A Machine Learning Approach for Depression Detection in Sinhala-English Code-Mixed Language

Fahd Fazal, Cassim Farook

***Abstract*— The most significant and prevalent mental illness in the world today is depression. Depression is a treatable condition. In Sri Lanka, people would not recognize depression in its earliest stages, when left un-treated contributes to a high rate of suicidal behaviour. Most people tend to share depressive feelings and thoughts intentionally or unintentionally on social media. Social media is a viable source of data for researchers engaged in attempts for early detection of depression. There are numerous studies done in the field of depression detection using English language. But depression detection is less researched using code-mixed languages which is a prevalent trend on most social media users. Therefore, the objective of this research is proposing a machine learning approach to detect depression in Sinhala-English code-mixed textual contents posted in social media. Several base and ensemble machine learning algorithms were tested to detect depression using a dataset collected from social media platforms. Evaluating the performance results, it was found that ExtraTreesClassifier outperformed other machine learning algorithms with the highest accuracy of 79.13%, which was then incorporated into the research prototype.**

***Keywords*— Depression Detection, Machine Learning, Deep Learning, Natural language Processing, Ensemble Classifier, Sinhala-English code-mixed language.**

I. Introduction

A. Depression

Depression is a mood disorder that involves a persistent feeling of sadness and loss of interest. It is a serious mental illness, and it negatively affects the entire body [1]. Most of the common symptoms of depression are feelings of sad and irritation, loss of interest in everything, hopeless, difficulty in sleeping and suicidal thoughts [2]. Based on the severity level of symptoms, depression can be classified as mild, moderate, or severe. The causes of depression can be categorized as bullying, relationships, family issues, occupation and financial instability [3]. According to the World Health Organization (WHO), approximately 280 million of people suffer from depression. It is equivalent to 3.8% of the total global population which includes 5.0% among adults and 5.7% among adults older than 60 years [2].

In Sri Lanka, 4.1% of total population suffers from depression which includes mostly age group of 20 years and over [4]. Due to dissatisfaction, helplessness, and pessimism after retirement, depression affects the majority of senior people in Sri Lanka [5]. Also, university students also have a greater risk of depression than the general population [6]. During the COVID pandemic, a greater percentage of persons had depressive symptoms due to their concern of contracting COVID-19 [7]. Because many individuals were financially and emotionally impacted by the epidemic, some people may experience post-pandemic depression. Furthermore, the economic crisis has become a key factor for a considerable increase in depression rates among people due to dissatisfaction with the government [8]. People are still failing to recognize depression at an early stage due to a lack of understanding and social stigma. As a result, several suicides and self-harm occurrences have been documented throughout the years [4]. Therefore, depression is a major issue in Sri Lanka.

B. Depression and Social Media

Due to the internet's explosive expansion and widespread use of mobile devices, everyone in the modern era uses social media platform to express and share their feelings, thoughts, mood, emotions, and interests with others through posts, comments, and stories. This means that people openly discuss their mental health and problems to seek or provide assistance and to combat the stigma associated with mental disorders [9].

Since some of the well-known symptoms of depression can be extracted through the words used by the social media users [9], it can be considered that social media networks have developed into viable tools for researching various mental diseases including depression as well [10]. Moreover, the [11] research states that depression detection using social media data is more effective as it does not involve any physical interactions. This allowed researchers to focus more on early depression detection using social media platforms.

C. Sinhala-English Code-Mixed Language

Mixing two or more languages in a single statement or speech is referred to as code mixing. It is commonly used to improve communication in multilingual groups [12]. On social media platforms such as Facebook, Instagram, Snapchat, and Twitter, many users always combine both their native language and English to communicate [13].

Code mixing can be categorized into intra-sentential code-mixing and inter-sentential code-mixing [14]. Intra-sentential code-mixing occurs when there is no apparent shift in the subject within a sentence, while inter-sentential

Correspondence: Fahd Fazal (E-mail: fahad.2019656@iit.ac.lk) Received: 22-05-2023 Revised: 21-08-2024 Accepted: 10-10-2024

Fahd Fazal is a Computer Science graduate from Informatics Institute of Technology, Sri Lanka. His research interests are NLP and Deep Learning. (e-mail: fahad.2019656@iit.ac.lk).

Cassim Farook is a Senior Lecturer at the School of Computing, Informatics Institute of Technology, Sri Lanka. His current research focuses on text mining and language models. (e-mail: cassim.f@iit.ac.lk).

DOI: https://doi.org/10.4038/icter.v17i3.7282

code-mixing happens when a speaker switches between their native language and a second language to describe an experience [14]. Sri Lankans also engage in code mixing on social media, where they combine their native language, Sinhala, with English, which is commonly referred to as Singlish [15]. Among Sri Lankan social media users, various styles of code mixing are employed, with the most prevalent one being the use of Sinhala words written in the English language (Singlish) [16]. The TABLE I shows different styles of Sinhala-English code mixing with some examples.

TABLE I
DIFFERENT SINGLISH CODE MIXING STYLES

Code-Mixing Style	Example
Sinhala is written in English (Singlish), with some English terms in the sentence.	Mage phone eka denna
Unicode characters with mix of Singlish	Mage දුරකථනය denna
English words + Unicode characters + Singlish	මගේ phone eka දෙන්න

Even though depression detection using English language is a highly researched area, there are only very few studies done in depression detection using Sinhala and other code-mixed languages. Therefore, the authors proposed a machine learning approach to detect depression in Singlish text. The remaining sections of the paper are organized after an explanation of the literature review and relevant works, methodology which includes details of the proposed Machine Learning (ML) architecture. The evaluation outcomes of the employed algorithms and limitations encountered are included in the paper's later section. Then discussion about the prototype implementation and evaluation. Finally, the paper concludes with a summary that explains the overall research and its potential future improvements.

II. LITERATURE REVIEW

Since depression detection using social media data has been increasingly popular among researchers in recent years, numerous attempts have been made to detect depression in its earliest stages. These existing works can be broadly categorized into 2 main categories such as depression detection using machine learning approaches and depression detection using deep learning approaches. As this research involves Sinhala-English code-mixed text data and depression detection research works have not been done in Sinhala-English code-mixed text data, the authors will discuss existing depression detection works done in English, Sinhala and other code-mixed languages under this section. Additionally, other detection works using Singlish text data will be discussed to understand how Singlish data have been used for text classification purposes.

A. *Depression Detection in English and Other Languages*

H. S. Alsagri and M. Ykhlef [17] focused on depression detection using tweets and Twitter profile activity. This research used a supervised machine learning approach and explored different machine learning classification algorithms. The dataset consists of tweets of depressed and non-depressed users. Numerous pre-processing operations were carried out, such as data preparation and alignment, data labelling, feature extraction, and feature selection. Support Vector Machines (SVM) has achieved a higher accuracy of 82% than other classifiers. Also, due to the variety and depth of its feature set, which this study has added by combining three additional new features with the previously presented ones, it was proved that the proposed model performs better than the previously suggested ones in terms of accuracy. Since this paper used a larger dataset, it is better to explore deep learning techniques.

In the paper [18], different base and ensemble ML classifiers were experimented to detect signs of depression using social media texts. This research utilized two different twitter datasets obtained from previous research. BOW was used to directly extract features from social media messages. It was discovered that Random Forest model outperforms other classifiers in detecting depression from generic texts. It was observed that dynamic sampling could reduce the accuracy of the more populous class while increasing the accuracy of the less populous class in a highly unbalanced dataset. However, this observation should be tested with some more datasets as the behaviour could change with the nature and the size of the dataset.

Another study [19] proposed a ML approach to detect depression of a user using reddit data. The CLEF 2020 dataset from the e-risk lab has been used in this study. With an accuracy of 83%, the Nave Bayes Classifier performed better than the other classifiers employed in this study. Its accuracy is 4% higher than the Support Vector Machine Classifier's and 2% higher than both the Logistic Regression model and the Random Forest Classifier's. In the literature, it was observed that variations of NB classifiers performed well with Facebook and Twitter data as well. Therefore, it can be assumed that NB classifier could be the best to detect depression using social media text data but still it needs to be clarified through proper research.

[20] performed depression detection using Russian text collected from a social network called Vkontakte. It was found that SVM with Psycholinguistic markers achieved 66.40% of F1-score in depression classification of the text. The term "psycholinguistic markers" refers to linguistic elements in a document that represent the author's psychological traits and may also indicate any psychiatric illnesses. Psycholinguistic markers performed well on the data, and term clustering strategies could improve the performance of n-grams models. As a result of the excessively noisy Vkontakte textual data and the relatively poor classification results obtained by tf-idf based models, the technique adopted in this paper may not be the best to employ. Therefore, this approach needs lot of improvements based on the readings such as reducing the noisiness in data and the feature set dictionaries should be rewritten and filtered.

B. Verma et al. [21] introduced a hybrid model using CNN & LSTM deep learning models for depression detection among people using typical conversation-based text data gathered from twitter. The dataset used in this study contains 15000 tweets. In the proposed model, the CNN was used to extract features and LSTM model for the classification of tweets into depressive and non-depressive classes using the extracted features. It was observed that the proposed model achieved an accuracy of 92% in comparison with the Decision Tree classifier that gives the highest accuracy of 83% out of all five classifiers employed. The study stated that in comparison to the machine learning models, deep learning models perform better with textual data. Therefore, this study proved that use of hybrid model cancels the disadvantages of one model by the and performs better.

In the [22] paper, a Bi-LSTM classifier was used to detect depression of individuals using reddit data. Early Detection of Depression in CLEF eRisk 2017 dataset was used to train the depression detection model. The Bi-LSTM classifier was experimented with different combinations of feature sets. It was found that Word2VecEmbed+Meta features performed well with Bi-LSTM classifier. Even though the proposed model performs well in terms of accuracy, the time taken to detect depression of a user is very high. This limitation should be addressed in order to consider the proposed approach at an implementation level.

[23] pioneered a multitask learning-based strategy for predicting depressed Sina Weibo users. A large Weibo User depression detection dataset - WU3D which contains around 2,191,910 of data was created and published in this study. The proposed multitask learning Deep Neural Network (DNN) classifier named FusionNet was used to handle both the statistical feature classification problem and the word vector classification task at once because some statistical features like social behaviours and picture-based features were utilized with text features to improve the quality of the detection result. When compared to the widely used models in existing works, FusionNet has significantly outperformed them with an F1-Score of 0.9772. Therefore, when dealing with several classification tasks concurrently, it has demonstrated to be the best classification model. However, this proposed FusionNet model should be tested with data from other social media platforms to ensure its performance on different datasets.

Another study [24] proposed a productive model by implementing a deep RNN-LSTM model to detect depression in text using twitter data. The dataset was taken from Kaggle which contains more than 4000 tweets. A robust one-hot and PCA technique was used for data pre-processing and feature extraction with other NLP techniques as well. This approach was not experimented in previous studies but an accuracy of 99% was attained with a reduced false positive rate. The proposed solution was evaluated by comparing it with traditional ML classifiers where most of them achieved accuracy more than 90%. This brings the question that whether the dataset is highly imbalanced where very less non-depressing tweets are present. Therefore, the proposed approach cannot be considered as good until it is tested with a new balanced dataset.

[25] analyses the efficiency of the BERT (Bidirectional Encoder Representations from Transformers) approach for diagnosing depression via tweet analysis. This study utilized a dataset of 3,867 tweets in Indonesian. The BERT model accurately detected tweets related to depression, with an accuracy of 71.76%. It was found that BERT can handle lengthy texts successfully and outperforms older approaches like CNN. However, the dataset may be biased due to the demographic and linguistic features of users. Proper validation with larger and more varied datasets is required to generalize these findings.

Suzan Elmajali and Irfan Ahmad [26] presents a novel approach for detecting depression symptoms in Arabic tweets using pretrained transformer models, specifically AraBERT and MARBERT. A custom dataset of 1,290 Arabic tweets supplemented with data generated by ChatGPT were used to classify tweets into 10 categories. This is the first study to focus on symptom recognition rather than simply identifying depression as a binary state. AraBERT achieved impressive results with accuracy of 99.3%, while MARBERT showed slightly lower accuracy at 98.3%. Even though data augmentation techniques were employed to address potential biases in the dataset, it was limited to short tweets without emojis, which could impact the model's performance. Additionally, the lack of historical tweet data for users poses a challenge for tracking symptoms over time. Therefore, as the dataset is small further research needs to be performed with large datasets as these results have higher chances of being data biased.

B. Depression Detection in Sinhala Language

In the research [4], a supervised learning approach is used to detect Sinhala depressive posts in twitter. KNN classifier was used to build the depression detection system. The dataset used in this study consists of 1005 Sinhala texts collected from Facebook and Twitter. It was observed that KNN algorithm has the best performance (70% accuracy) on depression detection using Sinhala text data out of all the algorithms such as SVM, Multinomial Naïve Bayes (MNB), Random Forest and Decision Tree. According to the experiment results TF-IDF vectorizer had better results for each algorithm than the Counter Vectorizer. However, the proposed approach can be improved a lot in terms of its performance by increasing the size of the dataset and applying pre-processing techniques such as stop words removal and stemming.

C. Depression Detection in Other Code-Mixed Languages

The study [27] introduced a classifier model using TF-IDF and multinomial Naive Bayes to detect depression of tweets in Hindi-English code-mixed language. The classifier model uses a supervised machine learning approach. 670 raw data were collected from Twitter to train the classifier model. The proposed model has achieved an accuracy of 96.15 and comparatively higher than other models evaluated in the paper such as Char-LSTM, Subword-LSTM, and CNN-BiLSTM. It is observed that only this paper has focused on depression detection using code mixed data and their proposed model has performed well in the classification. Also, it can be concluded that multinomial naive bayes works better than traditional naive bayes through this work. But as the size of the dataset being comparatively small which is less than 1000, it is better to perform cross validation with other datasets to evaluate its performance even more.

D. Other Detection Works using Sinhala-English Code-Mixed Text Data

O. Liyanage and K. Jayakumar [28] focused on hate speech detection in Sinhala-English mixed language using ensemble supervised ML approaches. The dataset used in this study was collected from YouTube and Facebook which contains 1375 comments. Eight base ML classifiers and three ensemble classifiers were experimented with different ranges of character and word n-grams where five of them have not been explored in the previous studies of this domain. It was observed that the hard voting classifier with character n-grams outperformed all the other classifiers with an accuracy of 84%. This kind of approach was not seen in depression detection studies that used machine learning approaches which makes the proposed approach to be explored in the depression detection domain as well. However, a contradiction appears as the author had stated that stop words removal had a negative impact on the accuracy but the research (Hettiarachchi, Weerasinghe and Pushpanda, 2020) has achieved a good accuracy with the application of stop words removal. Therefore, this study can be improved by applying stop words removal and increasing the size of the dataset to try out deep learning approaches.

Another study [29] focused on a language detection model to detect Sinhala and English words in code-mixed data using XGB classifier and a CRF model for sequence labeling. There were no suitable datasets to be used because this was the first study on Sinhala-English code-mixing analysis at the time this paper was produced. Consequently, a new dataset was built using Facebook data to carry out this research which contains 7,500 sentences. It was observed that XGB classifier outperformed with an accuracy of 92.1% with bigram features when compared with other ML and DL models trained with different n-grams and bag of words. This study stated that tree-based model with n-gram methods performs better in Singlish-English code-mixed text classification tasks. So, the author believes that this can be explored in the depression detection domain as well to get better results. Moreover, hybrid deep learning models can be explored to improve the performance of the language detection model.

[30] proposed a comprehensive approach to detecting hate speech in Sinhala-English code-mixed text data which employs a range of machine learning and deep learning models, including transformers. BERT and GPT-2 demonstrated the best performance out of all models by achieving accuracies of 82% and 80% respectively. This indicates that transformer models are highly effective in handling the complexities of code-mixed languages. However, the study could be improved by diversifying data sources, addressing annotation biases, delving deeper into ethical problems, and investigating methods for dealing with class imbalance.

In summary, majority of the research used supervised machine learning algorithms to detect depression from social media texts and had small datasets, however the suggested models performed well in depression detection. As a result, supervised machine learning algorithms outperform deep learning approaches for small datasets. Furthermore, the authors noticed that emojis were not included in depression detection by the researchers since emojis can differ from the emotions represented in the text, which might be difficult to handle. Also, existing studies shows that deep learning algorithms require a huge amount of training data to perform well. According to several research, the time required to identify depression using deep learning models such as LSTM is longer due to its feed forward and backward propagation characteristics.

Previous studies in the domain of Sinhala-English code-mixed data collected and created their own datasets through web scraping or surveys since there are no publicly available datasets for specific tasks. Same words having different spellings is one of the main issues stated in the literature when handling Singlish text data because different people code mix in different ways [31].

III. Methodology

As this study is focused on detecting depression in Singlish textual contents using social media data, the author experimented with different base and ensemble ML algorithms to find the best performing algorithm for the specific depression detection task. The methodology consists of main steps such as dataset collection, data pre-processing, feature extraction and classification as shown in the Fig 1.

A. Data Collection and Annotation

The authors had to manually gather data and construct an own dataset because there is no publicly available dataset for depression detection in Sinhala-English code-mixed language. So, twitter API was used to extract tweets by using set of keywords which are mostly used by depressive people such as “dukai”, “epa wela”, “marenna hithenawa” and “kalakirila”. Due to the reason that same Singlish word can have different spellings as combinations of some English letters produce the same sound, different variations of same Singlish words were used during keyword searching.

Since tweets were in limited numbers, the author had to collect Singlish texts from Facebook by using Facebook Scraper which is a python library created based on the Facebook API and by manually searching in psychology groups on Facebook. The author was able to annotate the collected total number of Singlish texts with the help of a counsellor psychologist. All the tweets and posts collected were publicly accessible and the identity of the user who posted the text was ignored due to privacy issues.

Even though there are different styles of Sinhala-English code-mixing as mentioned in TABLE I, the dataset only consists of 1177 data records where Sinhala words written in English with some English terms in the sentence as shown in TABLE II. The label column is to label the Singlish texts as non-depressive or depressive. ‘1’ denotes the ‘depressive texts while ‘0’ denotes the non-depressive texts.

TABLE III
Sample Records of the Dataset

Text	Label
mata manasika gataluwak thiyenawa. Please katha krnna puluwan kenek innawada	1
Life eke hithana dewalma wenne natha	0

B. Data Pre-processing

Raw data collected from social media should be pre-processed by removing the unnecessary and noisy elements in the data to extract features from it [32]. Different Natural Language Processing (NLP) techniques are used for pre-processing in text classification tasks. Tokenization, stop words removal, and removal of capitalization, punctuation, special characters, non-alphabetic characters, numbers, HTML tags, links and emojis were followed as pre-processing steps.

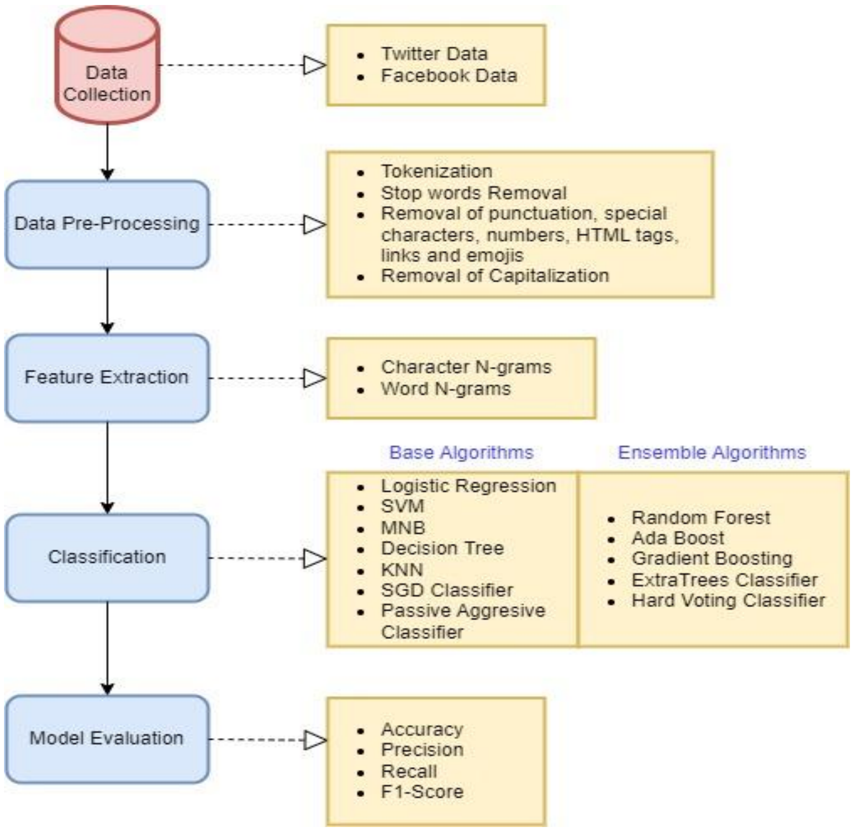

Fig. 1 Proposed Approach Architecture

Even though stemming and lemmatization are widely used for pre-processing, experts in Singlish domain suggest that it is highly difficulty to perform them on a Singlish dataset. Therefore, stemming and lemmatization were ignored in this study.

1) Tokenization:

Tokenization is the process of converting a stream of text into tokens. As most languages contain unique tokens that must be recognized as words, the tokenizer must be adapted to the application's requirements [33]. Since the Singlish text data contains English letters, tokenization may be accomplished quickly and efficiently using accessible Python ML modules. It was discovered that the majority of current works used tokenization at the pre-processing step.

2) Stop-words Removal:

Stop-words are often occurring terms that do not add any sense to the data [34]. Stop-word removal is an important pre-processing technique that increases the efficiency of feature extraction while decreasing the dimensionality of the dataset [31]. Moreover, for stop words removal, custom Singlish stop words list was created using Sinhala and English stop words where Sinhala words were translated into Singlish and common variations of them were inserted into the Singlish stop words list.

3) Removal of Special Characters and Emojis:

Noise in social media text data includes user references (such as @john), numerals, punctuation, special characters, and hashtags, which should be eliminated for appropriate feature extraction [35]. Although emojis may convey the meaning and emotion of a text, social media users have a habit of combining emoticons in ways that contradict the intention and emotion of the text. Emojis are therefore regarded as noise in the data and are removed during pre-processing [36].

4) Removal of Non-alphabetical characters and Capitalization:

Since the Singlish data contains only English letters, it is necessary to remove non-alphabetical characters as they are irrelevant to the context and can affect the performance of the ML model negatively. Also, the Singlish texts should be converted into lowercase to remove capitalization to avoid using alternative word forms. [35]

C. Feature Extraction

Feature extraction is an important step to be performed before classification as it involves transforming text data into a numerical representation, known as features, that can be used as inputs to machine learning models. Bag of Words (BoW), word n-grams, character n-grams, Term Frequency and Inverse Document Frequency (TF-IDF) and word embeddings are some of the features used for feature extraction in the previous studies of depression detection domain.

1) Bag of Words:

This is the simplest and most direct method for identifying features in texts. Bag of Words (BOW) tracks how frequently words appear in the text [37]. Several studies used BOW and achieved good accuracy in detecting depression using social media texts study. However, BOW has some limitations, including the fact that it disregards grammatical relationships and word sequence [32].

2) Word N-grams:

Word n-grams are continuous sequences of n words retrieved from a text, where n is an integer denoting the number of consecutive words in the sequence [38]. Word n-grams are preferable than BOWs because ML algorithms favour organized words and it is better at creating structured fixed size vectors [4]. Since word n-grams are sensitive to out-of-vocabulary words (OOV), an OOV word may not be correctly represented in an n-gram or may be totally eliminated.

3) Character N-grams:

Character n-grams are extracted continuous sequences of n characters from a text, where n is a number indicating the length of the sequence in characters. Character n-grams can detect patterns in texts with a high rate of misspelled words. Because of this, character n-grams fared well in studies that employed Singlish data for detection.

The semantic meaning of words is not captured by character n-grams. As a result, their capacity to do tasks like semantic parsing or sentiment analysis, which need a detailed understanding of word meanings or entailment connections, may be constrained. [39]

4) TF-IDF:

TF-IDF reveals subjects based on the correlation relevance of words in documents without taking into account the context or synonyms. It measures by taking two metrics into account where Term frequency (TF) measures how frequently a word appears in a publication and Inverse document frequency (IDF) calculates a word's frequency within a corpus. TF-IDF value is the result of multiplying TF by IDF. [40]

Since TF-IDF is widely used in text classification, it has also been successful in depression detection domain. Some of the major drawbacks of TF-IDF are lack of semantic understanding and ignorance of word sequence and context [32].

5) Word Embeddings:

Word embeddings represent each word as a real-valued vector in a condensed space, capturing the inter-word semantics. Each word is represented as a real-valued vector of tens to hundreds of dimensions. Word2vec, FastText, Elmo, GloveEmbed are some of the word embeddings used for depression classification. [32]

However, based on an experiment done by the authors, it was observed that ML algorithms perform well with unigram to five-gram character n-gram features than word n-gram features. These features should be extracted from the dataset and vectorized using a vectorizer. In the studies [27], [4] and [28], it was proved that TF-IDF vectorizer works better than Count vectorizer with ML algorithms.

Therefore, TF-IDF vectorizer in the scikit-learn package was used to extract and vectorize character n-gram features in this study. Since the vectorizer performs the task of an encoder where it converts text into a sparse matrix of TF-IDF scores representing each word or n-gram as a floating-point value [41], the authors have not implemented separate encoding or embedding techniques along with the vectorizer before feeding the data into the ML model.

.

D. Classification

Due to the reason that deep learning approaches requires larger dataset for training, seven base and five ensemble supervised ML algorithms were considered in this study.

1) Base Machine Learning Algorithms:

Base machine learning algorithms are essential approaches utilized in a variety of data analysis and prediction tasks. They serve as the foundation for more sophisticated algorithms and methodologies such as gradient boosting, deep learning, and reinforcement learning. Each algorithm has its own set of advantages, disadvantages, and areas of application.

Most of the studies experimented base machine learning algorithms for the depression detection task. Therefore, Logistic Regression, SVM, Multinomial Naïve Bayes, KNN, Decision Tree, SGD Classifier and Passive Aggressive Classifier are the base ML algorithms considered in this study.

2) Ensemble Machine Learning Algorithms:

Ensemble learning is a machine learning approach that mixes numerous base models to create predictions. Ensemble approaches exploit the strengths of numerous models while reducing the impact of individual model flaws, resulting in increased performance, robustness, and generalization. However, it is crucial to highlight that ensemble approaches may increase computing complexity and model interpretability when compared to utilizing a single model.

Use of ensemble ML algorithms is less when compared to the use of base ML algorithms for depression detection using social media data. But ensemble algorithms can be very effective in depression detection due to their capacity to identify varied patterns while reducing the influence of individual model biases. They can successfully deal with the complexity and variety of depression-related data. Therefore, the authors considered ensemble ML algorithms such as Random Forest (RF), Ada Boost, Gradient Boosting, Hard voting classifier and ExtraTreesClassifier in this study.

3) Hyper-parameter Tuning:

Hyper-parameter tuning is one of the crucial stages in building machine learning solutions [42]. The process of determining the ideal values for the hyperparameters of a machine learning algorithm in order to maximize its performance on a given task or dataset is known as hyperparameter tuning [43]. Hyperparameters are configuration settings that are not learnt from data but are established before the training process. They have an impact on the model's behaviour and performance [42]. Most commonly used hyper-parameter tuning methods are manual search, grid search, random search, and Bayesian Optimization [43]. Hyperparameter tuning can be a time-consuming and resource-consuming process when the search space is large [44].

In order to find the best parameters for each algorithm utilized in this study, hyper-parameter tuning was performed using grid search with five-fold cross validation which estimates the accuracy of machine learning models by splitting the given dataset into five folds. Grid search was selected over other methods as it checks for all possible combinations of hyper-parameter values from the defined search space. All the algorithms were implemented using Scikit Learn and results are discussed in section IV.

IV. Results and Discussion

All the algorithms experimented in this study were selected based on an analysis done similar to the TABLE III which summarizes the dataset sizes, models used, and results from various studies on depression detection. Performance of these algorithms were evaluated using classification metrics such as accuracy, precision, recall and f1-score. Each algorithm was trained with unigram to five-gram character n-gram features using 80% of the created dataset and was tested using 20% of the created dataset.

According to the performance results shown in TABLE IV, it can be observed that the implemented ExtraTreesClassifier algorithm with character n-gram features performs better than other ML algorithms used as it has achieved higher accuracy of 79.13% and f1-score of 0.71 out of all.

Since false positive and negative values are quite low in the confusion matrix shown in Fig 2, good prediction results can be expected from the ExtraTreesClassifier but it is always better to have false negative values lesser than false positive values as false negative values has high impact on the classification.

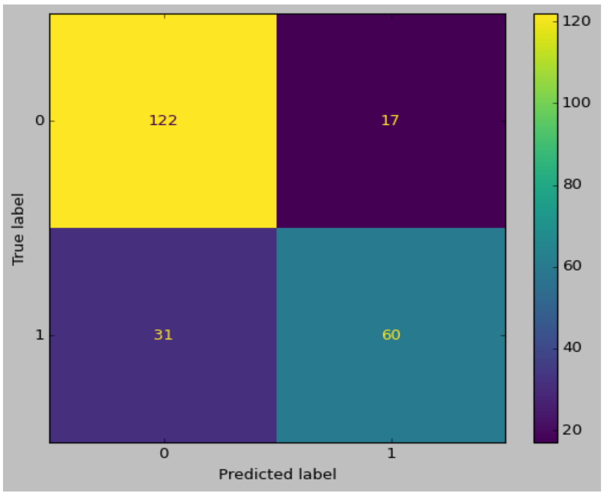

Fig. 2 Confusion Matrix of the ExtraTreesClassifier

TABLE IIIII
Comparative Analysis of Dataset Sizes, Data Types, and Model Performance Across Studies

Study	Dataset Details	Dataset Size	Model/s Used	Results/Findings
[4]	Sinhala only texts collected from Facebook and Twitter	1005	KNN	70% accuracy, TF-IDF vectorizer performs better than Counter Vectorizer
[27]	Hindi-English code-mixed tweets	670	MNB	High accuracy of 96.15 than traditional NB, Char-LSTM, Subword LSTM, and CNN-BiLSTM,
[28]	Sinhala-English mix comments from Facebook and YouTube	1375	Hard Voting Classifier	Performed better than 13 models with accuracy of 84%, No stop words removal used
[29]	Sentences collected from Facebook that contains Sinhala and English words	7500	XGB classifier	92.1% accuracy, tree-based model with n-gram methods performs better in Singlish-English code-mixed text
[30]	Facebook comments in Sinhala-English code-mixed content	2,205	BERT	High accuracy of 82% than GPT-2 and ROC value of 90%
[17]	English tweets of depressed and non-depressed users	300,000+	SVM Classifier	Higher accuracy of 82% than other classifiers, introduced three additional new features
[19]	Texts collected from reddit (CLEF 2020)	763	NB Classifier	Gained best accuracy of 83%
[21]	Conversation-based English tweets	15,000	CNN + LSTM	Achieved best accuracy of 92%
[25]	Indonesian tweets	3867	BERT	71.76% of accuracy than traditional CNN
[26]	Arabic Tweets	1290	AraBERT & MABERT	AraBERT gained 99.3% accuracy and MARBERT gained 98.3% accuracy

TABLE IVV
PERFORMANCE RESULTS OF THE ML ALGORITHMS

Machine Learning Algorithm	Accuracy	Precision	Recall	F1-score
Logistic Regression	0.74	0.71	0.58	0.64
Support Vector Machines (SVM)	0.75	0.73	0.60	0.66
Multinomial Naïve Bayes	0.73	0.66	0.67	0.66
Decision Tree	0.57	0.46	0.48	0.47
Random Forest	0.73	0.68	0.59	0.63
K-Nearest Neighbor (KNN)	0.66	0.57	0.67	0.61
SGD Classifier	0.66	0.60	0.49	0.54
Passive Aggressive Classifier	0.71	0.63	0.68	0.65
Ada Boost	0.68	0.59	0.63	0.61
Gradient Boosting	0.70	0.63	0.57	0.60
ExtraTreesClassifier	**0.79**	**0.77**	**0.65**	**0.71**
Hard voting classifier	0.74	0.72	0.59	0.65

In general, a larger AUC value implies a stronger classifier, with a classifier that has an AUC close to 1.0 having good discriminating power and a classifier that has an AUC near to 0.5 having weak discrimination power [45]. According to the Fig 3, since the AUC value is 0.80 which is closer to 1.0, it can be concluded that the ExtraTreesClassifier algorithm performs well in classifying Singlish text into depressive or non-depressive classes.

In order to validate the achieved accuracy of ExtraTreesClassifier, the authors manual inserted several inputs and tested whether expected results were shown by the algorithm. In most cases, it was able to properly distinguish between depressive and non-depressive text inputs. Additionally for depressive texts, rough depression rate along with the prediction was shown which is calculated by averaging the class probabilities predicted by each estimator, using specified weights then the highest averaged probability is selected as the rate value. This is displayed in the TABLE V where the output value 1 means input is identified as depressive text and the output value 0 means non-depressive text.

TABLE V
SAMPLE INPUTS AND THEIR OUTPUTS

Input Text	Output
mata jeewithe epa wela...mata marenda hithenawa friends	1, 89%
mang loku manasika peedanayaka inne. eka parasnayak iwara unama tawa prasnayak enawa....life eka mata epa wela	1, 75%
mama sathutin innawa shape eke. kisi awlak naha	0
mata samaweyalla waradiyata hithanna epa	0

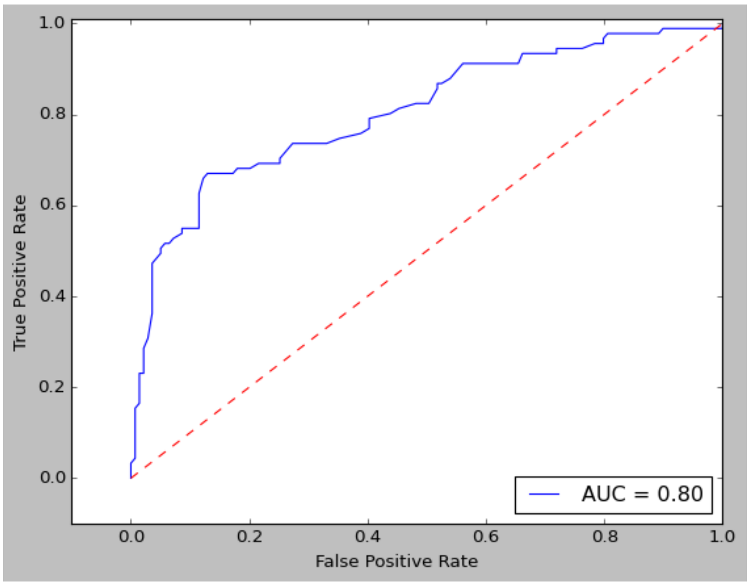

Fig. 3 AUC-ROC Curve of ExtraTreesClassifier

Although authors considered only ML models, it was identified through the analysis in the TABLE III that the BERT transformer model has performed well with limited data despite being a DL model. Therefore, authors trained a BERT model later in this study with the same Singlish dataset where it was split into 80% for training and 20% for testing. The BERT model achieved accuracies of 64% and 41% with and without fine-tuning respectively. Moreover, evaluation loss is around 0.7034 which suggests that the model is able to make most of the predictions accurately. However, still ML models are able to outperform the BERT transformer model as shown in TABLE IV.

One of the main limitations of this research is that a smaller dataset with 1177 Singlish texts was used to train the depression classification model which is very low compared to other datasets used in existing depression works which have more than 10,000 texts. This was caused due to unavailability of public Singlish datasets for depression detection purposes. The intention behind the text has a big impact on the depression detection because sometimes people does not actual mean what they write. This was also a limitation of this research as depression detection model was unable to analyse the intention of a complete Singlish text when detecting the depression. As explained in TABLE I, although there are different styles of Singlish code-mixing used by people on social media, only Singlish texts which contains Sinhala words written using English letters with a mix of few English words were considered for the depression detection task because dataset only contained this specific style of Singlish.

V. Prototype Implementation and Evaluation

As per the results, it was found that ExtraTreesClassifier algorithm works better than other base and ensemble ML algorithms for depression detection in Singlish. Therefore, using this approach, the authors developed a web application as a prototype for depression detection in Singlish textual content. The web application was designed based on the API architecture. The API backend of the system was developed using python Flask and frontend was developed using Bootstrap and jQuery.

Quantitative evaluation of the prototype was performed through model evaluation in section IV. Since depression detection is a sensitive task, it is required to evaluate the effectiveness of the prototype using a qualitative approach. So, the qualitative evaluation was carried out by selecting mental health professionals and ML experts as evaluators.

During the evaluation process, evaluators were presented with a demo of the prototype and allowed to use the prototype which was hosted in Azure cloud. Finally, the evaluation feedback was given based on the criteria listed in the TABLE VI and it was analysed using thematic analysis.

TABLE VI
Qualitative Evaluation Criteria

Criteria	Evaluation Purpose
Overall concept of the research project	To get a review on the whole idea behind this research project
Scope of the project	To validate whether the scope covered in this research project is acceptable for a depression detection system
Solution and prototype	To validate the effectiveness of the proposed solution and whether the prototype represents the overall concept of the project
Performance of the prototype	To validate the performance of the developed system in terms of quality of the outputs
Limitations and Possible Improvements	To identify the future enhancements and the drawbacks of the depression detection system
Usability and UI/UX	To ensure that the product created for demonstration is user-friendly

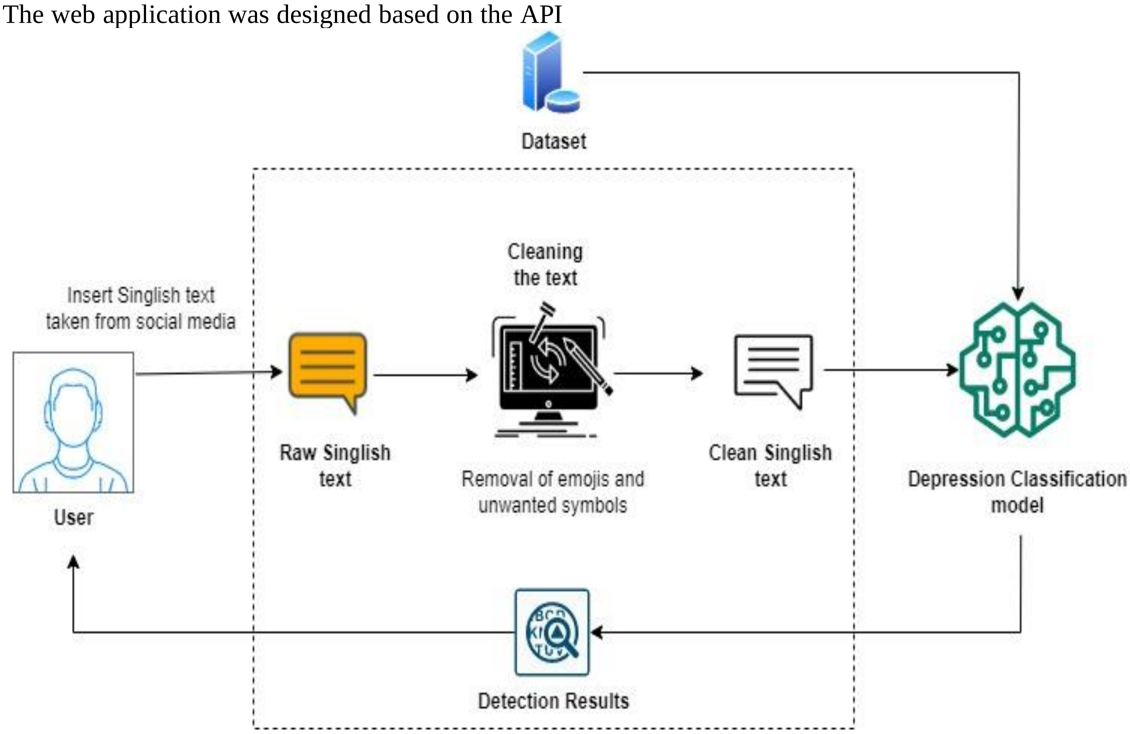

Fig. 4 Prototype Diagram

According to the qualitative evaluation results, the solution provided by utilizing ensemble machine learning algorithm is good and effective on a certain level because diagnosing process of clinical depression should happen with a diagnostic criterion. The prototype detects depression properly for most scenarios of Singlish texts but in some tricky scenarios it fails to detect properly. Overall performance is good as it has less processing time and decent accuracy. Finally, it was identified this will aid professionals to easily detect depressive people in psychological groups on social media through their posts which will save lives of numerous people.

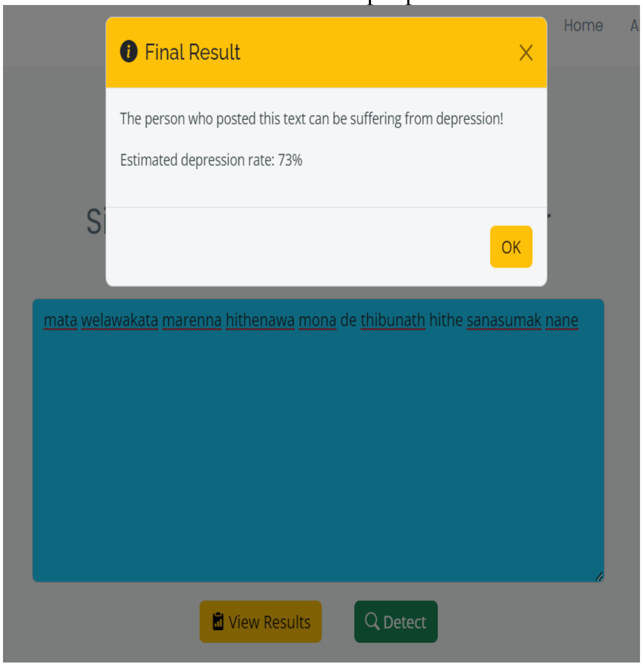

Fig. 5 User Interface of the Prototype

VI. Conclusion and Future Work

This study has proposed a machine learning approach for detecting depression in Sinhala-English code-mixed language. A dataset was created by gathering tweets and posts from Twitter and Facebook which was annotated with the help of a counsellor psychologist. In data pre-processing, a custom Singlish stop words list was created using Sinhala and English stop words for stop words removal task. The implemented ExtraTreesClassifier outperformed all the other classifiers with an accuracy of 79.13% and f1-score of 0.71. Unigram to five-gram character n-gram features performed well with machine learning algorithms in depression detection in Sinhala-English code-mixed language.

As future work, the size of the dataset can be increased with a significant amount which will eventually improve the accuracy of the depression classification model. Since only machine learning approaches were experimented, different deep learning approaches can be explored for better performance, but this requires a dataset with at least 10,000 training data otherwise the deep learning approach will not produce better results. Consideration of different elements in a social media textual content are important for a depression detection task. Emojis are one such element which depicts the overall emotion behind the text but these emojis can be misused with different combinations of emojis as they might contradict with the actual meaning of the text. Another element is the posted time of the Singlish text, this will help to identify the period that the social media user is having depression feelings. This may also help to recognize the severity of depression.

Another major future enhancement would be classifying the detected depression into different levels from mild to severe based on the severity of the symptoms which will help mental healthcare professionals to easily identify what cure should be given to the depressed individual. Use of a spell checker in the depression classification model to overcome the limitations mentioned regarding the different spellings of the same Singlish word. Moreover, the depression detection system can be integrated into social media platforms which will allow end users to easily interact with the system and depression detection system can be upgraded into an interactive chatbot which will also provide guidance to overcome depression.

Acknowledgment

Authors would like to give special thanks to Mrs. Arosha Dasanayaka and Mrs. Eshama Kariyawarna for sharing their domain knowledge and providing support to make this research a success. Also sincere thanks to all the evaluators for allocating their precious time to give their feedback on this research project.

References

[1] F. Torres, “What Is Depression?,” psychiatry.org. Accessed: Oct. 04, 2022. [Online]. Available: https://psychiatry.org:443/patients-families/depression/what-is-depression

[2] “Depression,” World Health Organization. Accessed: Oct. 04, 2022. [Online]. Available: https://www.who.int/news-room/fact-sheets/detail/depression

[3] N. F. Abd Yusof, C. Lin, and F. Guerin, “Analysing the Causes of Depressed Mood from Depression Vulnerable Individuals,” in *Proceedings of the International Workshop on Digital Disease Detection using Social Media 2017 (DDDSM-2017)*, Taipei, Taiwan: Association for Computational Linguistics, Nov. 2017, pp. 9–17. Accessed: Oct. 05, 2022. [Online]. Available: https://aclanthology.org/W17-5802

[4] L. Rathnayake and I. A. N. Arachchige, “Supervised Learning Approach for Detection of Sinhala Depressive Posts based on Twitter,” in *2021 21st International Conference on Advances in ICT for Emerging Regions (ICter)*, Dec. 2021, pp. 111–116. doi: 10.1109/ICter53630.2021.9774819.

[5] L. I. Ekanayake and A. Bulathwatta, “An exploratory study on factors affect on Elderly Depression in Sri Lanka,” *IJRISS*, vol. 06, no. 04, pp. 197–206, 2022, doi: 10.47772/IJRISS.2022.6414.

[6] S. D. Amarasuriya, A. F. Jorm, and N. J. Reavley, “Predicting intentions to seek help for depression among undergraduates in Sri Lanka,” *BMC Psychiatry*, vol. 18, no. 1, p. 122, May 2018, doi: 10.1186/s12888-018-1700-4.

[7] B. Perera, B. Wickramarachchi, C. Samanmalie, and M. Hettiarachchi, “Psychological experiences of healthcare professionals in Sri Lanka during COVID-19,” *BMC Psychology*, vol. 9, no. 1, p. 49, Mar. 2021, doi: 10.1186/s40359-021-00526-5.

[8] The Lancet Regional Health – Southeast Asia, “Sri Lanka - Health in the middle of a crisis,” *The Lancet Regional Health - Southeast Asia*, vol. 4, p. 100067, Sep. 2022, doi: 10.1016/j.lansea.2022.100067.

[9] R. Martínez-Castaño, J. C. Pichel, and D. E. Losada, “A Big Data Platform for Real Time Analysis of Signs of Depression in Social Media,” *IJERPH*, vol. 17, no. 13, p. 4752, Jul. 2020, doi: 10.3390/ijerph17134752.

[10] E. A. Ríssola, S. A. Bahrainian, and F. Crestani, “A Dataset for Research on Depression in Social Media,” in *Proceedings of the 28th ACM Conference on User Modeling, Adaptation and Personalization*, in UMAP ’20. New York, NY, USA: Association for Computing Machinery, Jul. 2020, pp. 338–342. doi: 10.1145/3340631.3394879.

[11] S. Kayalvizhi and D. Thenmozhi, "Data set creation and empirical analysis for detecting signs of depression from social media postings," *ArXiv*, Feb. 2022, Accessed: Apr. 12, 2023. [Online]. Available: https://www.semanticscholar.org/paper/fb1cefdcc6d630c2fc38a78f5f5d39cd0ca29902

[12] D. R. Mabule, "What is this? Is It Code Switching, Code Mixing or Language Alternating?," *JESR*, Jan. 2015, doi: 10.5901/jesr.2015.v5n1p339.

[13] N. Tarihoran, E. Fachriyah, Tressyalina, and I. R. Sumirat, "The Impact of Social Media on the Use of Code Mixing by Generation Z," *International Journal of Interactive Mobile Technologies (iJIM)*, vol. 16, no. 07, Art. no. 07, Apr. 2022, doi: 10.3991/ijim.v16i07.27659.

[14] K. Shanmugalingam and S. Sumathipala, "Language identification at word level in Sinhala-English code-mixed social media text," in *2019 International Research Conference on Smart Computing and Systems Engineering (SCSE)*, Mar. 2019, pp. 113–118. doi: 10.23919/SCSE.2019.8842795.

[15] L. de Silva and S. Ahangama, "Singlish to Sinhala Transliteration using Rule-based Approach," in *2021 IEEE 16th International Conference on Industrial and Information Systems (ICIIS)*, Dec. 2021, pp. 162–167. doi: 10.1109/ICIIS53135.2021.9660744.

[16] I. Smith and U. Thayasivam, "Sinhala-English Code-Mixed Data Analysis: A Review on Data Collection Process," in *2019 19th International Conference on Advances in ICT for Emerging Regions (ICTer)*, Sep. 2019, pp. 1–6. doi: 10.1109/ICTer48817.2019.9023739.

[17] H. S. Alsagri and M. Ykhlef, "Machine Learning-Based Approach for Depression Detection in Twitter Using Content and Activity Features," *IEICE Trans. Inf. & Syst.*, vol. E103.D, no. 8, pp. 1825–1832, Aug. 2020, doi: 10.1587/transinf.2020EDP7023.

[18] R. Chiong, G. S. Budhi, S. Dhakal, and F. Chiong, "A textual-based featuring approach for depression detection using machine learning classifiers and social media texts," *Computers in Biology and Medicine*, vol. 135, p. 104499, Aug. 2021, doi: 10.1016/j.compbiomed.2021.104499.

[19] S. Samanvitha, A. R. Bindiya, S. Sudhanva, and B. S. Mahanand, "Naïve Bayes Classifier for depression detection using text data," in *2021 5th International Conference on Electrical, Electronics, Communication, Computer Technologies and Optimization Techniques (ICEECCOT)*, Dec. 2021, pp. 418–421. doi: 10.1109/ICEECCOT52851.2021.9708014.

[20] M. Stankevich, A. Latyshev, E. Kuminskaya, I. Smirnov, and O. Grigoriev, "Depression Detection from Social Media Texts," p. 11, 2019.

[21] B. Verma, S. Gupta, and L. Goel, "A Neural Network Based Hybrid Model for Depression Detection in Twitter," in *Advances in Computing and Data Sciences*, M. Singh, P. K. Gupta, V. Tyagi, J. Flusser, T. Ören, and G. Valentino, Eds., in Communications in Computer and Information Science. Singapore: Springer, 2020, pp. 164–175. doi: 10.1007/978-981-15-6634-9_16.

[22] F. M. Shah *et al.*, "Early Depression Detection from Social Network Using Deep Learning Techniques," in *2020 IEEE Region 10 Symposium (TENSYMP)*, Jun. 2020, pp. 823–826. doi: 10.1109/TENSYMP50017.2020.9231008.

[23] Y. Wang, Z. Wang, C. Li, Y. Zhang, and H. Wang, "A Multitask Deep Learning Approach for User Depression Detection on Sina Weibo," Aug. 26, 2020, *arXiv*: arXiv:2008.11708. Accessed: Oct. 09, 2022. [Online]. Available: http://arxiv.org/abs/2008.11708

[24] A. Amanat *et al.*, "Deep Learning for Depression Detection from Textual Data," *Electronics*, vol. 11, no. 5, Art. no. 5, Jan. 2022, doi: 10.3390/electronics11050676.

[25] Fikri Ilham and Warih Maharani, "Analyze Detection Depression In Social Media Twitter Using Bidirectional Encoder Representations from Transformers," *Journal of Information System Research (JOSH)*, vol. 3, no. 4, pp. 476–482, Jul. 2022, doi: 10.47065/josh.v3i4.1885.

[26] Suzan Elmajali and Irfan Ahmad, "Toward Early Detection of Depression: Detecting Depression Symptoms in Arabic Tweets Using Pretrained Transformers," *IEEE Access*, 2024, doi: 10.1109/access.2024.3417821.

[27] C. M J, R. Subburam, D. Venkatesan, A. K, and A. Ganeshan, "Linguistic Analysis of Hindi-English Mixed Tweets for Depression Detection," *Journal of Mathematics*, vol. 2022, pp. 1–7, Apr. 2022, doi: 10.1155/2022/3225920.

[28] O. Liyanage and K. Jayakumar, "Hate Speech Detection in Sinhala-English Code-Mixed Language," in *2021 21st International Conference on Advances in ICT for Emerging Regions (ICter)*, Dec. 2021, pp. 225–230. doi: 10.1109/ICter53630.2021.9774816.

[29] I. Smith and U. Thayasivam, "Language Detection in Sinhala-English Code-mixed Data," in *2019 International Conference on Asian Language Processing (IALP)*, Nov. 2019, pp. 228–233. doi: 10.1109/IALP48816.2019.9037680.

[30] M. Muthuthanthri and R. I. Smith, "Hate Speech Detection for Transliterated English and Sinhala Code-Mixed Data," in *2024 4th International Conference on Advanced Research in Computing (ICARC)*, Feb. 2024, pp. 155–160. doi: 10.1109/ICARC61713.2024.10499768.

[31] B. Sharounthan, D. P. Nawinna, and R. De Silva, "Singlish Sentiment Analysis Based Rating For Public Transportation," in *2021 International Conference on Computer Communication and Informatics (ICCCI)*, Jan. 2021, pp. 1–7. doi: 10.1109/ICCCI50826.2021.9402548.

[32] S. R. Begum and S. Y. Sait, "Effective Techniques for Depression Detection on Social Media: A Comprehensive Review," in *2022 International Conference on Computer Communication and Informatics (ICCCI)*, Jan. 2022, pp. 1–9. doi: 10.1109/ICCCI54379.2022.9740754.

[33] S. Koshy and R. Padmajavalli, "Evaluating Techniques for Pre-Processing of Unstructured Text For Text Classification," vol. 6, no. 8, pp. 151–160, Aug. 2018.

[34] E. kah Anoual and I. Zeroual, "The effects of Pre-Processing Techniques on Arabic Text Classification," *International Journal of Advanced Trends in Computer Science and Engineering*, vol. 10, pp. 41–48, Feb. 2021, doi: 10.30534/ijatcse/2021/061012021.

[35] U. Sandamini *et al.*, "A Singlish Supported Post Recommendation Approach for Social Media:," in *Proceedings of the 14th International Conference on Agents and Artificial Intelligence*, Online Streaming, --- Select a Country ---: SCITEPRESS - Science and Technology Publications, 2022, pp. 412–419. doi: 10.5220/0010829700003116.

[36] J. A. D. U. Shalinda and L. Munasinghe, "Hate Words Detection Among Sri Lankan Social Media Text Messages," in *2022 International Research Conference on Smart Computing and Systems Engineering (SCSE)*, Sep. 2022, pp. 55–60. doi: 10.1109/SCSE56529.2022.9905129.

[37] M. Mounika, N. Srinivasa Gupta, and B. Valarmathi, "Detection of Depression Related Posts in Tweets Using Classification Methods – A Comparative Analysis," in *Proceeding of the International Conference on Computer Networks, Big Data and IoT (ICCBI - 2019)*, A. P. Pandian, R. Palanisamy, and K. Ntalianis, Eds., in Lecture Notes on Data Engineering and Communications Technologies. Cham: Springer International Publishing, 2020, pp. 620–630. doi: 10.1007/978-3-030-43192-1_70.

[38] M. Stankevich, V. Isakov, D. Devyatkin, and I. Smirnov, "Feature Engineering for Depression Detection in Social Media:," in *Proceedings of the 7th International Conference on Pattern Recognition Applications and Methods*, Funchal, Madeira, Portugal: SCITEPRESS - Science and Technology Publications, 2018, pp. 426–431. doi: 10.5220/0006598604260431.

[39] M. A. Abdulhayoglu, B. Thijs, and W. Jeuris, "Using character n-grams to match a list of publications to references in bibliographic databases," *Scientometrics*, vol. 109, no. 3, pp. 1525–1546, Dec. 2016, doi: 10.1007/s11192-016-2066-3.

[40] C. Fang, G. Dianatobing, T. Atara, I. S. Edbert, and D. Suhartono, "Feature Extraction Methods for Depression Detection Through Social Media Text," in *2022 6th International Conference on Informatics and Computational Sciences (ICICoS)*, Sep. 2022, pp. 117–121. doi: 10.1109/ICICoS56336.2022.9930596.

[41] J. Ramos, "Using TF-IDF to determine word relevance in document queries," Jan. 2003.

[42] A. Esmaeili, Z. Ghorrati, and E. Matson, "Hierarchical Collaborative Hyper-parameter Tuning," May 11, 2022, *arXiv*: arXiv:2205.05272. Accessed: May 17, 2023. [Online]. Available: http://arxiv.org/abs/2205.05272

[43] E. Elgeldawi, A. Sayed, A. R. Galal, and A. M. Zaki, "Hyperparameter Tuning for Machine Learning Algorithms Used for Arabic Sentiment Analysis," *Informatics*, vol. 8, no. 4, Art. no. 4, Dec. 2021, doi: 10.3390/informatics8040079.

[44] K. Lee and J. Yim, "Hyperparameter Optimization with Neural Network Pruning," May 17, 2022, *arXiv*: arXiv:2205.08695. Accessed: May 17, 2023. [Online]. Available: http://arxiv.org/abs/2205.08695

[45] J. M. Lobo, A. Jiménez-Valverde, and R. Real, "AUC: a misleading measure of the performance of predictive distribution models," *Global Ecology and Biogeography*, vol. 17, no. 2, pp. 145–151, 2008, doi: 10.1111/j.1466-8238.2007.00358.x.