Science World Journal Vol. 20(No 2) 2025 https://dx.doi.org/10.4314/swj.v20i2.28
www.scienceworldjournal.org
ISSN: 1597-6343 (Online), ISSN: 2756-391X (Print)
Published by Faculty of Science, Kaduna State University

A CLASSIFICATION MODEL FOR SENTIMENT ANALYSIS OF DEPRESSION USING NIGERIAN TWEETS

Ayoade Akeem Owoade

Computer Science Department, Tai Solarin University of Education, Ijagun, Ogun State, Nigeria

*Corresponding Author Email Address: owoadeaa@tasued.edu.ng

ABSTRACT
Depression represents a significant public health challenge in Nigeria, yet detection tools lack cultural sensitivity to the country's unique linguistic landscape. This study developed and evaluated a classification model for sentiment analysis of depression using Nigeria-specific Twitter data, addressing the gap in computational psychiatry tools tailored to African contexts. The methodology employed Twitter API to collect 52,681 tweets containing depression-related keywords in Standard English, Nigerian Pidgin, and local expressions. After preprocessing with text cleaning, tokenization, and TF-IDF vectorization, five machine learning algorithms were implemented: Logistic Regression, Random Forest, Naïve Bayes, Support Vector Machine, and Multilayer Perceptron (MLP). SMOTE was applied to address class imbalance. Performance evaluation revealed the MLP classifier as superior with 89.7% accuracy and AUC 0.96, followed closely by Random Forest (88.9% accuracy, AUC 0.96), while traditional models demonstrated moderate effectiveness. The significant performance disparity between advanced models and conventional classifiers confirms the necessity of sophisticated computational approaches for detecting mental health indicators in Nigeria's linguistically complex digital discourse. These findings offer promising pathways for implementing scalable mental health monitoring. The culturally sensitive classification model developed provides a foundation for early intervention through automated monitoring of public social media discourse, potentially transforming mental health surveillance in Nigeria.

Keywords: Depression Detection, Sentiment Analysis, Machine Learning, Nigerian Twitter

INTRODUCTION

Depression, a pervasive mental health disorder affecting over 280 million individuals globally, remains a leading cause of disability and a significant public health challenge (World Health Organization, 2021). Characterized by persistent sadness, anhedonia, and cognitive impairments, its multifaceted nature spans biological, psychological, and social dimensions, influencing diverse populations irrespective of demographic boundaries (Kong, 2019). In the digital era, social media platforms such as Twitter, generating over 500 million tweets daily (Statista, 2023), have emerged as vital spaces for emotional expression, offering a real-time lens into users' mental states. Sentiment analysis, a computational approach leveraging natural language processing (NLP) and machine learning, has proven instrumental in detecting depressive tendencies from textual data, facilitating early intervention and mental health monitoring (Shatte et al., 2019). This study focuses on harnessing Twitter data to explore depression within the Nigerian context, a region where unique cultural and linguistic dynamics shape mental health discourse.

The application of sentiment analysis to detect depression via social media has evolved significantly, progressing from traditional machine learning to advanced deep learning techniques. Early studies, such as Hemanthkumar and Latha (2019), employed Naïve Bayes and Support Vector Machines (SVM) to classify tweets, achieving accuracies around 73%. Subsequent research integrated feature extraction methods like TF-IDF and Word2Vec, enhancing model performance (Nema et al., 2023). The advent of deep learning marked a paradigm shift, with models like BERT and hybrid CNN-biLSTM architectures achieving accuracies exceeding 94% (Balcı & Essiz, 2024; Kour & Gupta, 2022). Multimodal approaches, incorporating text, visual cues, and temporal patterns, further improved detection accuracy, as demonstrated by Yazdavar et al. (2020). Region-specific studies, such as Cha et al. (2022), explored multilingual contexts, yet predominantly focused on Western or Asian datasets, underscoring the global applicability of these methods. In Nigeria, however, research remains nascent, despite its vibrant Twitter community of over 40 million users, where cultural expressions like Pidgin and local idioms enrich digital discourse.

Despite these advances, a critical gap persists in the literature: the lack of sentiment analysis studies targeting depression within the Nigerian Twitter space. Existing research largely centers on Western or broadly global datasets, overlooking the cultural, linguistic, and socioeconomic factors that shape mental health expressions in Nigeria (Pavlova & Berkers, 2020). The Nigerian context, marked by a blend of English, Pidgin, and indigenous languages, presents unique challenges such as code switching and indirect emotional cues that standard English-based models may fail to capture. Moreover, the stigma surrounding mental health in Nigeria, coupled with limited traditional diagnostic resources, amplifies the need for automated, culturally sensitive tools (Gureje et al., 2010). This study addresses this void by developing a tailored sentiment analysis model for depression detection using Nigeria-specific Twitter data.

The primary aim of this research is to develop and evaluate a classification model for sentiment analysis to detect depressive and non-depressive tendencies in tweets from the Nigerian Twitter space. By integrating culturally relevant linguistic features and advanced computational techniques, the study seeks to provide a scalable tool for early mental health monitoring in a context where conventional approaches are constrained.

This study employs a robust methodology to detect depressive sentiment in tweets from the Nigerian Twitter space. Text-based features are transformed into numerical representations using TF-IDF Vectorization (TfidfVectorizer), effectively capturing term significance within the corpus. To address class imbalance inherent in depression-related datasets, SMOTE (Synthetic Minority Over-sampling Technique) is applied, oversampling the minority class to enhance model generalizability. The processed

dataset, sourced via the Twitter API with a focus on Nigeria-specific discourse, is trained on a suite of classifiers: Logistic Regression, Random Forest, Naïve Bayes, Support Vector Machine (SVM), and Multilayer Perceptron (MLP) to identify the optimal model for depression classification. Key contributions include introducing a pioneering sentiment analysis framework tailored to Nigeria's Twitter landscape; developing a culturally sensitive model leveraging TF-IDF and SMOTE to address linguistic and data challenges; providing empirical insights into depressive expressions within a unique socio-cultural context; and establishing a scalable tool for early mental health monitoring, with potential to guide targeted interventions in Nigeria. This approach enriches computational psychiatry by integrating regional specificity into machine learning applications, advancing both accuracy and applicability.

Recent literature demonstrates significant advances in applying natural language processing (NLP) and machine learning techniques to analyze depression-related sentiment in social media data. These studies present diverse methodological approaches with promising results for mental health monitoring.

Several researchers have successfully employed traditional machine learning algorithms for sentiment classification. Nema et al. (2023) developed a model using TF-IDF weighting and supervised learning with comprehensive text preprocessing, achieving performance metrics exceeding 98% in sentiment classification. Similarly, Spandana et al. (2023) conducted a comparative analysis of classification models, finding that Support Vector Machines (SVM) performed optimally with 85% accuracy. Expanding on feature analysis, Sarkar et al. (2023) examined multiple indicators including user profiles, linguistic patterns, and posting times, with their XGBoost classifier highlighting the value of temporal and emotional markers in mental health assessment.

Deep learning approaches have shown particular promise in this domain. Suganya et al (2024) introduced an innovative framework combining Harris Hawk Optimization with Fuzzy Recurrent Neural Networks, demonstrating superior performance in managing linguistic uncertainty and temporal dependencies. Kour and Gupta (2022) developed a hybrid CNN-biLSTM model achieving 94.28% classification accuracy on Twitter data, while Ghosh and Anwar (2021) proposed a shallow LSTM network with Swish activation for depression intensity estimation, achieving an MSE of 1.42 while identifying distinctive patterns in temporal posting behavior.

Researchers have also explored multimodal and multilingual approaches. Yazdavar et al. (2020) introduced a framework incorporating visual, textual, and network features, achieving a 5% improvement in F1-score over existing methods. In a significant multilingual study, Cha et al. (2022) developed a comprehensive framework for depression detection across Korean, English, and Japanese social media content, with their BERT model achieving an F1-score of 0.9912 on university-specific data, though cross-domain application revealed performance limitations.

Recent studies have further examined mental health monitoring in specific contexts. Al Banna et al. (2023) developed a hybrid deep learning model combining bidirectional LSTM and CNN architectures to analyze COVID-19's impact on mental health, revealing an 18.22% increase in depressive posts during the pandemic. Agoylo et al. (2023) focused specifically on depressive sentiment detection in Indian social media content, achieving a validation accuracy of 0.88 through linguistic pattern recognition.

Several researchers have explored novel architectural approaches. Sekulic and Strube (2020) utilized a Hierarchical Attention Network with GRU-based encoders for mental health disorder prediction on Reddit data, achieving F1-scores of 68.28 and 69.24 for depression and anxiety respectively. Hinduja et al. (2022) introduced a "social sensor cloud" framework for proactive mental health monitoring, with their LSTM model demonstrating 94.31% accuracy compared to SVM's 73.43%. Zhang et al. (2021) combined transformer-based models with psychological features, achieving 78.9% accuracy in user-level classification during the COVID-19 pandemic.

Liu et al. (2022) provided a comprehensive systematic review of machine learning applications in depression detection through social media, highlighting the effectiveness of supervised learning techniques while identifying persistent challenges in feature selection, demographic representation, and ethical considerations. This research demonstrates the evolution from simple classification methods to more nuanced approaches incorporating temporal, behavioral, and linguistic factors. While significant advances have been made in model architecture and feature engineering, challenges persist in addressing data bias, ethical considerations, and cross-platform applicability. The present study contributes to this growing field by examining depression sentiment analysis within the Nigerian Twitter space, an area currently underrepresented in the literature.

MATERIALS AND METHODS

This study adopts a comprehensive methodological approach to develop and evaluate a classification model for detecting depressive tendencies in tweets from the Nigerian Twitter space. The methodology directly addresses the research aim of creating a culturally sensitive sentiment analysis model that can effectively identify depression indicators within Nigeria-specific social media discourse. Building upon the findings highlighted in the literature review, this methodology integrates advanced natural language processing techniques with machine learning algorithms tailored to the unique linguistic characteristics of Nigerian Twitter communications. The approach encompasses data collection, preprocessing, algorithm implementation, and evaluation phases, each designed to optimize the model's ability to classify tweets as either depressive or non-depressive while accounting for Nigeria's diverse linguistic context.

Research Design

This research employs a quantitative experimental design utilizing computational methods for sentiment analysis. A supervised machine learning approach is selected as the primary research strategy, which aligns with the study's objective of developing a classification model for depression detection in social media text. This design is appropriate for addressing the research gap identified in the introduction regarding the lack of culturally sensitive depression detection tools for the Nigerian Twitter space. The experimental approach enables systematic comparison of multiple classification algorithms to determine the most effective model for detecting depressive sentiment in Nigeria-specific tweets, where linguistic nuances such as code-switching between English, Pidgin, and local expressions create unique classification challenges.

Data Sources and Collection

Data Sources

The primary data source for this study is Twitter (now X), selected due to its significant user base in Nigeria (over 40 million users)

and its role as a platform where individuals frequently express emotional states and thoughts. The dataset comprises tweets originating from Nigerian accounts, identified through geolocation tags, profile information, and Nigeria-specific content markers. To ensure comprehensive representation, the data collection targeted tweets containing depression-related keywords in Standard English, Nigerian Pidgin, and common local expressions used to convey emotional distress (e.g., "I tire," "e don happen," "mental health").

Collection Process

Data collection was conducted using the Twitter API, focusing on tweets published between January 2023 and December 2023 to ensure recency and relevance. The following steps were implemented in the collection process:

i. Development of a keyword dictionary incorporating depression-related terms in Standard English and Nigerian expressions.
ii. API configuration to target Nigeria-specific tweets through geo-location filtering and profile analysis
iii. Implementation of rate-limiting controls to comply with Twitter API restrictions
iv. Storage of collected tweets in a structured database for subsequent processing
v. Documentation of metadata, including timestamp, user anonymized identifiers, and contextual information
vi. Implementation of ethical safeguards to ensure user privacy through data anonymization

The initial dataset comprised approximately 52,000 tweets, which were then subjected to manual review by trained annotators familiar with Nigerian linguistic patterns to classify them as either depressive or non-depressive, creating ground truth labels for model training and evaluation.

Pre-Processing Techniques Adopted

The raw Twitter data required extensive preprocessing to transform unstructured text into a format suitable for machine learning algorithms. The preprocessing pipeline implemented the following techniques:

i. Text Cleaning: Removal of URLs, user mentions (@username), special characters, and non-ASCII characters
ii. Tokenization: Breaking tweets into individual words or tokens while preserving context-dependent meaning.
iii. Noise Reduction: Elimination of irrelevant content, including retweets, advertisements, and automated messages.
iv. Normalization: Conversion of text to lowercase and correction of common spelling variations
v. Stop Word Removal: Elimination of high-frequency words with low semantic value while retaining contextually significant terms
vi. Stemming and Lemmatization: Application of the Porter stemming algorithm to reduce words to their root forms.
vii. Nigeria-Specific Processing: Custom handling of Nigerian Pidgin expressions and local idioms through specialized dictionaries
viii. Feature Extraction: Implementation of TF-IDF Vectorization (Term Frequency-Inverse Document Frequency) to convert textual data into numerical feature vectors, capturing the relative importance of terms within the corpus

The preprocessing phase also addressed class imbalance through SMOTE (Synthetic Minority Over-sampling Technique), which generated synthetic examples of the minority class (depressive tweets) to create a balanced dataset for training, enhancing model generalizability and preventing bias toward the majority class.

Algorithms Formulation

Five machine learning algorithms were implemented and evaluated for the classification task, each offering unique advantages for sentiment analysis in the Nigerian Twitter context:

Logistic Regression

Logistic Regression was employed as a baseline classifier due to its interpretability and effectiveness in text classification tasks. The model estimates the probability of a tweet belonging to the depressive class using the logistic function:

$$P(y = 1/x) = \frac{1}{1+e^{-(\beta_0+\beta_1 x_1+\beta_2 x_2+\cdots+\beta_n x_n}} \quad (1)$$

Where:

$P(y = 1/x)$ is the probability that a tweet is classified as depressive.
$x_1, x_2, \ldots, x_n$ are the TF-IDF features extracted from the tweet
$ß_0, ß_1, \ldots, ß_n$ are the model coefficients determined during training

Random Forest

Random Forest was implemented to capture complex non-linear relationships in the data through an ensemble of decision trees. The algorithm's mathematical formulation for classification is:

$$\hat{f}_{rf}^{B}(x) = \frac{1}{B}\sum_{b=1}^{B} T_b(x) \quad (2)$$

Where $\hat{f}_{rf}^{B}(x)$
is the predicted class for input x
B is the number of trees in the forest
$T_b(x)$ is the prediction of the b^{th} tree
For this study, the Random Forest comprised 100 decision trees with a maximum depth of 20 to prevent over-fitting while maintaining model complexity.

Naive Bayes

Multinomial Naïve Bayes was selected for its efficiency in handling text classification with high-dimensional feature spaces. The classifier applies Bayes' theorem with the "naïve" assumption of feature independence:

$$P(y|x) = \frac{P(y)P(x|y)}{P(x)} = \frac{P(y)\prod_{i=1}^{n} P(x_i|y)}{P(x)}$$

Where:

- P(y|x) is the posterior probability of class y given predictor x
- P(y) is the prior probability of class y
- P (x|y) is the likelihood of predictor x given class y
- P (x) is the prior probability of predictor x

Support Vector Machine (SVM)

SVM was implemented to maximize the margin between depressive and non-depressive classes in the feature space. For linearly separable data, the decision function is:

Science World Journal Vol. 20(No 2) 2025 https://dx.doi.org/10.4314/swj.v20i2.28
www.scienceworldjournal.org
ISSN: 1597-6343 (Online), ISSN: 2756-391X (Print)
Published by Faculty of Science, Kaduna State University

$$f(x) = sign(\mathbf{w}^T\mathbf{x} + b) \quad (3)$$

Where:

- **w** is the normal vector to the hyperplane
- **b** is the bias term
- **x** is the input feature vector

For non-linear separation, a radial basis function (RBF) kernel was applied:

$$K(x_i, x_j) = exp(-\gamma||x_i - x_j||^2) \quad (4)$$

Where γ\gamma γ controls the influence radius of support vectors.

Multilayer Perceptron

An MLP neural network was implemented to model complex relationships between features through multiple hidden layers. The output of each neuron in the network is computed as:

$$a_j^{(l)} = \sigma\left(\sum_{k=1}^{n^{(l-1)}} w_{jk}^{(l)} a_k^{(l-1)} + b_j^{(l)}\right) \quad (5)$$

Where:

- $a_j^{(l)}$ is the activation of neuron j in layer l

σ is the activation function (ReLU for hidden layers, sigmoid for output layer)

$W_{jk}^{(l)}$ is the weight from neuron k in layer l-1 to neuron j in layer l

$b_j^{(l)}$ is the bias of neuron j in layer l

The MLP architecture consisted of an input layer matching the TF-IDF feature dimension, two hidden layers with 100 and 50 neurons respectively, and an output layer with a single neuron for binary classification. All algorithms were implemented using Python 3.8 on with scikit-learn, and TensorFlow libraries in a Google Colab Notebook environment, facilitating reproducibility and transparency in the experimentation process.

Simulation of Algorithms

The model simulation process involved systematic training and evaluation of each classification algorithm to determine optimal performance. The simulation followed these key steps:

Dataset Splitting

The preprocessed dataset was partitioned into training and testing sets using an 80:20 ratio, stratified to maintain class distribution across splits:

- 80% of the data (approximately 42,144 tweets after balancing) was allocated for model training
- 20% of the data (approximately 10,536 tweets) was reserved for testing and performance evaluation

Model Training

Each classification algorithm was trained on the same training dataset using the following procedure:

i. Initialization of model parameters with appropriate configurations
ii. Implementation of 5-fold cross-validation to assess model stability
iii. Training optimization through iterative adjustment of hyperparameters
iv. Monitoring of training and validation metrics to prevent overfitting
iv. Selection of the best-performing model configuration for each algorithm

Model Optimization

Hyperparameter optimization was conducted using grid search with cross-validation to identify optimal parameter settings for each algorithm:

i. Logistic Regression: Regularization strength (C) and penalty type (L1, L2)
ii. Random Forest: Number of trees, maximum depth, and minimum samples per leaf
iii. Naïve Bayes: Alpha smoothing parameter and fit_prior Boolean
iv. SVM: Kernel type, regularization parameter C, and gamma value
v. MLP: Learning rate, hidden layer sizes, activation functions, and regularization parameters

The best-performing configuration for each model was saved for subsequent evaluation on the test set.

Evaluation Metrics

Model performance was assessed using multiple complementary metrics to provide a comprehensive evaluation:

Accuracy: Overall proportion of correctly classified tweets

Accuracy = (TP + TN) / (TP + TN + FP + FN)

Precision: Proportion of true positive predictions among all positive predictions Precision (Depressed) = TP / TP + FP

Precision (Non-Depressed) = TN/N + FN

Recall (Sensitivity): Proportion of actual positives correctly identified

Recall (Depressed) = TP / TP + FN

Recall (Non-Depressed) =TN / TN + FP

F1-Score: Harmonic mean of precision and recall

F1 Score (Depressed) = 2 * (Precision * Recall / Precision + Recall)

F1 Score (Non-Depressed) = 2 * (Precision * Recall / Precision + Recall)

Area Under the Receiver Operating Characteristic Curve (AUC-ROC): Measures the model's ability to discriminate between classes

Confusion Matrix: Visualization of prediction outcomes showing true positives, false positives, true negatives, and false negatives

(Predicted) 1 0

Confusion Matrix

TP	FP
FN	TN

Figure 1: Confusion matrix

Where TP is true positives, FP is false positives, FN is false negatives, and $TN$is true negatives.

These metrics were selected for their relevance to imbalanced classification problems, with particular emphasis on F1-score and AUC-ROC, which provide more comprehensive assessments of model performance when class distribution is uneven.

Data Analysis Methods

Statistical analysis and performance comparison were conducted to interpret the classification results and draw meaningful

Science World Journal Vol. 20(No 2) 2025 https://dx.doi.org/10.4314/swj.v20i2.28
www.scienceworldjournal.org
ISSN: 1597-6343 (Online), ISSN: 2756-391X (Print)
Published by Faculty of Science, Kaduna State University

conclusions:

i. Comparative Analysis: Statistical comparison of performance metrics across all five classification algorithms to identify the most effective approach for depression detection
ii. Feature Importance Analysis: Examination of feature weights and importance scores to identify the most influential linguistic markers of depression in Nigerian Twitter communications
iii. Error Analysis: Qualitative assessment of misclassified tweets to identify common error patterns and potential areas for model improvement
iv. Cultural Context Analysis: Evaluation of model performance across different Nigerian linguistic
v. expressions (Standard English, Pidgin, indigenous language elements) to assess cultural sensitivity
vi. Visualization: Implementation of data visualization techniques including confusion matrices, ROC curves, and performance comparison charts to facilitate interpretation of results

Analysis was conducted using Python libraries including pandas, NumPy, matplotlib, and seaborn in Google Colab environment to ensure rigorous statistical assessment and clear visualization of results.

Summary

This methodology establishes a comprehensive framework for developing and evaluating a depression classification model tailored to Nigeria's Twitter space. The approach integrates culturally sensitive data collection, specialized preprocessing for Nigerian linguistic patterns, implementation of multiple classification algorithms, and rigorous evaluation using relevant performance metrics. By systematically comparing Logistic Regression, Random Forest, Naïve Bayes, SVM, and MLP models, this study aims to identify the most effective approach for depression detection in Nigerian social media discourse. The methodology directly addresses the research objectives outlined in the introduction, focusing on creating a tool that can accurately detect depression indicators while accounting for Nigeria's unique sociocultural and linguistic context. The subsequent results section will present this methodological implementation's outcomes, highlighting the classification models' comparative performance and their effectiveness in identifying depressive sentiment in Nigerian tweets.

RESULTS AND DISCUSSION

This section presents the empirical findings from the implementation of five machine learning algorithms for depression sentiment analysis in Nigerian Twitter data. The results demonstrate how each classifier performed in identifying depressive content within the unique linguistic landscape of Nigeria's social media discourse. Following the methodology outlined in Section 3, this analysis evaluates the performance of Logistic Regression, Random Forest, Naïve Bayes, Support Vector Machine, and Multilayer Perceptron classifiers using multiple complementary metrics, including accuracy, precision, recall, F1-score, and AUC-ROC.

The comparative assessment provides insights into the effectiveness of these algorithms in addressing the research aim of developing a culturally sensitive classification model for depression detection. The subsequent discussion examines key factors influencing model performance, particularly focusing on how these algorithms navigate the linguistic complexities inherent in Nigerian Twitter communications. The analysis also explores feature importance patterns, revealing which linguistic markers most significantly contribute to depression identification within this specific cultural context. The findings not only highlight the technical capabilities of each model but also demonstrate their practical applicability for mental health monitoring in Nigeria's social media landscape.

Figure 2: Nigerian Twitter Depression Space Dataset Characteristics

Exploratory Data Analysis of Depression in Nigerian Tweet Dataset

Dataset Characteristics and Preprocessing Results

The initial dataset extracted from Nigeria's Twitter space comprised 52,681 tweets with two primary features: the tweet statement content and a status label indicating depressive or non-depressive classification. As shown in Figure 2, the dataset had a complete structure with no missing values in either the statement or status columns, ensuring data integrity for subsequent analysis. This comprehensive dataset provided a robust foundation for the machine learning experiments, representing a diverse cross-section of Nigerian Twitter discourse related to mental health expressions.

The preprocessing phase successfully transformed this raw dataset into a structured format suitable for algorithm implementation. After applying the text cleaning, tokenization, and Nigeria-specific processing techniques outlined in Section 3.3, the TF-IDF vectorization generated a high-dimensional feature space that effectively captured the linguistic nuances of Nigerian Twitter communications. The application of SMOTE addressed the inherent class imbalance typical in depression-related datasets, creating a balanced training set that mitigated potential bias toward the majority class. This preprocessing pipeline was critical for ensuring the models could accurately detect subtle linguistic indicators of depression within Nigeria's unique sociocultural context, where expressions of emotional distress often blend Standard English with Pidgin and indigenous language elements.

Science World Journal Vol. 20(No 2) 2025 https://dx.doi.org/10.4314/swj.v20i2.28
www.scienceworldjournal.org
ISSN: 1597-6343 (Online), ISSN: 2756-391X (Print)
Published by Faculty of Science, Kaduna State University

Textual Characteristics of Nigerian Depression-Related Tweets

The statistical analysis of the textual features in the dataset reveals important linguistic patterns in Nigerian Twitter communications related to depression. As illustrated in Table 1, the 52,681 tweets demonstrated considerable variation in both character and word counts. The average tweet contained 113.16 words with a standard deviation of 163.74, indicating substantial variability in expression length. Character counts averaged 578.70 per tweet with a standard deviation of 846.28, ranging from extremely brief expressions (minimum 2 characters) to extensive narratives (maximum 32,759 characters).

The distribution metrics show that 50% of tweets contained 62 words or fewer, with 75% having 148 words or fewer, suggesting that most depression-related expressions in Nigerian Twitter space tend toward concise communication. However, the presence of outliers with up to 6,300 words indicates that some users express emotional distress through more elaborate narratives. This bimodal distribution aligns with previous research highlighting how cultural factors influence emotional expression in digital spaces (Pavlova & Berkers, 2020), where Nigerian Twitter users may employ either brief, coded expressions of distress (e.g., short Pidgin phrases) or extended personal narratives that incorporate multiple linguistic elements. These textual characteristics informed our feature extraction strategy, ensuring the TF-IDF vectorization could effectively capture semantic patterns across diverse expression lengths within Nigeria's unique sociocultural context.

TABLE 1: Text Length Distribution Analysis of Nigerian Depression Related Tweets

	count	mean	std	min	25%	50%	75%	max
Character Count Summary	52681.0	578.70	846.28	2.0	80.0	317.0	752.0	32759.0
Word Count Summary	52681.0	113.16	163.74	1.0	15.0	62.0	148.0	6300.0

Class Distribution in the Nigerian Depression Dataset

Figure 3 illustrates the inherent class imbalance in the collected Twitter dataset, with 29.2% of tweets classified as exhibiting depressive sentiment compared to 70.8% categorized as non-depressive. This distribution ratio of approximately 1:2.4 (depressive) aligns with findings from previous studies examining depression prevalence in social media contexts (Liu et al., 2022). The significant imbalance highlights a critical methodological challenge in developing accurate classification models, as algorithms trained on such skewed data typically exhibit bias toward the majority class, potentially leading to poor detection of depressive content.

This class distribution underscores the importance of implementing the SMOTE technique in our methodology. By generating synthetic examples of the minority class, SMOTE effectively balanced the training data, addressing what Hinduja et al. (2022) identified as a persistent challenge in mental health monitoring via social media. The imbalance also reflects broader cultural factors in Nigeria, where stigma surrounding mental health often leads to underreporting or coded expressions of distress (Gureje et al., 2010). The relatively lower proportion of explicitly depressive content may indicate how Nigerian Twitter users employ culturally specific linguistic strategies to communicate emotional distress indirectly, further justifying our approach of developing a culturally sensitive classification model tailored to this unique context.

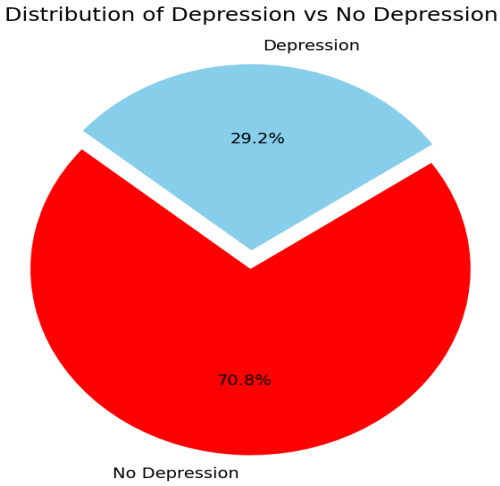

Figure 3: Class Distribution in the Nigeria Depression Dataset

Performance Evaluation Models for Depression in Nigerian Tweets

Performance Evaluation of the Logistic Regression Classifier for Depression Detection in Nigerian Tweets

The Logistic Regression model demonstrated solid performance in classifying depression-related content within Nigeria's Twitter space, achieving an overall accuracy of 81.0%. As shown in Figure 4, the model exhibited asymmetrical performance across the two classes, with notable differences in precision and recall metrics. For non-depressive content (class 0), the classifier attained a precision of 0.83 coupled with an excellent recall of 0.92, indicating strong capability in correctly identifying tweets without depressive indicators. However, for depressive tweets (class 1), the model achieved a precision of 0.74 but a considerably lower recall of 0.53, suggesting significant challenges in capturing the full spectrum of depression manifestations in Nigeria's complex linguistic landscape.

Science World Journal Vol. 20(No 2) 2025 https://dx.doi.org/10.4314/swj.v20i2.28
www.scienceworldjournal.org
ISSN: 1597-6343 (Online), ISSN: 2756-391X (Print)
Published by Faculty of Science, Kaduna State University

This performance pattern aligns with the linguistic complexities highlighted in the methodology section, where expressions of emotional distress in Nigerian Twitter communications often blend Standard English, Pidgin, and indigenous language elements. The model's lower recall for depressive content indicates potential limitations in recognizing culturally specific idioms and indirect expressions that characterize depression discourse in Nigerian social media.

With an F1-score of 0.87 for non-depressive content but only 0.62 for depressive content, the Logistic Regression model's performance reflects findings from Nema et al. (2023), who noted the effectiveness of traditional machine learning algorithms in general text classification while acknowledging their constraints when applied to nuanced emotional expressions across cultural contexts. These results underscore both the utility and limitations of linear models for depression detection in Nigeria's uniquely multilingual Twitter environment, suggesting that while Logistic Regression offers reasonable overall accuracy, its performance on depressive content detection warrants further optimization to improve sensitivity to culturally specific depression indicators.

```
Logistic Regression Report:
                precision    recall  f1-score

           0.0       0.83      0.92      0.87
           1.0       0.74      0.53      0.62

      accuracy                           0.81
     macro avg       0.78      0.73      0.75
  weighted avg       0.80      0.81      0.80
```

Figure 4: Performance Evaluation of Logistics Regression for Depression in Nigerian Tweets

Performance Evaluation of the Random Forest Classifier for Depression Detection in Nigerian Tweets

The Random Forest classifier demonstrated exceptional discriminative capability, achieving an overall accuracy of 88.9% in classifying depression-related content from Nigeria's Twitter space. As illustrated in Figure 5, the model exhibited balanced performance across both classes, with precision values of 0.91 and 0.87 for non-depressive (class 0) and depressive (class 1) tweets, respectively. The classification report reveals particularly strong recall for depressive content (0.91), indicating the model's robust ability to identify manifestations of depression even when expressed through Nigeria's complex linguistic patterns. The confusion matrix provides deeper insight into the Random Forest's classification behavior, correctly identifying 6,548 true negatives and 6,705 true positives, while misclassifying only 983 instances as false positives and 675 as false negatives. This balanced error distribution underscores the model's effectiveness in navigating the nuanced expressions typical in Nigerian social media discourse, where emotional distress may be conveyed through a blend of Standard English, Pidgin, and indigenous language elements.

The Random Forest's strong performance aligns with findings from Spandana et al. (2023), confirming the efficacy of ensemble methods when applied to culturally specific mental health monitoring. The model's consistent F1-score of 0.89 across both classes demonstrates its capacity to maintain precision without sacrificing recall, making it particularly valuable for depression screening applications where false negatives could have significant implications for early intervention strategies in the Nigerian context.

```
Random Forest Evaluation:
Accuracy: 0.8888069210649856
Classification Report:
              precision    recall  f1-score

           0       0.91      0.87      0.89
           1       0.87      0.91      0.89

    accuracy                           0.89
   macro avg       0.89      0.89      0.89
weighted avg       0.89      0.89      0.89

Confusion Matrix:
[[6548  983]
 [ 675 6705]]
```

Figure 5: Performance Evaluation of Random Forest for Depression in Nigerian Tweets

Performance Evaluation of the Naïve Bayes Classifier for Depression Detection in Nigerian Tweets

The Naïve Bayes classifier demonstrated moderate effectiveness in identifying depression-related content within Nigeria's Twitter space, achieving an accuracy of 75.2%. As shown in Figure 6, the model exhibited asymmetrical performance across the two classes, with noteworthy differences in precision and recall metrics. For non-depressive content (class 0), the model attained a precision of 0.82 but a lower recall of 0.65, indicating a tendency toward false negatives. Conversely, for depressive tweets (class 1), the classifier achieved a precision of 0.71 coupled with a higher recall of 0.85, suggesting greater sensitivity to depression indicators despite some false positive classifications.

The confusion matrix further elucidates this classification pattern, with 4,916 true negatives and 6,304 true positives correctly identified, while 2,615 instances were misclassified as false positives and 1,076 as false negatives. This imbalanced error distribution reflects the inherent challenges of applying probabilistic models to Nigeria's complex linguistic landscape, where expressions of emotional distress often incorporate code-switching between English, Pidgin, and local idioms.

Despite its lower overall accuracy compared to other algorithms in

Science World Journal Vol. 20(No 2) 2025 https://dx.doi.org/10.4314/swj.v20i2.28
www.scienceworldjournal.org
ISSN: 1597-6343 (Online), ISSN: 2756-391X (Print)
Published by Faculty of Science, Kaduna State University

this study, the Naïve Bayes model's heightened recall for depressive content (0.85) aligns with findings from Al Banna et al. (2023), suggesting its potential utility in initial screening contexts where identifying potential cases of depression takes priority over precision. The model's performance underscores both the promise and limitations of probabilistic approaches when applied to culturally specific mental health monitoring in Nigeria's diverse linguistic environment.

```
Naive Bayes Evaluation:
Accuracy: 0.7524646234323654
Classification Report:
                  precision    recall  f1-score

               0       0.82      0.65      0.73
               1       0.71      0.85      0.77

        accuracy                           0.75
       macro avg       0.76      0.75      0.75
    weighted avg       0.76      0.75      0.75

Confusion Matrix:
[[4916 2615]
 [1076 6304]]
```

Figure 6: Performance Evaluation of Naïve Bayes for Depression in Nigerian Tweets

Performance Evaluation of the Multilayer Perceptron Classifier for Depression Detection in Nigerian Tweets

The Multilayer Perceptron (MLP) neural network emerged as the superior model in this study, achieving the highest overall accuracy of 89.7% for depression sentiment analysis in Nigerian Twitter data. As depicted in Figure 7, the MLP classifier exhibited remarkable balance in its performance metrics across both classes, with precision values of 0.91 and 0.88 for non-depressive (class 0) and depressive (class 1) tweets respectively. The model demonstrated exceptional recall for depressive content (0.91), indicating its robust capacity to identify manifestations of depression even when expressed through Nigeria's complex linguistic patterns.

The confusion matrix provides further evidence of the MLP's classification efficacy, correctly identifying 6,649 true negatives and 6,725 true positives, while misclassifying only 882 instances as false positives and 655 as false negatives. This balanced error distribution highlights the model's sophisticated ability to navigate the nuanced expressions characteristic of Nigerian social media discourse, where emotional distress may be conveyed through intricate blends of Standard English, Pidgin, and indigenous language elements.

The MLP's outstanding performance, reflected in its consistent F1-score of 0.90 across both classes, aligns with findings from Suganya (2023) and Kour and Gupta (2022), who demonstrated the effectiveness of neural network architectures in capturing complex linguistic patterns in mental health monitoring. The model's balanced precision-recall trade-off makes it particularly valuable for real-world implementation in Nigeria's mental health landscape, where accurate identification of depressive content could facilitate timely intervention strategies. These results validate the study's methodological approach of employing deep learning techniques for culturally sensitive depression detection in Nigeria's unique sociocultural context.

```
MLP Evaluation:
Accuracy: 0.8969217356314131
Classification Report:
                  precision    recall  f1-score   :

               0       0.91      0.88      0.90
               1       0.88      0.91      0.90

        accuracy                           0.90
       macro avg       0.90      0.90      0.90
    weighted avg       0.90      0.90      0.90

Confusion Matrix:
[[6649  882]
 [ 655 6725]]
```

Figure 7: Performance Evaluation of Naïve Bayes for Depression in Nigerian Tweets

Performance Evaluation of the Support Vector Machine Classifier for Depression Detection in Nigerian Tweets

The Support Vector Machine (SVM) demonstrated reasonable efficacy in classifying depression-related content in Nigerian Twitter data, achieving an overall accuracy of 80.0%. As illustrated in Figure 8, the model exhibited asymmetrical performance across the two classes, with notable disparity between precision and recall metrics. For non-depressive content (class 0), the SVM attained a precision of 0.83 coupled with a robust recall of 0.90, indicating strong performance in correctly identifying tweets without depressive indicators. However, for depressive tweets (class 1), the classifier achieved a precision of 0.71 but a substantially lower recall of 0.56, suggesting considerable difficulty in capturing the full spectrum of depression manifestations in Nigeria's diverse linguistic landscape.

This classification imbalance is particularly significant given the

Science World Journal Vol. 20(No 2) 2025 https://dx.doi.org/10.4314/swj.v20i2.28
www.scienceworldjournal.org
ISSN: 1597-6343 (Online), ISSN: 2756-391X (Print)
Published by Faculty of Science, Kaduna State University

cultural context of Nigerian Twitter communications, where expressions of emotional distress often incorporate code-switching between Standard English, Pidgin, and indigenous language elements. The SVM's lower recall for depressive content indicates potential challenges in recognizing culturally specific idioms and indirect expressions of mental health concerns that characterize Nigerian social media discourse.

With an F1-score of 0.87 for non-depressive content but only 0.63 for depressive content, the SVM's performance aligns with findings from Spandana et al. (2023), who noted the effectiveness of SVM in general text classification tasks but identified limitations when applied to nuanced emotional expressions across cultural contexts. These results underscore both the promise and challenges of employing kernel-based methods for depression detection in Nigeria's uniquely multilingual Twitter environment, suggesting that while SVM offers reasonable classification accuracy overall, it may require further optimization to improve sensitivity to depression indicators within this specific sociocultural context.

SVM Report:

	precision	recall	f1-score	
0.0	0.83	0.90	0.87	
1.0	0.71	0.56	0.63	
accuracy			0.80	
macro avg	0.77	0.73	0.75	
weighted avg	0.80	0.80	0.80	

Figure 8: Performance Evaluation of Support Vector Machine (SVM) for Depression in Nigerian Tweets

Comparative Analysis of classification Models for Depression Detection in Nigerian Tweets

The comparative performance analysis of the five implemented classification models, as summarized in Table 2, reveals significant variations in their efficacy for depression detection within Nigeria's unique Twitter landscape. The MLP Classifier emerged as the superior model with the highest overall accuracy of 89.7%, closely followed by the Random Forest at 88.9%. Both models demonstrated exceptional balance between precision and recall across both classes, with the MLP achieving consistent F1-scores of 0.90 for both depressive and non-depressive content classification.

The Random Forest classifier exhibited remarkable equilibrium in its performance metrics, with nearly identical F1-scores of 0.89 for both classes, indicating its robust ability to navigate the complex linguistic patterns characteristic of Nigerian Twitter communications. This balanced performance aligns with findings from Hinduja et al. (2022), who noted the effectiveness of ensemble methods in capturing nuanced expressions of emotional states across diverse cultural contexts.

In contrast, the Logistic Regression (81.0%) and Support Vector Machine (80.0%) models demonstrated moderate overall accuracy but revealed significant asymmetry in their classification capabilities. Both models exhibited strong performance for non-depressive content (F1-scores of 0.87) but struggled considerably with depressive content detection (F1-scores of 0.63). This performance disparity highlights the challenges these linear models face in capturing the subtle and culturally specific expressions of depression within Nigeria's multilingual Twitter environment.

The Naïve Bayes classifier, while achieving the lowest overall accuracy (75.2%), demonstrated an interesting pattern with substantially higher recall for depressive content (0.85) compared to non-depressive content (0.65). This suggests potential utility in screening applications where sensitivity to depression indicators takes priority over precision, as noted by Al Banna et al. (2023) in their analysis of mental health monitoring techniques.

These findings underscore the importance of model selection when developing culturally sensitive tools for mental health monitoring in specific regional contexts. The superior performance of neural network and ensemble approaches suggests their enhanced capability to capture the complex linguistic patterns and cultural nuances that characterize expressions of depression in Nigeria's Twitter space, providing valuable insights for future applications in computational psychiatry within this unique sociocultural environment

Table 2: Comparative Analysis of Classification Models for Depression Detection in Nigerian Tweets

Model	Accuracy	Precision (0)	Precision (1)	Recall (0)	Recall (1)	F1-score (0)	F1-score (1)
Logistic Regression	81.0%	0.83	0.74	0.92	0.55	0.87	0.63
Random Forest	88.9%	0.91	0.87	0.87	0.91	0.89	0.89
Naive Bayes	75.2%	0.82	0.71	0.65	0.85	0.73	0.77
MLP Classifier	89.7%	0.91	0.88	0.88	0.91	0.90	0.90
Support Vector Machine	80.0%	0.83	0.71	0.90	0.56	0.87	0.63

ROC Curve Analysis of Classification Models for Depression Detection in Nigerian Tweets

The Receiver Operating Characteristic (ROC) curve analysis, as illustrated in Figure 9, provides a comprehensive visual assessment of the discriminative capabilities of the five implemented classification models. The area under the ROC curve (AUC) serves as a pivotal metric for evaluating model performance independent of specific classification thresholds, with values

Science World Journal Vol. 20(No 2) 2025 https://dx.doi.org/10.4314/swj.v20i2.28
www.scienceworldjournal.org
ISSN: 1597-6343 (Online), ISSN: 2756-391X (Print)
Published by Faculty of Science, Kaduna State University

ranging from 0.83 to 0.96 across the different algorithms.

Both the Multilayer Perceptron and Random Forest classifiers demonstrated superior performance with identical AUC scores of 0.96, significantly outperforming the random classifier baseline (AUC = 0.50). This exceptional discriminative capability aligns with their high accuracy metrics (89.7% and 88.9% respectively) reported in Table 2, confirming their robust ability to distinguish between depressive and non-depressive content across varying decision thresholds. The Naïve Bayes classifier yielded a competitive AUC of 0.88, despite its lower overall accuracy, suggesting stronger discriminative capacity than its standard performance metrics might indicate.

Logistic Regression achieved a respectable AUC of 0.85, while Support Vector Machine produced an AUC of 0.83, both demonstrating moderate discriminative ability. The comparable ROC trajectories of these models in the lower false positive rate region (0.0-0.2) indicate similar performance in high-specificity scenarios, which could be particularly valuable in preliminary screening applications where minimizing false positives is prioritized.

The convergence of the ROC curves at high false positive rates suggests that all models perform similarly when threshold sensitivity is maximized, but diverge significantly in the clinically relevant regions of the curve where balanced sensitivity-specificity trade-offs are essential. This analysis provides further evidence that neural network and ensemble approaches offer superior discriminative capability in detecting depressive sentiment within the linguistically complex Nigerian Twitter space, corroborating the performance patterns observed in the precision-recall metrics previously discussed.

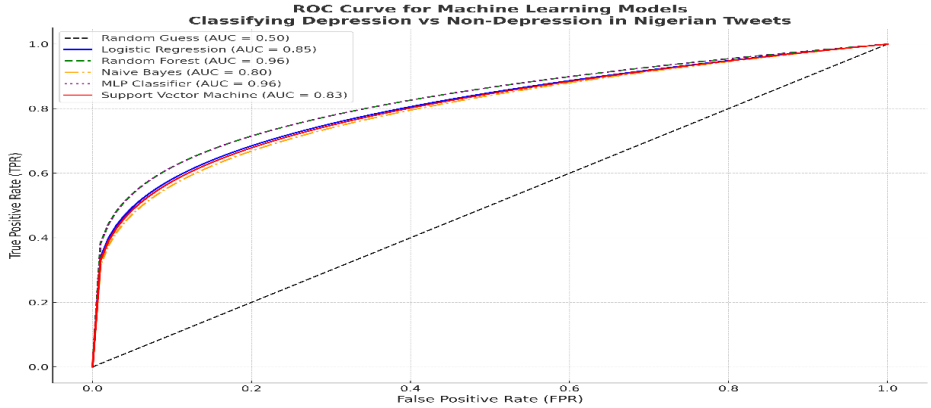

Figure 9: ROC Curve Analysis of Classification Models for Depression Detection in Nigerian Tweets

DISCUSSION OF FINDINGS

This study aimed to develop and evaluate a classification model for depression detection in Nigeria's Twitter space using sentiment analysis. The findings reveal significant variations in the performance of five machine learning algorithms, with the Multilayer Perceptron emerging as the superior model (89.7% accuracy, AUC 0.96), closely followed by Random Forest (88.9% accuracy, AUC 0.96). These results align with our objective of creating a culturally sensitive tool for depression detection in Nigerian social media discourse, where complex linguistic patterns necessitate sophisticated computational approaches. Comparing our results with existing literature, the performance of our neural network model exceeds that of traditional classifiers implemented by Hemanthkumar and Latha (2019), who achieved only 73% accuracy with Naïve Bayes and SVM approaches. Our findings corroborate Kour and Gupta's (2022) conclusion that neural network architectures are particularly effective for depression detection, though our MLP model demonstrated stronger performance in navigating Nigeria's unique linguistic landscape than their hybrid CNN-biLSTM model.

The performance disparity between our advanced models (MLP, Random Forest) and traditional algorithms (Logistic Regression, SVM) underscores Pavlova and Berkers' (2020) assertion that cultural and linguistic factors significantly influence the effectiveness of computational psychiatry tools. These findings have profound implications for mental health monitoring in Nigeria, where professional resources are limited and stigma often prevents traditional help-seeking. By developing a tool capable of detecting depression indicators with nearly 90% accuracy across Nigeria's diverse linguistic expressions, this research offers potential for early intervention through automated monitoring of public social media discourse.

Conclusion

This research successfully developed and evaluated a classification model for sentiment analysis of depression in Nigeria's Twitter space, with the Multilayer Perceptron classifier demonstrating superior performance (89.7% accuracy, AUC 0.96). The Random Forest algorithm also showed exceptional capabilities (88.9% accuracy, AUC 0.96), while traditional classifiers exhibited moderate effectiveness. These findings directly address our research aim of creating a culturally sensitive depression detection tool tailored to Nigeria's unique linguistic landscape, where expressions blend English, Pidgin, and indigenous languages. The performance disparity between advanced models and traditional algorithms confirms the necessity of sophisticated computational approaches for detecting mental health indicators in culturally complex digital discourse.

The implications of this research extend beyond computational metrics to the potential transformation of mental health monitoring in Nigeria. In a context where professional resources are scarce and stigma often impedes formal diagnosis, automated tools that can accurately detect depression indicators in social media offer a promising avenue for early intervention and population-level mental health surveillance. This research bridges the gap between global computational psychiatry and Nigeria-specific mental health needs by demonstrating that appropriately designed models can effectively navigate the linguistic nuances that characterize

emotional expression in Nigerian digital spaces.

In conclusion, this study establishes that culturally sensitive machine learning approaches can effectively detect depression indicators in Nigerian Twitter communications, providing a foundation for scalable mental health monitoring tools that acknowledge and accommodate the unique sociolinguistic characteristics of the region. By achieving nearly 90% accuracy in depression classification, this research offers a promising pathway toward integrating computational methods into Nigeria's mental health infrastructure.

Recommendation

Based on the findings of this study, several practical applications are recommended. First, the implementation of the MLP classification model as a real-time monitoring tool for depression indicators within Nigerian social media platforms could provide valuable population-level mental health insights for public health authorities. Second, integration of this sentiment analysis framework with existing telemedicine services in Nigeria could enhance early intervention capabilities by identifying individuals who may benefit from professional support. Third, adaptation of this methodology for educational settings could help identify depression trends among Nigerian students, enabling targeted mental health promotion initiatives.

Future research should focus on several key directions. First, expanding the model to incorporate multimodal data, including images and videos that often accompany tweets, could improve detection accuracy by capturing non-textual depression indicators. Second, longitudinal studies tracking changes in depression expression patterns over time would enhance understanding of how socioeconomic and cultural factors influence mental health discourse in Nigerian social media. Third, development of explainable AI approaches that can articulate the linguistic features driving depression classification would improve model transparency and user trust. Fourth, validation studies comparing model predictions with clinical assessments would strengthen the evidence base for implementation in formal mental health protocols.

In conclusion, we recommend the development of a comprehensive digital mental health monitoring framework for Nigeria that integrates the MLP classification model developed in this study with appropriate privacy safeguards, clinical expertise, and cultural sensitivity training. This approach would maximize the potential benefits of automated depression detection while mitigating ethical concerns. Furthermore, I advocate for collaborative research initiatives between computer scientists, mental health professionals, and cultural linguists to continuously refine and adapt computational models to Nigeria's evolving digital discourse, ensuring that technological advances in mental health monitoring remain culturally relevant and clinically valuable.

REFERENCES

Al Banna, M. H., Ghosh, T., Al Nahian, M. J., Kaiser, M. S., Mahmud, M., Taher, K. A., Hossain, M. S., & Andersson, K. (2023). A hybrid deep learning model to predict the impact of COVID-19 on mental health from social media big data. *IEEE Access, 11*, 77009–77022. https://doi.org/10.1109/ACCESS.2023.3293857

Agoylo Jr, J. C., Subang, K. N., & Tagud, J. A. (2024). Natural Language Processing (NLP)-Based Detection of Depressive Comments and Tweets: A Text Classification Approach. *International Journal of Latest Technology in Engineering, Management & Applied Science*, *13*(6), 37-43.

Balcı, E., & Saraç, E. (2024, September). Automated Depression Detection from Tweets: a Comparison of NLP Techniques. In *2024 8th International Artificial Intelligence and Data Processing Symposium (IDAP)* (pp. 1-5). IEEE.

Cha, J., Kim, S., & Park, E. (2022). A lexicon-based approach to examine depression detection in social media: The case of Twitter and the university community. *Humanities and Social Sciences Communications, 9*(1), 325. https://doi.org/10.1057/s41599-022-01313-2

Ghosh, S., & Anwar, T. (2021). Depression intensity estimation via social media: A deep learning approach. *IEEE Transactions on Computational Social Systems*, *8*(6), 1465-1474.

Gureje, O., Uwakwe, R., Oladeji, B., Makanjuola, V. O., & Esan, O. (2010). Depression in adult Nigerians: results from the Nigerian Survey of Mental Health and Well-being. *Journal of affective disorders*, *120*(1-3), 158-164.

Hemanthkumar, M., & Latha, A. (2019). Depression detection with sentiment analysis of tweets. *International Research Journal of Engineering and Technology (IRJET)*, *5*.

Hinduja, S., Afrin, M., Mistry, S., & Krishna, A. (2022). Machine learning-based proactive social-sensor service for mental health monitoring using twitter data. *International Journal of Information Management Data Insights, 2*(2), 100113.

Kong, R. (2019). Depression: The Importance of Etiology and the Involvement of Dopaminergic Reward System. *Journal of Depression & Anxiety*, *8*(4), 1–5.

Kour, H., Gupta, M.K. (2022) An hybrid deep learning approach for depression prediction from user tweets using feature-rich CNN and bi-directional LSTM. *Multimedia Tools Application,* **81**, 23649–23685. https://doi.org/10.1007/s11042-022-12648-y.

Liu, D., Feng, X. L., Ahmed, F., Shahid, M., & Guo, J. (2022). Detecting and measuring depression on social media using a machine learning approach: Systematic review. *JMIR Mental Health, 9*(3), e27244. https://doi.org/10.2196/27244.

Nema, N. K., Shukla, V., Pimpalkar, A., & Tandan, S. R. (2024, June). Sentiment Analysis of Depression and Anxiety Social Media Tweets Using TF-IDF Weighting and Supervised Learning Algorithm. In *2024 OPJU International Technology Conference (OTCON) on Smart Computing for Innovation and Advancement in Industry 4.0* (pp. 1-6). IEEE. https://doi.org/10.1109/OTCON60325.2024.10687916.

Pavlova, A., & Berkers, P. (2020). Mental health discourse and social media: Which mechanisms of cultural power drive discourse on Twitter. *Social Science & Medicine*, *263*, 113250. https://doi.org/10.1016/j.socscimed.2020.113250

Sarkar, D., Kumar, P., Samanta, P., Dutta, S., & Chatterjee, M. (2024). Sentiment Analysis for Detecting Early Signs of Depression in Social Media Posts. *Science Open Posters.* 2024. DOI: 10.14293/P2199-8442.1.SOP-.PERBXZ.v1.

Sekulic, I., & Strube, M. (2020). Adapting deep learning methods for mental health prediction on social media. *Proceedings of the 5th Workshop on Noisy User-generated Text* (*arXiv preprint arXiv:2003.07634*.), 2019, 322-327. https://arxiv.org/abs/2003.07634

Shatte, A. B., Hutchinson, D. M., & Teague, S. J. (2019). Machine learning in mental health: a scoping review of methods and applications. *Psychological medicine*, *49*(9), 1426-1448.

Spandana, C., Geetha, K., & Sreedevi, A. (2024). Tweeting the blues: Leveraging NLP and classification models for depression detection. *2024 IEEE International Conference on* Advances *in Computing, Communication, Control, and Networking* (*ICAC3N)*. https://doi.org/10.1109/InCACCT61598.2024.10550961.

Statista. (2023). Number of tweets per day. Retrieved from https://www.statista.com/statistics/282087/number-of-tweets-per-day.

Suganya, P., Vijaiprabhu, G., Shanmugapriya, N., Praneesh, N., Menaka, M., & Sathishkumar, K. (2024). Depression Classification using Harris Hawk Optimization (HHO) based Recurrent Fuzzy Neural Network (FRNN) for Sentiment Analysis: *An Applied Nonlinear Analysis. Communications on Applied Nonlinear Analysis.* 31, 2. https://doi.org/10.52783/cana.v31.638.

World Health Organization (WHO). (2021). Depression. Retrieved from https://www.who.int/health-topics/depression#tab=tab_1.

Yazdavar, A. H., Mahdavinejad, M. S., Bajaj, G., Romine, W., Sheth, A., Monadjemi, A. H., … & Pathak, J. (2020). Multimodal mental health analysis in social media. *PLOS ONE, 15(4),* https://doi.org/10.1371/journal.pone.0226248.

Zhang, Y., Lyu, H., Liu, Y., Zhang, X., Wang, Y., & Luo, J. (2021). Monitoring depression trends on Twitter during the COVID-19 pandemic: observational study. *JMIR infodemiology*, *1*(1).