method widely adopted on depression detection is based on precision, recall and f-measure according to Wang et.al (2013) and Tsugawa et.al (2015) which was adopted in this research.

2.5. Comparison and Conclusion

A comparative study was conducted to compare different sample size, criteria, features, algorithms, evaluation metrics, results and platform from different authors from the literature. The results of this comparison are presented in Table 2.

Table 2. Reviewed literature based on Sample size, Criteria Features, Algorithms, Evaluation metrics, Results and Platforms.

Sample Size	Criteria	Features	Algorithm with best performance	Evaluation Metrics	Results Best Performance	Platform	Authors and Year
476	Survey (CESD + BDI)	LIWC and Sentiment	SVM with PCA	Recall	0.72	Twitter	De Choudhry et.al. (2013)
9611	Self-declared with tweet	n-grams	Neural Network	AUC	Depression 76% Bipolar 75% Suicide 83%	Twitter	Benton et.al. (2017)
900	Self-declared with tweet	n-grams	Naïve Bayes	AUC	70%	Twitter	Nadeem (2016)
165	Survey (PHQ-9)	LIWC and Sentiment	Logistic Regression	Pseudo R2	0.36	Facebook	De Choudhry et.al. (2013)
209	Survey (CESD)	n-grams, LIWC, Sentiment and Topics	SVM	Precision, Recall and F-measure	0.69	Twitter	Tsugawa et.al. (2015)
6013	Self-declared	Sentiment, Linguistic rules, vocabulary construction	Bayes, Tree J48 and Rules Decision	F-measure	Bayes 0.85 Tree J48 0.76 Rules Decision 0.81	SINA Microblog	Wang et.al. (2013)
28,749	Survey (Personality)	n-grams, LIWC and Topics	Ridge Regression	Correlation	0.38	Facebook	Schwartz et.al. (2014)
378	Survey (CESD)	LIWC, Sentiment	Random Forest	AUC	Depression 87% PTSD 89%	Twitter	Recce et.al (2016)
1957	Self-declared with tweet	n-grams, LIWC, Sentiment and Topics	Logistic Regression	AUC	Depression 85% PTSD 91%	Twitter	Pietro et.al. (2015)
21,103 tweets 69 volunteers	Self-declared with tweet and Survey (CESD)	LIWC and Sentiment	Multiple Linear Regression	p-value	0.0022 < 0.05	Twitter associated with Geo-Location API	Park et.al. (2012)

It is evident from the literature that depression detection is a broad research and different technique and feature are used to achieve the aim. The researchers focused on detecting depression among social media users and only Prieto et.al (2014) did a confined research towards Spain and Portugal. Based the reviewed literature, knowledge are gathered including the challenges of creating dataset from the scratch and proper exploratory data analysis, this research took a supervised classification method (linear discriminant analysis, classification and regression tree, random forest, extreme gradient boosting, adaptive boosting, generalized linear model, principal component analysis and c5.0 decision tree) to detect depression among Nigerians on twitter using the extracted dataset where precision, recall and f-measure was adopted for evaluation.

3. Scientific Methodology and Design Specification

This chapter explains the procedure and method embraced in carrying out this research project as well as the architectural design and the purpose of the research was to develop classification method that can be used to detect depression post among Nigerians on twitter. This research project adopted the depression methodology which was altered to fit the process and purpose of this research and likewise the design specification of this research project was a 3-tier architectural design which was embraced because the implementation of the project requires a data persistent tier for creation of dataset through application programming interface.

3.1 Depression Methodology Approach

According to Berry et.al (1997), Data Mining is a technique which uses data that can provide good and greater insight to support decision making in order to acquire important ideas. Many methodologies are adopted by researchers for the purpose of data mining (Azevedo et.al, 2008) such as knowledge discovery in databases (KDD), cross-industry standard process for data mining (CRISP-DM) and in this research project, a knowledge discovery in databases was modified to suit several process of actualizing the depression detection solutions and the modified and altered knowledge discovery in databases (KDD) methodology is presented in the figure 1.

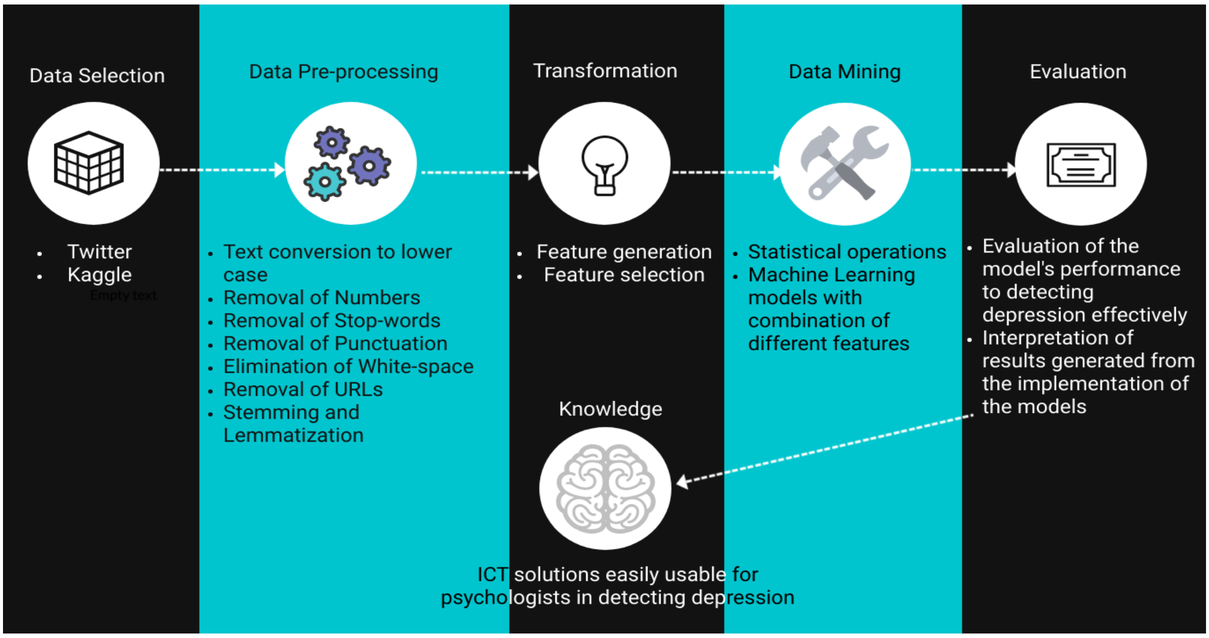


Figure 1: Depression Methodology

Data Selection: This process is important because it is the most important step of the research because the data controls the whole research project. There are 2 sources of data used for this research project namely:

- **Twitter dataset:** Twitter was utilized as a social media platform for extraction of individual's opinion and thoughts on subjects using various keywords and hashtag to extract those opinion and thoughts according to (Park et.al, 2012), (Jamil, 2007) and (Prieto et.al, 2014). Extraction of depressive tweets was performed using several keywords and hashtags such as (depressive, anger, anxiety, suicidal thoughts, suicide, rage, mental health, mental disorder, mental illness) using twitter application programming interface granted for this research work by twitter.

- **Kaggle dataset:** An available dataset was sourced and downloaded from kaggle called sentiment 140 (Bhayani et.al, 2009) which consists of various random tweets from twitter user and the sentiment polarity of each tweets which are annotated as Negative = 0 and Positive = 4 and Neutral = 2 and was selected in order to complement the created dataset from Twitter.

Data Pre-processing: The second stage was mainly about the filtering, cleaning and merging of the created dataset (Twitter dataset)and the dataset that was available in kaggle repository (Kaggle dataset) where the created dataset (Twitter dataset) was pre-processed using python3 on google collaboratory by undergoing processes such as text conversion to lower case, removal of numbers, removal of stop words, removal of punctuation, removal of whitespace, removal of special characters, calculation of the text's sentiment polarity score and stemming while the available dataset (Kaggle dataset) was pre-processed using R-studio and both datasets were merged using R-studio.

Transformation and Feature extraction: This process involves generating useful features from the pre-processed dataset which was used to train and test the classifiers for detecting depression and sentiment polarity, tf-idf, n-grams, hash and sparse data were used.

Data Mining: Statistical operations and various machine learning techniques using classifiers such as linear discriminant analysis, classification and regression tree, random forest, extreme gradient boosting, adaptive boosting, generalized linear model, principal component analysis and c5.0 decision tree were used with different features extracted such as sentiment polarity, text to vector, n-grams and word embedding were trained to detect depression.

Evaluation: Effectiveness and performance of each of these classifiers with different features is evaluated using precision, recall and f-measure.

3.2 Project Design Specification

Research work that deals with back-end processes should adopt a 3-tier design according to (Praveen, 2009) and in this research work, twitter application programming interface (API) was used for the extraction of tweets which was a back-end process. The design structure shows the full implementation of the project as presented in figure 2.

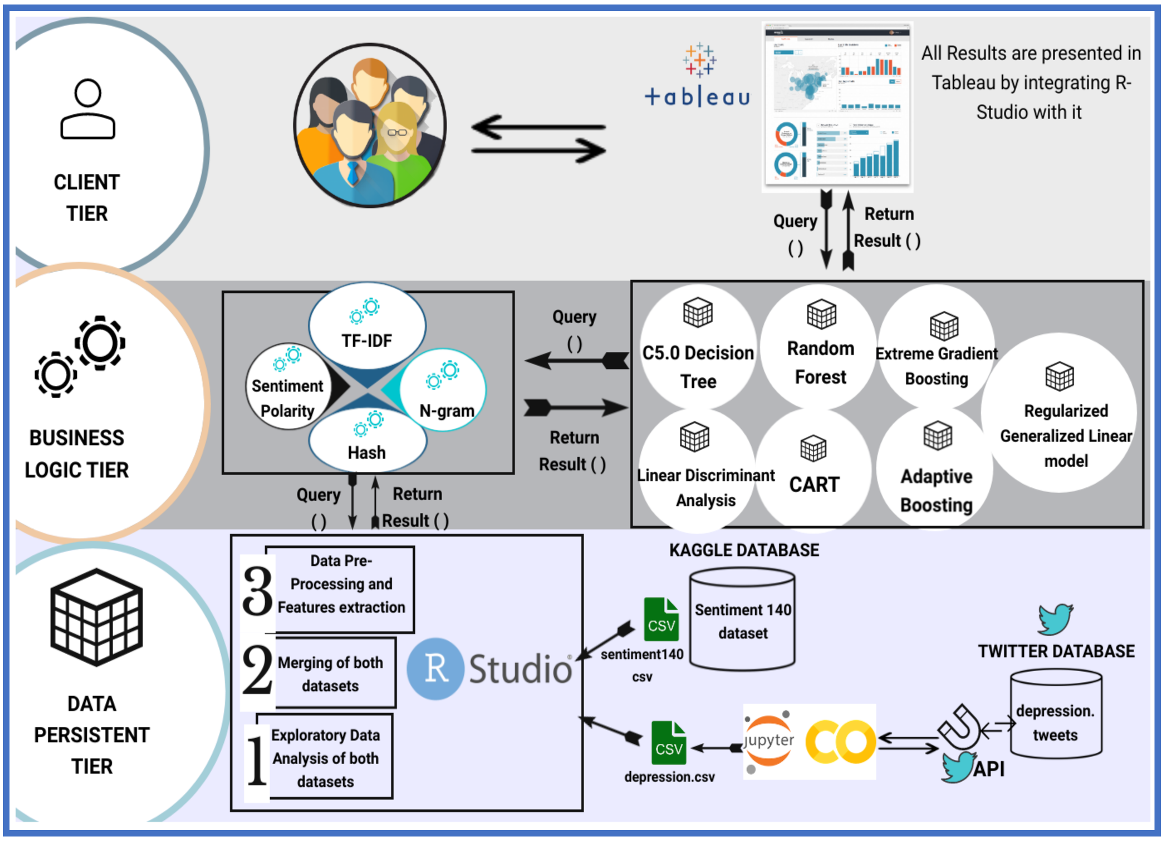


Figure 2: Design Specification of the research project

The first tier shows the visualization of results to highly skilled psychologist in the Nigerian health sector using Tableau.

The second tier shows the features extracted from the text through necessary pre-processing steps and feature extraction methods and how they were used to train the classification models.

The third tier shows the back-end process of how the twitter dataset was created and how the sentiment 140 dataset was downloaded and merged likewise how both datasets were pre-processed.

In this research project, depression detection methodology which was adopted from KDD was segmented into a 3-tier architectural design approach which includes the data persistent tier for the back-end processes, business logic tier for modelling and client tier for visualization of results. This methodology was used as a step by step guide during the implementation of the depression detection models.

4. Implementation, Evaluation and Result of Depression Detection Models

4.1 Introduction

The process of extracting tweets ethically and the way the tweets were made useful to provide context to the research questions are described in this section. The next two subsections show the processes involved in extracting the tweets as well as initial selection of usable features. The fourth subsection explores the datasets employed in this research while describing how the data were pre-processed and transformed. The rest of the subsections describes how the results were derived from the sampling to the comparison of the models first, within the research and later with other existing research in literature. In this research implementation features extracted from the text for the development of the models are:

- **Sentiment polarity**: Sentiment polarity score of each tweet were calculated using library (TextBlob) in python in order to know the polarity of each tweets (positive tweets, negative tweets).

- **TF-IDF:** Each tweets (text) were converted to vector by term frequency inverse document frequency (TF-IDF) using library (text2vec) in R-studio.

- **N-grams:** N-grams were extracted which consists of word n-grams by term frequency inverse document frequency (TF-IDF) using tokenization library (text2vec) in R-studio.

- **Hash:** Hash vectorizer was created with a hash size of 2^14 and a bigram. This was used to create the DTM used and was created using library (text2vec) in R studio.

- **Sparse Data:** Words that appeared only 0.5% of the time were removed from the document term matrix using library (tm) in R studio.

In order to evaluate the results of the implementation, evaluation metrics were adopted according to the literature.

Precision = True positive(TP) / True Positive(TP) + False Positive(FP)
where TP is the tweets that are positive to depression and predicted to be positive
and FP is the tweets that are negative to depression but predicted to be positive.

This means the ability of developed depression detection model to identify only the relevant data points i.e. it expresses the proportion of the data points in the tweets that our model says was depressive and were truly depressive.

Recall = True positive(TP) / True Positive(TP) + False Negative(FN)
where TP is the tweets that are positive to depression and predicted to be positive

and FN is the tweets that are positive to depression but predicted to be negative.
This means the ability of the developed depression detection model to find all the detected depression cases within the tweets.

F-measure = 2(Precision*Recall / Precision + Recall)

This is the single metric that combines recall and precision using the harmonic mean (Islam et.al, 2018).

4.2 Data Extraction

The process of extracting the tweets ethically was conducted which involves steps which are important in order not to violate twitter user's privacy (Twitter, 2018)[1]. The following steps were put into practice:

- A twitter developer's account was obtained, and twitter API application was developed for the purpose of the research.
- Necessary authentication keys in pulling tweets from twitter were obtained ethically through the twitter API application.

The depressive tweets were pulled from twitter database using hashtags and keywords according to (Park et.al, 2012), (Jamil, 2007) and (Prieto et.al, 2014) such as (depressive, anger, anxiety, suicidal thoughts, suicide, rage, mental health, mental disorder, mental illness) relating to depression. The tweets were extracted using python on google colaboratory and pre-processing was done in order to generate a clean text using library (tweet pre-processor) and calculating the polarity and subjectivity of the tweets in context to depression using library(textblob).

4.3 Feature selection

After the extraction of the tweets, the dataset derived had 18 attributes (columns) and 44179 instances (tweets). All of these were not considered necessary for the research and they were removed from the dataset. Overall, 6 attributes were considered usable for both modelling and exploration where polarity column (polarity score of the tweet) was considered as the dependent variable for the research. Figure 3 shows the initial list of attributes and the features selected. They were removed because they would not contribute any importance to the objectives of this research.

---

[1] Twitter (2018), Twitter privacy policy. Available at: https://cdn.cmstwdigitalassets.com/content/dam/legal-twitter/siteassets/privacy-pagegdpr/pdfs/PPQ22018AprilEN.pdf

```
 [1] "X"                "id"                    "created_at"       "source"
 [5] "original_text"    "clean_text"            "sentiment"        "polarity"
 [9] "subjectivity"     "lang"                  "favorite_count"   "retweet_count"
[13] "original_author"  "possibly_sensitive"    "hashtags"         "user_mentions"
[17] "place"            "place_coord_boundaries"
```

```
[1] "X"          "id"          "created_at" "clean_text" "polarity"   "place"
```

Figure 3: Feature selection of the important attributes

4.4 Exploratory Data Analysis, Data Pre-processing and Data transformation

Initial exploration of data included checking for missing values, duplicated tweets and the structure of the dataset and the following were observed:

- No missing values were found.
- No duplicated tweets were found
- The structures of the features selected were not accurate.

In a bid to understand the dataset, 4 questions were posed and to answer each of the questions, the data were processed and transformed. The questions are necessary to provide a context for the research. They include

- EDA1 - Where are the tweets coming from?
- EDA2 - What is the probability of positiveness of Nigerians' tweet?
- EDA3 - What are the common words in a depressive tweet from Nigeria?
- EDA4 - What are the common words in a positive tweet from Nigeria?

To answer EDA1 as presented in Figure 4, the necessary attributes were the "place" and "x" (id) columns. In the plotting of the graph, the location column was coerced into a character format and the strings were split to pick only the countries rather than the state and country as derived from the extraction of tweets. From the bar graph, there are 15234 tweets coming from Nigeria which is about 80% of the total size of the dataset. The next explorations were solely based on Nigerian's tweets.

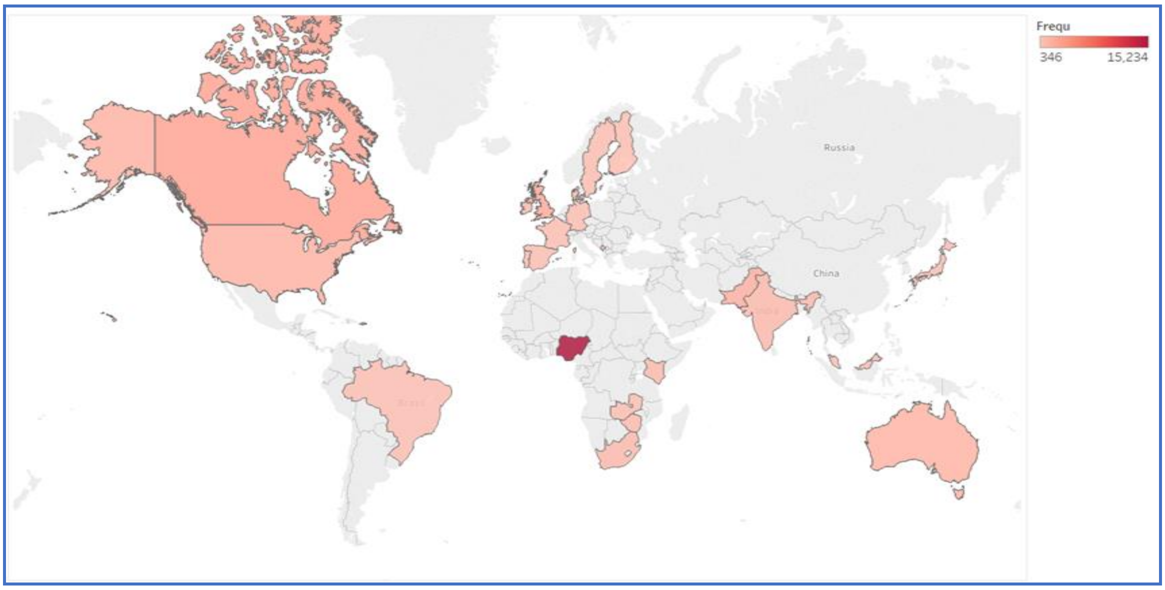


Figure 4: Map showing the origin of the tweets (Countries) and their frequency

Figure 5 presents the probability of the tweets being depressive by measuring the tweet sentiment polarity rate. From the visualization using a polarity scale of 0 to 1 (0 indicates depressive state and 1 indicates positivity). Therefore, the dataset has lots of depressive tweets. Figure 6 presents the collapsed class of depressiveness of a tweet in order to investigate the class imbalance. Out of the 15234 Nigerian tweets, 2141 were positive and 13,093 were depressive.

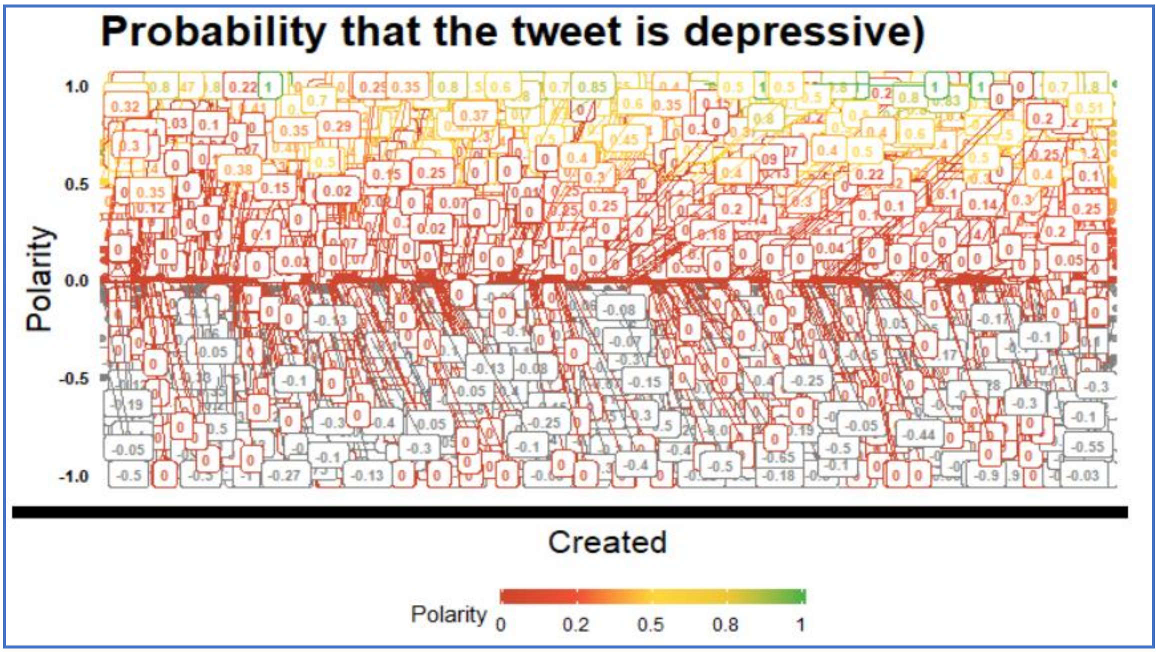


Figure 5: Tweets sentiment polarity score (Probability of Positivity and Negativity)

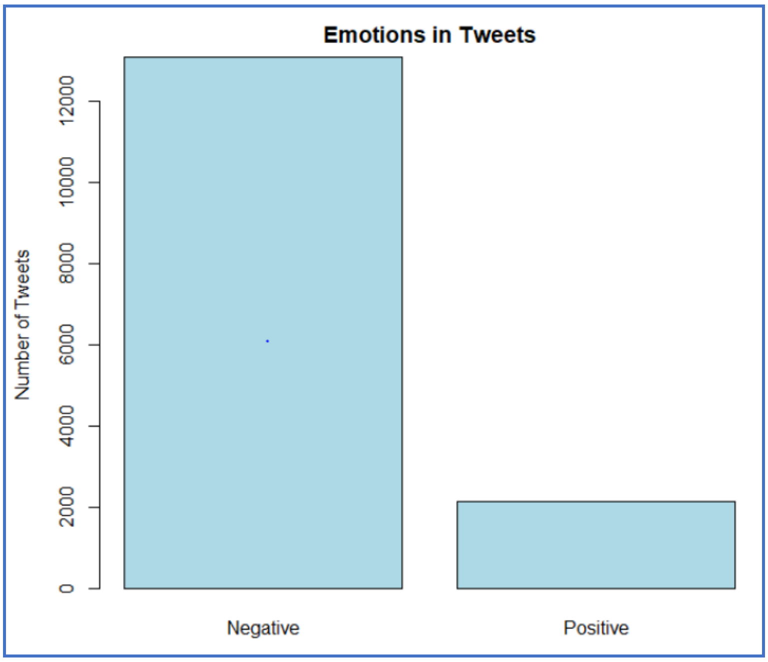


Figure 6: The frequency of Negative and Positive depressive tweet showing class imbalance

In answering EDA3, Figure 7 presents a word cloud of keywords in all the tweets and bar graph showing the top 10 words found across all the tweets after the removal of the hashtags initially used to harvest the tweets. Other pre-processing steps to achieve this included the converting the tweets into a corpus, removal of whitespaces, stopwords, punctuations, numbers, special characters, twitter characters and the conversion of text to lower case. It appears that depressive keywords can also include suicide, help, boys etc.

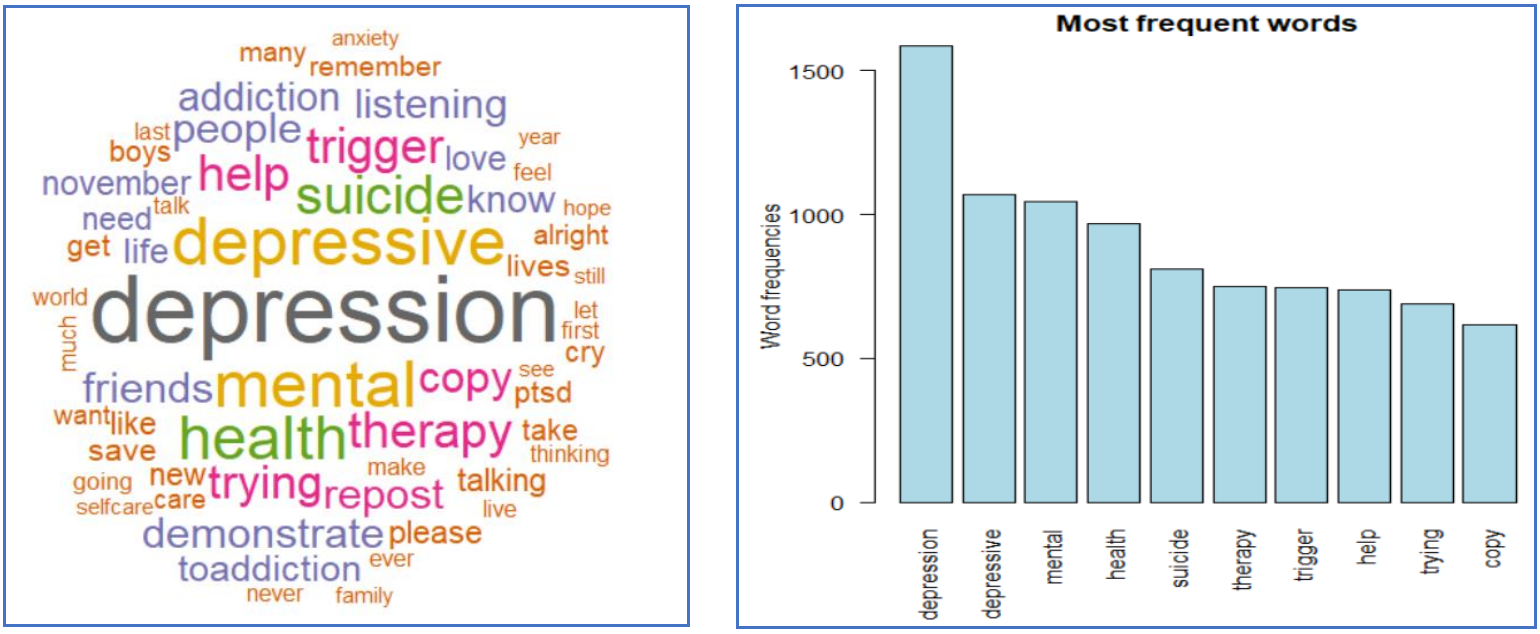


Figure 7: Word cloud and Bar-plot showing the frequency of the common words in the tweet

To investigate EDA4, a word cloud was also employed with similar pre-processing technique. However, to solve for the class imbalance observed in EDA2, a new dataset was introduced. The reasons are as follows:

- Using balancing methods like Synthetic Minority Over-sampling Technique (SMOTE) will generate new synthesized examples and this can reduce the intrinsic accuracy of the classification. SMOTE or any other sampling technique cannot create new tweets.
- Using balancing methods in this research is not practical because the data is high dimensional. In feature extraction, the data are tokenized leading to very high dimensionality.
- Positive tweets are the same regardless of the hashtags used to extract them. Also, the relevant data point in this research are the depressive tweets.

In view of the above reasons, positive tweets from a popular sentiment analysis dataset called sentiment140 were extracted to balance the dataset for the creation of models. 10952 positive tweets were extracted from this dataset and added to the other positive tweets in the original dataset as a form of transformation. After the addition, the common words for positive tweets were investigated using word cloud. Figure 8 presents the visualizations. It appears that words such as good, love, great expresses positiveness.

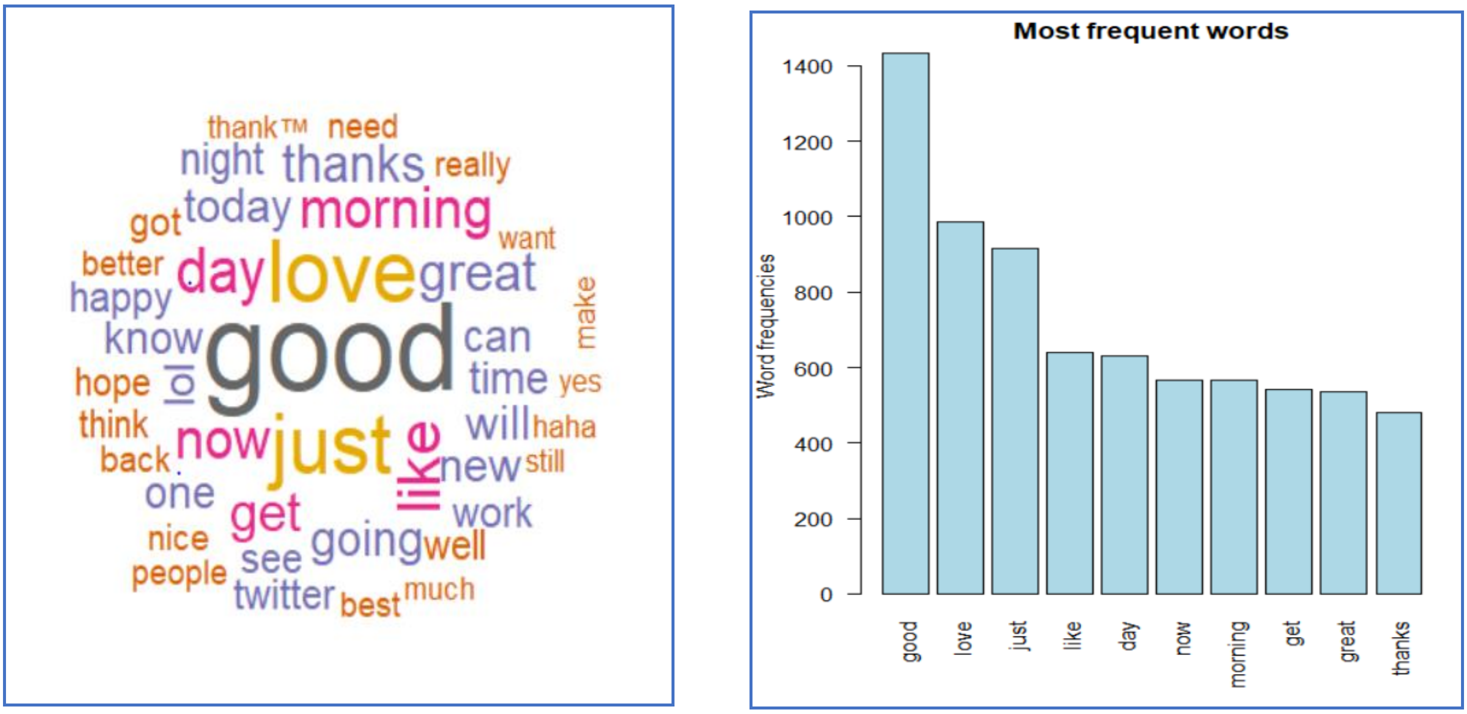


Figure 8: Word-Cloud and Bar-plot showing the common words and frequency of words in the sentiment 140 dataset

Other data pre-processing and data transformation processes were employed when building the models. They include:

- In creating models with sparse data, words that appeared only 0.5% of the time were removed from the document term matrix (DTM).
- In creating a TF-IDF classifier, the tweets were tokenized and a vectorized vocabulary was created. A DTM was generated from the usage of this using the “texttovec” library in R
- In creating a Ngrams classifier, the tweets were tokenized, and a vocabulary was created to train the model using bigrams. The bigrams were then vectorized and used in the creation of the DTM used.
- Similar process was employed in the creation of a hash classifier. The classifier’s vectorizer was created with a hash size of 2^14 and a bigram. This was used to create the DTM used.
- In normalization, the tweets used to train the model were normalized with “l1” as the method used to normalize the term vectors
- A general pre-processing step in the creation of the models is the splitting of the data into training and test data. A ratio of 80:20 was employed.

4.5 Implementation, Evaluation and Results of Algorithms on Sparse Tweets

In this research, the tweets were converted to a sparse data matrix. This was to investigate the performances of the seven developed algorithms highlighted above on a sparse tweet. The result of the algorithms in detecting depression as well as the best parameter and hyperparameter tuning are reported in this section. The results were repeatedly cross validated 10 times to ensure that the results were not by chance. The algorithms used include classification and regression trees (cart_SpD), linear discriminant analysis (lda_SpD), C5.O (c50_SpD), regularised generalized linear model (glmnet_SpD), adaptive boosting (adb_SpD), extreme gradient boost (xgb_SpD) and random forest (rf_SpD). The summary is presented in the Table 3.

Table 3. Algorithm with best tuning parameter, sampling and results using sparse tweet

ALGORITHM	BEST TUNE	SAMPLING	RESULT
cart_SpD	Cp = 0.02196	Train data (80%) Test data (20%)	Precision = 0.95 Recall = 0.17 F-measure = 0.30
lda_SpD	None	Train data (80%) Test data (20%)	Precision = 0.75 Recall = 0.86 F-measure = 0.81
c50_SpD	Trial = 1, model = tree, winnow = FALSE	Train data (80%) Test data (20%)	Precision = 0.75 Recall = 0.87 F-measure = 0.81
glmnet_SpD	Alpha = 0.1, lambda = 0.00199113	Train data (80%) Test data (20%)	Precision = 0.75 Recall = 0.86 F-measure = 0.80
adb_SpD	Mfinal = 150, max depth = 3	Train data (80%) Test data (20%)	Precision = 0.95 Recall = 0.18 F-measure = 0.30
xgb_SpD	Nrounds = 150, max depth =3, eta = 0.4, gamma = 0, col sample by tree = 0.8	Train data (80%) Test data (20%)	Precision = 0.76 Recall = 0.88 F-measure = 0.81
rf_SpD	Mtry = 113, split rule = extra tree, min node size = 1	Train data (80%) Test data (20%)	Precision = 0.79 Recall = 0.88 F-measure = 0.83

In summary, adb_SpD and cart_SpD performed better than all the algorithms in terms of precision on the sparse data because they both favoured precision of the model over the ability of the model to predict the relevant data point (depressive tweets). They both utilize a somewhat summarization approach to classify the tweets. In context, the single terms in the sparse DTM are collapsed into a single strong classifier and used to detect depression quickly. The downside of this is that the model would not generalize well on individual terms hence, the low recall. F-measure can be referred to as the balanced score (harmonic mean) of precision and recall and if recall is far below precision value, the F measure will be low as well. The other algorithms have a better spread of scores because they treat each term as a single classifier for detection of depression. These algorithms can be used to build a model that can effectively detect depression if the tweets are converted to a sparse DTM. Overall, xgb_SpD and rf_SpD had the highest scores across the evaluation metrics.

4.6 Implementation, Evaluation and Results of Regularized Generalized Linear Model on Tokenized Tweets

In contrast to the previous section, regularised generalized linear model was made to fit into tokenized tweets using TF_IDF (glmnet_TFIDF), hashing (glmnet_HSL), and Ngrams (glmnet_Ngrams). The algorithm was chosen because it was one of the best algorithms on the sparse data and an algorithm that is not constantly used in literature. To further improve the accuracy of the resultant models, the methods were implemented on normalized tokenized tweets (glmnet_TFIDF_Norm, glmnet_HSL_Norm, glmnet_Ngrams_Norm). The model was tuned as follows – 10-fold cross validation with a thresh value of 0.001 and maximum iteration at 1000 iterations. This section presents the results of the model and its corresponding parameter tuning as presented in Table 4 and 5.

Table 4. Results of GLMNET with TF_IDF, Hashing and Ngrams

FEATURE	SAMPLING	RESULTS
TF_IDF	Train data (80%) Test data (20%)	Precision = 0.92 Recall = 0.89 F-measure = 0.90
HSL	Train data (80%) Test data (20%)	Precision = 0.88 Recall = 0.86 F-measure = 0.87
NGRAM	Train data (80%) Test data (20%)	Precision = 0.90 Recall = 0.87 F-measure = 0.88

Table 5. Results of GLMNET with TF_IDF, Hashing and Ngrams using normalized data

FEATURE	SAMPLING	RESULTS
TF_IDF_NOR M	Train data (80%) Test data (20%)	Precision = 0.89 Recall = 0.91 F-measure = 0.90
HSL_NORM	Train data (80%) Test data (20%)	Precision = 0.87 Recall = 0.86 F-measure = 0.87
NGRAM_NO RM	Train data (80%) Test data (20%)	Precision = 0.87 Recall = 0.89 F-measure = 0.88

In summary, glmnet using different methods out-performed the use of sparse data. This is because creation of sparse data can lead to the removal of terms that can help in the detection of depression i.e. A term not common in tweets does not necessarily connote irrelevancy as some terms can have association with other terms that have high predictive ability. An observation of Table 5 showed an improvement in recall which is the ability of the classifier to locate the relevant data point (depressive tweets). It is safe to say that normalization of the train data can improve recall. Overall, the combination of glmnet and TF_IDF gave the best scores across the evaluation metrics.

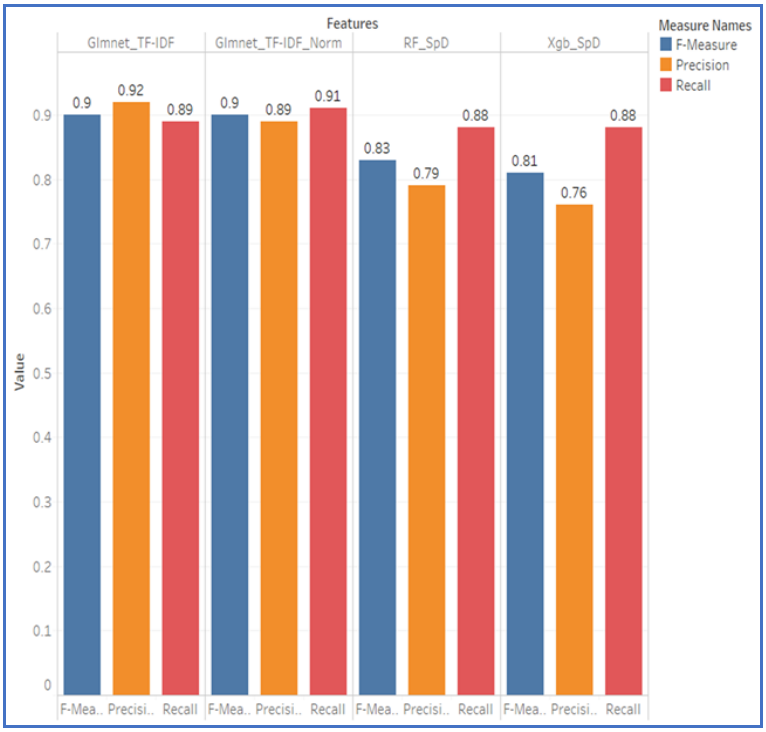
From the implementation stage, large sample size and its sampling as well as proper pre-processing of twitter specific pre-processing, proper best tune parameter and hyperparameter before training the models, tokenization and normalization process enhance and improve the effectiveness of the developed depression detection model. Evidently, the research question in (chapter 1, subsection 1.2) has been solved with the results of the implementation and all objective 1 and 2 in Table 1 have been implemented.

5. Discussion and Comparative Study

A comparative study was conducted in order to know the best performing and effective which can be embraced by the highly skilled psychologists in Nigerian health sector to detect depression among Nigerians.

5.1 Comparison of Depression Detection Models

In Figure 9, the combination of glmnet and the feature extraction methods out-performed the use of the 7 algorithms on a sparse DTM. The figure shows the comparison of the two best algorithms and two worst algorithms in both processes in terms of the specified evaluation metrics. It is also imperative to state that the combinations considered as worst in the detection of depression using feature extraction methods is not worse literally. It only had a lower evaluation metric's value when compared to the other 4 combinations. Hence, it can be used to detect depression.



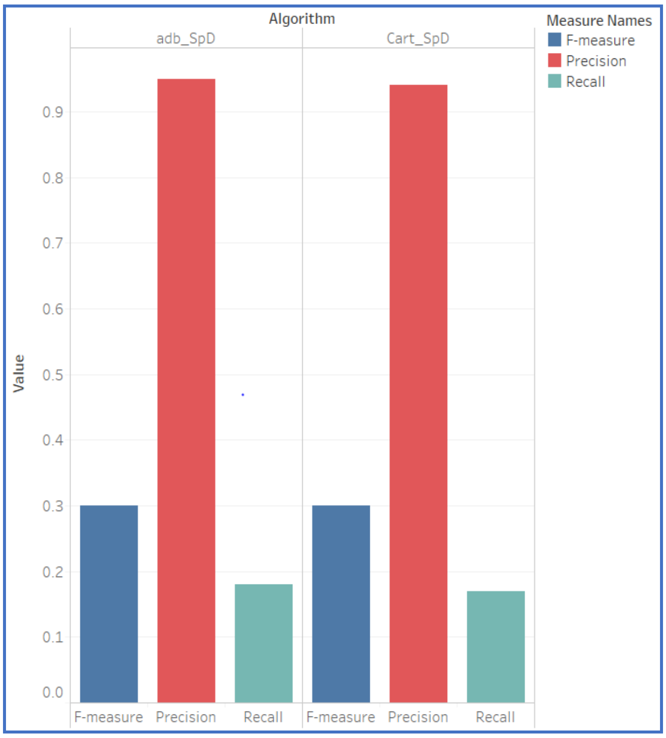


Figure 9: Comparison of 4 best effective models with 2 less effective developed model

5.2 Comparison of Depression Detection Models with Literature

Table 6 and 7 presents comparison of the models developed in the literature and the two best models developed in this research using recall one of the specified evaluation metrics. It is seen that the models out-performed some of the results in the literature.

Table 6. Results of developed depression detection models from existing literature

Author	Sample Size	Platform	Results
Tsugawa et.al (2015)	209	Twitter	Recall = 0.69
Islam et.al (2018)	7145	Facebook	Recall = 0.98
De Choudhry et.al (2013)	476	Twitter	Recall = 0.72

Table 7. Best results from the developed depression detection models

Best Models	Sample size	Result
Glmnet_TFIDF	15,234	Recall = 0.89
Glmnet_TFIDF_Norm	15,234	Recall = 0.91
Glmnet_Ngram_Norm	15,234	Recall = 0.89

The reason for the better values in this research when compared to the ones in literature could be one of the following:

- The use of only Nigerian tweets.
- The balancing method utilized (combing the positive tweets in the dataset with positive tweets from a different dataset)
- The caret package[2] in R – It searches for the best tuned parameter and hyperparameter before training the method using grid search.
- The normalization of the train data when combining it with glmnet and other feature extraction methods

[2] https://cran.r-project.org/web/packages/caret/caret.pdf

- The size of the dataset, the proportion of split used (80:20) as well as the folds of cross validation used.

From the implementation stage, large sample size and its sampling as well as proper pre-processing of twitter specific pre-processing, proper best tune parameter and hyperparameter before training the models, tokenization and normalization process enhance and improve the effectiveness of the developed depression detection model.

6. Conclusion and Recommended Future Work

Based on all the analysis, implementation of depression detection models, evaluation using proper evaluation metrics, results and cross fold validation of the results, the research project has fully solved the research question asked in (chapter 1, subsection 1.2) of this technical report as well as achieving all the set research objectives in (Chapter 1, subsection 1.3, Table 1). The reason for the better values in this research project and findings was that the balancing method utilized (combing the positive tweets in the dataset with positive tweets from a different dataset) and not creating synthetic data to solve the class imbalance as well as using the best tune parameters and normalization of the data is a great innovation which helped and enhance the good outcome of this research project. The overall outcome of the research project was encouraging an useful with good precision, recall and f measure value in most of the developed depression detection models especially in regularized generalized linear model with term frequency and inverse document frequency and normalized data (glmnet_tfdif_norm) giving recall of 0.91 which will highly useful for highly skilled psychologists in the Nigerian health sector to identify Nigerians who are suffering from depression from their expression on twitter.

Future Work :

This research project can be further improved by adopting other process of pre-processing such as leaving punctuation, emoticons and special characters for the models to understand the context of the tweets. As the implementation of the research project was accomplished in three months which was a big barrier, features extraction method such as Linguistic word count, word embedding would have been used to extract psycholinguistic features, syntactic features, character n-gram to train the models as well as using word embedding features to train a principal component analysis, and topic modelling so as to know the precise words that contributes and depict depression. In future, these processes can be used to enhance and improve the detection of depression among Nigerians on Twitter.

Acknowledgement

My deep sincere appreciation goes to my supervisor Dr Catherine Mulwa for her guidance, and quality teaching throughout the course of the research project. Thanking my parents can't be enough for their trust in me, I owe you both a great deal . I would also like to acknowledge my dearest friends Akano Lekan, Olaleye Eniola, Ayo-Akere Nelson, Aijotan Iyabo, Idowu Oluwasurefunmi, Jayeoba Omolola and Adedokun Salam for their input and immerse support throughout this research project.

References

American Psychiatric Association (2013) 'Diagnostic and statistical manual of mental disorders', (5th ed.). American Psychiatric Publishing, Washington.

Ashiekpe, J.A. and Mojaye, E.M. (2017) 'Influence of social media use on the social development of Nigerian adolescents and youths', Journal of social and management sciences, 12(1).

Azevedo, A. I. R. L. and Santos, M. F. (2008) 'Kdd, semma and crisp-dm: a parallel overview', IADIS European Conference Data Mining (IADS-DM) 2008, pp. 182–185.

Benton, A., Mitchell, M. and Hovy, D. (2017) 'Multi-task learning for mental health using social media text', arXiv preprint arXiv:1712.03538.

Berry, M. J. A. and Lino, G. S. (1997) 'Data mining techniques: for marketing, sales, and customer support', John Wiley & Sons, Inc.

Bhayani, R., Go, A., and Huang, L. (2009) 'Twitter sentiment classification using distant supervision' CS224N Project Report, Stanford, 1(12), p.2009.

De Choudhury, M., Gamon, M., Counts, S., and Horvitz, E. (2013) 'Predicting depression via social media', In ICWSM, page 2.

De Choudhury, M., Counts, S., Horvitz, E.J. and Hoff, A. (2014) 'Characterizing and predicting postpartum depression from shared Facebook data', In Proceedings of the 17th ACM conference on Computer supported cooperative work & social computing, pp. 626-638.

Hassan, A.U., Hussain, J., Hussain, M., Sadiq, M. and Lee, S. (2017) 'Sentiment analysis of social networking sites (SNS) data using machine learning approach for the measurement of depression', In 2017 International Conference on Information and Communication Technology Convergence (ICTC), pp. 138-140. IEEE.

Healthline (2018), How social media increases Depression and Loneliness. Available at: https://www.healthline.com/health-news/social-media-use-increases-depressionandloneliness [Accessed 23 November 2019]

Islam, M.R., Kabir, M.A., Ahmed, A., Kamal, A.R.M., Wang, H. and Ulhaq, A. (2018) ‘Depression detection from social network data using machine learning techniques’, Health information science and systems, 6(1), p.8.

Jamil, Z. (2017) ‘Monitoring tweets for depression to detect at-risk users’, Doctoral dissertation, University of Ottawa.

Karanasou, M., Ampla, A., Doulkeridis, C. and Halkidi, M. (2016) ‘Scalable and realtime sentiment analysis of twitter data’, In 2016 IEEE 16th International Conference on Data Mining Workshops (ICDMW), pp. 944-951. IEEE.

MonkeyLearn (2019), Sentiment Analysis: Nearly all you need to know. Available at: https://monkeylearn.com/sentiment-analysis/ [Accessed 27 November 2019]

Nadeem, M. (2016) ‘Identifying depression on Twitter’, arXiv preprint arXiv:1607.07384.

Naf'an, M.Z., Bimantara, A.A., Larasati, A., Risondang, E.M. and Nugraha, N.A.S. (2019) ‘Sentiment Analysis of Cyberbullying on Instagram User Comments’, Journal of Data Science and Its Applications, 2(1), pp.88-98.

Pak, A. and Paroubek, P. (2010) ‘Twitter as a corpus for sentiment analysis and opinion mining’, In LREc, Vol. 10, No. 2010, pp. 1320-1326.

Park, M., Cha, C. and Cha, M. (2012) ‘Depressive moods of users portrayed in Twitter’, In Proceedings of the ACM SIGKDD Workshop on healthcare informatics (HI-KDD), Vol. 2012, pp. 1-8.

Preotiuc-Pietro, D., Eichstaedt, J., Park, G., Sap, M., Smith, L., Tobolsky, V., Schwartz, H.A. and Ungar, L. (2015) ‘The role of personality, age, and gender in tweeting about mental illness’, In Proceedings of the 2nd workshop on computational linguistics and clinical psychology: From linguistic signal to clinical reality, pp. 21-30.

Praveen, S., (2009) ‘Advanced applications of data warehousing using 3-tier architecture’, DESIDOC Journal of Library & Information Technology, 29(2), pp.61-66.

Reece, A.G., Reagan, A.J., Lix, K.L., Dodds, P.S., Danforth, C.M. and Langer, E.J. (2017) ‘Forecasting the onset and course of mental illness with Twitter data. Scientific reports’, 7(1), p.13006.

Sau, A. and Bhakta, I. (2018) ‘Screening of anxiety and depression among the seafarers using machine learning technology’, Informatics in Medicine Unlocked, p.100149.

Schwartz, H.A., Eichstaedt, J., Kern, M.L., Park, G., Sap, M., Stillwell, D., Kosinski, M. and Ungar, L. (2014) ‘Towards assessing changes in degree of depression through Facebook’, In Proceedings of the Workshop on Computational Linguistics and Clinical Psychology: From Linguistic Signal to Clinical Reality, pp. 118-125.

Tsugawa, S., Kikuchi, Y., Kishino, F., Nakajima, K., Itoh, Y., and Ohsaki, H. (2015) ‘Recognizing depression from twitter activity’, In Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems, CHI ’15, pages 3187–3196, New York, NY, USA. ACM.

Wang, X., Zhang, C., Ji, Y., Sun, L., Wu, L. and Bao, Z. (2013) ‘A depression detection model based on sentiment analysis in micro-blog social network’, In Pacific-Asia Conference on Knowledge Discovery and Data Mining, (pp. 201-213). Springer, Berlin, Heidelberg.

WHO (2004), ‘Promoting mental health’, Concepts, emerging evidence, practice: Summary report.

WHO (2019), Depression, Available at: https://www.who.int/news-room/fact-sheets/detail/depression