

DEPRESSION DETECTION FROM SOCIAL MEDIA USING MACHINE
LEARNING TECHNIQUES

ABDULLAH AL RAFI

UNIVERSITI TEKNOLOGI MALAYSIA

UNIVERSITI TEKNOLOGI MALAYSIA**DECLARATION OF THESIS / UNDERGRADUATE PROJECT REPORT AND
COPYRIGHT**

Author's full name : Abdullah AL Rafi

Date of Birth : 23 September, 2001

Title : Depression Detection from Social Media using Machine Learning
Techniques

Academic Session : 2023-2024/2

I declare that this thesis is classified as:

☐**CONFIDENTIAL**(Contains confidential information under the
Official Secret Act 1972)*☐**RESTRICTED**(Contains restricted information as specified by the
organization where research was done)*☒**OPEN ACCESS**I agree that my thesis to be published as online
open access (full text)

1. I acknowledged that Universiti Teknologi Malaysia reserves the right as follows:
 - a. The thesis is the property of Universiti Teknologi Malaysia
 - b. The Library of Universiti Teknologi Malaysia has the right to make copies for the purpose of research only.
 - c. The Library has the right to make copies of the thesis for academic exchange.

Certified by:

**SIGNATURE OF STUDENT****SIGNATURE OF SUPERVISOR**

A21EC4006

MATRIX NUMBER

DR. SARINA BINTI SULAIMAN

NAME OF SUPERVISOR

NOTES : If the thesis is CONFIDENTIAL or RESTRICTED, please attach with the letter from the organization with period and reasons for confidentiality or restriction

“I hereby declare that we have read this thesis and in my
opinion this thesis is sufficient in term of scope and quality for the
award of the degree of Bachelor of Computer Science (Software Engineering)”

Signature : _____
Name of Supervisor : DR. SARINA BINTI SULAIMAN
Date : 15 MAY 2024

DEPRESSION DETECTION FROM SOCIAL MEDIA USING MACHINE
LEARNING TECHNIQUES

ABDULLAH AL RAFI

A thesis submitted in fulfilment of the
requirements for the award of the degree of
Bachelor of Computer Science (Software Engineering)

Faculty of Computing
Universiti Teknologi Malaysia

MAY 2024

DECLARATION

I declare that this thesis entitled “*Depression detection from social media using Machine Learning techniques*” is the result of my own research except as cited in the references. The thesis has not been accepted for any degree and is not concurrently submitted in candidature of any other degree.



Signature :

Name : ABDULLAH AL RAFI

Date : 15 MAY 2024

DEDICATION

This thesis is dedicated to my father, who taught me that the best kind of knowledge to have is that which is learned for its own sake. It is also dedicated to my mother, who taught me that even the largest task can be accomplished if it is done one step at a time.

ACKNOWLEDGEMENT

In preparing this thesis, I was in contact with many people, researchers, academicians, and practitioners. They have contributed towards my understanding and thoughts. In particular, I wish to express my sincere appreciation to my main thesis supervisor, Professor Dr. Sarina Binti Sulaiman, for encouragement, guidance, critics and friendship. Without their continued support and interest, this thesis would not have been the same as presented here.

My fellow student should also be recognised for their support. My sincere appreciation also extends to all my colleagues and others who have provided assistance at various occasions. Their views and tips are useful indeed. Unfortunately, it is not possible to list all of them in this limited space. I am grateful to all my family member.

ABSTRACT

Depression is currently thought to be the most frequent mental health issue worldwide and is worsening day by day with social media interactions. The paper presents a closer approach to detecting depression from social media, helping to utilize Recurrent Neural Networks (RNN), Support Vector Machines (SVM), and Naive Bayes (NB) classifiers. The current work surveys the relationship between online behaviour and symptoms of depression. Advanced machine learning techniques are then put to work whereby the text is mined for the emotional tone and linguistic patterns indicative of depression. Sentiment analysis, coupled with NLP techniques using SVM and RNN architecture, places the model in a supreme position for the accurate detection of persons at risk of falling into depression. With labelled datasets from sources such as Kaggle, models can be taught to identify subjects that are between the threshold of depression and non-depression by identifying people based on their social media activities. Many people use social media to gain reassurance and validation. Ironically, Social Media often becomes a source of affirmation, especially for people who are already experiencing depression, as it leads a person to present an idealized version of reality. Everyone is happy in their lives. This perception might further worsen the already low condition of a man, leading to more mental crises where depression worsens and reality is distorted. Further, this model provides easy access to appropriate mental health care for those embarrassed to seek help through normal channels. The other two significant aspects of the deployment strategy, which emanate ethical concerns in data privacy and algorithmic biases, are: the methodology proposed is on how machine learning can be applied to enhance mental health care using the data obtained by social media interactions to predict signs of depression early in a person.

ABSTRAK

Kemurungan kini dianggap sebagai masalah kesihatan mental yang paling kerap di seluruh dunia dan semakin memburuk dari hari ke hari dengan interaksi media sosial. Kertas ini membentangkan pendekatan yang lebih dekat untuk mengesan kemurungan daripada media sosial, membantu memanfaatkan Rangkaian Neural Berulang (RNN), Mesin Vektor Sokongan (SVM), dan pengelas Naive Bayes (NB). Kerja semasa mengkaji hubungan antara tingkah laku dalam talian dan gejala kemurungan. Teknik pembelajaran mesin yang maju kemudian digunakan di mana teks tersebut dilombong untuk nada emosi dan corak linguistik yang menunjukkan kemurungan. Analisis sentimen, bersama dengan teknik NLP menggunakan seni bina SVM dan RNN, meletakkan model dalam kedudukan unggul untuk pengesanan tepat orang yang berisiko jatuh ke dalam kemurungan. Dengan set data berlabel dari sumber seperti Kaggle, model dapat diajar untuk mengenal pasti subjek yang berada di antara ambang kemurungan dan bukan kemurungan dengan mengenal pasti orang berdasarkan aktiviti media sosial mereka. Ramai orang menggunakan media sosial untuk mendapatkan jaminan dan pengesahan. Ironinya, Media Sosial sering menjadi sumber pengesahan, terutama bagi orang yang sudah mengalami kemurungan, kerana ia membawa seseorang untuk menyajikan versi realiti yang ideal. Semua orang gembira dalam hidup mereka. Persepsi ini mungkin memburukkan lagi keadaan seseorang yang sudah rendah, yang menyebabkan lebih banyak krisis mental di mana kemurungan semakin teruk dan realiti menjadi terpesong. Selain itu, model ini menyediakan akses mudah kepada penjagaan kesihatan mental yang sesuai bagi mereka yang malu untuk mendapatkan bantuan melalui saluran biasa. Dua aspek penting lain dari strategi penggunaan, yang menimbulkan kebimbangan etika dalam privasi data dan bias algoritma, adalah: metodologi yang dicadangkan adalah mengenai bagaimana pembelajaran mesin dapat diterapkan untuk meningkatkan penjagaan kesihatan mental menggunakan data yang diperoleh melalui interaksi media sosial untuk meramalkan tanda-tanda kemurungan awal pada seseorang.

TABLE OF CONTENTS

	TITLE	PAGE
	DECLARATION	ii
	DEDICATION	iii
	ACKNOWLEDGEMENT	iv
	ABSTRACT	v
	ABSTRAK	vi
	TABLE OF CONTENTS	vii
	LIST OF TABLES	ix
	LIST OF FIGURES	x
	LIST OF ABBREVIATIONS	xi
	LIST OF SYMBOLS	xii
	LIST OF APPENDICES	xiii
CHAPTER 1	INTRODUCTION	1
1.1	Introduction	1
1.2	Problem Background	2
1.3	Research Aim	2
1.4	Research Question	3
1.5	Research Objectives	4
1.6	Research Scope	5
1.7	Research Contribution	6
1.8	Report Organization	7
CHAPTER 2	LITERATURE REVIEW	9
2.1	Introduction	9
2.2	Problem Formulation	9
2.2.1	Research Domain	9
2.2.2	Description of Related Studies	15

2.3	Machine Learning Approaches for Depression Detection	24
2.4	Chapter Summary	27
CHAPTER 3	RESEARCH METHODOLOGY	29
3.1	Introduction	29
3.2	Operational Framework/Research Workflow	29
3.3	Justification	33
3.4	Performance measurement	35
3.5	Chapter Summary	37
CHAPTER 4	RESEARCH DESIGN AND IMPLEMENTATION	38
4.1	Introduction	38
4.2	Proposed Solution	39
4.3	Experiment Design	41
4.4	Parameter and Testing Method	47
4.5	Chapter Summary	49
REFERENCES		55

LIST OF TABLES

TABLE NO.	TITLE	PAGE
Table 2.1	Comparison of Techniques	14
Table 2.2	Description of related studies	16

LIST OF FIGURES

FIGURE NO.	TITLE	PAGE
Figure 2.1	Simple artificial neural network architecture	24
Figure 3.1	Operational workflow	28
Figure 4.1	data processing workflow	37
Figure 2.2	Basic Structure of Naïve Bayes Approach	25
Figure 4.2	Data Pre-Processing	40
Figure 4.3	SVM Model Evaluation	41
Figure 4.4	RNN Model Evaluation	42
Figure 4.5	NB Evaluation Model	43
Figure 4.6	Cross-Validation and statistical Test	44

LIST OF ABBREVIATIONS

RNN	-	Recurrent Neural Network
SVM	-	Support Vector Machine
NB	-	Naïve Bayes

LIST OF SYMBOLS

LIST OF APPENDICES

APPENDIX	TITLE	PAGE
----------	-------	------

CHAPTER 1

INTRODUCTION

1.1 Overview

Due to stress, worry, and fast-moving modern life, the psychological health of people is seriously affected all over the world. Depression is a serious mental disorder irrespective of any age. It may cause major threats to health, resulting in many disorders like diabetes, high blood pressure, panic attacks, and even the fatal sequelae of heart attack. Since technology evolves, machine learning algorithms nowadays provide new ways for recognizing human emotions. One of the text-based techniques used for sentiment analysis is testing postings and tweets on various social media platforms to determine feelings and moods of users using the platform. Based on these counts, performance can be estimated for several machine learning algorithms like Recurrent Neural Networks (RNN), Support Vector Machines (SVM), and Naive Bayes (NB).

An accurate algorithm will thus be able to identify with a high score whether the sentiment is positive or negative. Emotions can also be drawn from speech, gestures, and facial expressions other than text analysis. The complexity of depression can nowhere be fully put across by the use of conventional clinical diagnosis techniques. However, machine learning methods offer a good substitute in the diagnosis of mental disorders. This is because precise identification and anticipation of depression symptoms make machine learning-based diagnostic methods a very invaluable tool for predictive analyses.

1.2 Problem Background

Because social media use has become ingrained in people's daily lives in the internet age, depression can arise from daily social media interaction. Fortunately, machine learning can help address this issue by monitoring social media activity and analyzing emotional tones. (Babu et al., 2021). The issue of diagnosing depression with outdated, conventional procedures has been increasing. It can, however, identify sadness with the use of machine learning. Social media platforms are extensively used as a means of communication, where the majority of people share their life stories and express their emotions. These platforms contain a lot of content about depressed users in addition to public information. This content can send out sensitive social signals that indicate whether a person is experiencing serious problems, such as self-harm, suicidal thoughts, or plans to commit an illegal act.

Designing an effective model to identify significant depressive symptoms early on that is aided by modern Natural Language Processing (NLP) and machine learning techniques for early depression identification. (Wani, Latif et al., & Shariq, 2023) Stress, worry, and the fast-paced, contemporary lifestyle have had a profound psychological impact on people's brains all around the globe throughout the years. In the field of mental health, machine learning techniques are being used to forecast the likelihood of mental illnesses and, therefore, to carry out possible treatment plans. The many machine learning methods used to identify and diagnose depression are listed in this review study.

The weight of mental health concerns demands an all-encompassing solution. An individual's socioeconomic level would be negatively impacted by depression. Individuals with depression are less likely to interact with others. Psychological therapy and counseling will aid in the battle against depression. Because machine learning algorithms can evaluate large amounts of diverse data and provide effective clinical insights, their applications in the healthcare industry have shown to be pragmatic. Machine learning techniques help mental health professionals make prediction decisions by offering an effective knowledge of mental illnesses. The

prediction results aid in the early diagnosis and treatment of individuals with high-risk medical disorders (Aleem, Alshamrani, & Alsehhri et al., 2022).

This mental health disorder affects millions across the world, and there are no effective treatments for it. Many among these affected individuals use social media as a platform to express and seek confirmation for their emotions. However, through social media, people find major consequences of using it as a source of affirmation. Much of this applies to sad individuals who present diverse realities. Everyone feels good and joyful; the perception that gives to life makes the sad people feel more inadequate, and that is where many mental crises are created. Depression only gets worse if someone is exposed to false reality all the time.

1.3 Research Aim

The objective of the study is to compare different machine learning algorithms with a view to establishing the best method for identifying young people with depression using social media. Specifically, in a bid to extend accessibility to mental health care and reach out to those who otherwise may not help themselves due to the social stigma associated with the act, the following methods are used: Recurrent Neural Networks, Support Vector Machines, and Naive Bayes.

1.4 Research Question

The following are the main research questions that this study seeks to address:

- How do Support Vector Machine and Recurrent Neural Network perform better than traditional diagnostic methods in detecting depression from social media?

- What influence would the early detection of depression from social media using Support Vector Machines and Recurrent Neural Networks have on emotional tone features and textual data analysis?
- What can be some of the challenges and advantages of applying Support vector machines, recurrent neural networks and Naive Bayes to social media such that a depression detection model based on machine learning is developed?

1.5 Research Objectives

The objectives of the research are as follows

- (a) To recognize depressed people by using machine learning methods such as Support Vector Machines, Recurrent Neural Networks, and Naive Bayes to evaluate social media sentiment and behavioural indicators.
- (b) To identify changes in posting frequency, engagement levels, emotion tone that signal symptoms of depression.
- (c) To evaluate and validate the effectiveness and accuracy of Support Vector Machines, Recurrent Neural Networks, and Naive Bayes through rigorous testing and validation

1.6 Research Scope

The scopes of the research are:

- a) The objective of this study is to develop and evaluate machine learning models to detect depression for which the dataset contains two major groups: condition group and control group. In specific, focus is attributed to three major algorithms related to machine learning throughout the study: Support Vector Machines, Recurrent Neural Networks, and Naive Bayes classifiers. It could be the dataset of tweets with metadata of users, such as the number of followers, friends, favorites, statuses, and the number of retweets that each tweet receives.
- b) The study will evaluate textual data from social media to detect and analyse depressive symptoms using natural language processing techniques. To make sure the models can accurately assess sentiment and emotional tone across various languages and populations, language usage will be a crucial factor to take into account.
- c) One limitation of this study is its dependence on publicly available datasets, which does not capture the entire range of socio-cultural nuances and variances in social media usage associated with depression. Furthermore, while the models aim to improve depression identification, their generalizability and usefulness in a variety of real-world contexts will be investigated, taking into account practical restrictions and ethical considerations.

1.7 Research Contribution

The following improvements to the field of mental health care and machine learning are anticipated as a result of this research:

- A better understanding of how social media interaction contribute to emotional states of individual particularly depression.
- The creation of highly effective machine learning model for depression detection using support vector machine and recurrent neural network which can be implemented to analyze social media data
- A comprehensive set of guidelines for integrating machine learning depression detection model with social media, ensuring ethical use, privacy protection.

These contributions have significant effect of early detection of depression and provide robust tools for mental health professional and have potential to enhance mental health care.

1.8 Report Organization

The following is how the rest of this report is organized:

- Chapter 1 provides the introduction investigates the global effects of modern lifestyles on mental health, addresses machine learning's function in detecting depression via social media analysis, and outlines research goals, questions, and objectives for increasing mental health treatment accessible.
- Chapter 2 investigates machine learning for detecting depression via social media. It evaluates the efficacy of SVM, RNN, and NB algorithms, focusing on data difficulties and ethical implications. This chapter intends to increase understanding and application of machine learning in mental health diagnosis using social media data.
- Chapter 3 discusses approaches for comparing machine learning models for diagnosing depression using social media data (e.g., Facebook and Instagram). It describes data collection, pre-processing, model selection (SVM, RNN, Naive Bayes), and performance assessment (accuracy, precision, recall, F1 Score). The chapter finishes with observations on model efficacy and ethical implications.
- Chapter 4 focuses on developing machine learning models to detect depression in social media data. It describes data pre-processing steps such dataset organization, standardization, and normalization. The chapter describes how to use Support Vector Machines (SVM), Recurrent Neural Networks (RNN), and Naive Bayes to analyse social media text for emotional expressions associated with depression. It describes the experimental design, which includes data collection methods, content analysis utilizing performance metrics (accuracy, precision, recall, F1 score), and parameter testing (model and batch size). The chapter finishes by reviewing the methodology's application and paving the way for future research and practical consequences.

1.9 Chapter Summary

The chapter covers how machine learning techniques aid in detection of depression from interfaces made in social media. Conventional treatment of the problem is more or less not valid, as depression is very intricate problem on digital communication platforms. In this context, the machine learning algorithms should be able to run the analyses of large pools of data in the social media for signs of emotional setup that should be precursors of depression. The research will be carried out conductively, with this conduction serving as a means to facilitate early detection, to validate the models, and to increase access to mental health care while being mindful of ethical and practical considerations.

CHAPTER 2

LITERATURE REVIEW

2.1 Introduction

This chapter's literature review focuses on machine learning, especially in regard to social media, to navigate this very broad area of mental health detection. No doubt that the way of socializing has a correlation with mental health today, when technology touches every aspect of our lives. This is an initial attempt to probe the scope of using machine learning techniques for the identification of depression on social media.

Now despair is one of the feelings expressed on a range of social media platforms, including Facebook and Instagram. These platforms are volumes of data, holding great promise but posing challenges in detecting mental health issues. The objective of this study on literature review was to analyse machine-learning methods in searching through data for traces of depression. In this regard, models of machine learning—namely, Naive Bayes (NB), Support Vector Machine (SVM), and lastly, Recurrent Neural Network (RNN) are taken into consideration to ascertain their efficiency and viability.

2.2 Problem Formulation

Determining how effectively the Support Vector Machines (SVM), Recurrent Neural Networks (RNN), and Naive Bayes (NB) algorithms will perform in estimating depression markers from social media data is the formulation issue. More precisely, it will provide performance metrics for identifying the accuracy, precision, recall, and F1 scores of related algorithms in identifying depression indications found in users' social media postings.

2.2.1 Research Domain

The study domain literature review clarifies the use of machine learning algorithms for depression diagnosis, particularly on social media platforms like Facebook and Instagram. The field of machine learning for mental health diagnoses has advanced significantly in the last several years, but there is still a lack of knowledge on the practical effectiveness of various methods, such support vector machines and recurrent neural networks.

Although there is a wealth of knowledge regarding conventional methods of diagnosing depression, the challenge is in examining the vast and complex data generated by social media sites to ascertain how well various machine learning models perform in detecting symptoms of depression, as well as comprehending the practical and ethical challenges associated with implementing these models. Large volumes of social media data to be analyzed by machine learning, which creates the foundation for the proposed research's major impact by assessing the effectiveness of recurrent neural networks and support vector machines and articulating the gaps in the literature.

2.2.1.1 Techniques for Depression Detection

Social media has enabled more opportunities for researchers to identify depression through the analysis of user-generated content. Various machine learning techniques are involved in the process of spotting depression from social media data. These include supervised learning that makes use of labelled datasets, unsupervised learning searching for hidden patterns, and deep learning models capable of extracting subtle features from text and images. Although researchers want to surpass the conventional diagnostic techniques through these machine learning technologies, their objectives remain to enhance scalability and accuracy in depression identification. This section explains how each of these vast depression diagnosis methodologies is applied in the framework of social media and uses that might be exclusive to the methodology.

2.2.1.2 Depression Detection in Social Media

Depression is one of the most serious public health concerns worldwide, with hundreds of millions of people affected globally. Traditional normative approaches to the diagnosis of depression—either clinician interviews or questionnaires by self-report should be subject to interviewer bias and lack fine-grained temporal resolution and poor accessibility. With the emergence of social media, further avenues have opened for detecting the first signs of depression by analysing online behaviour and content.

Twitter, Facebook, and Instagram are social media platforms that offer valuable data sources where users often share their emotions, opinions, and behaviours. As a result, machine learning (ML) methods have been developed to identify sadness by analysing text data obtained from these platforms. Applying machine learning techniques to social media data can provide more efficient and scalable solutions for detecting depression, in contrast to conventional methods Liu, D., et al. (2022).

2.2.1.3 Machine Learning Techniques

Several approaches of machine learning have been explored in recognizing depression from social media data: strategies with supervised learning, unsupervised learning, or deep learning methods.

2.2.1.4 Supervised Learning

In this case, the models would be trained using already tagged datasets, indicating depression or not. Several common techniques are applied in supervised learning, including SVM and Logistic Regression, Decision Trees, among many others. The models learn to classify new data by patterns found within the training data analysis. (Tariq et al., 2020).

2.2.1.5 Unsupervised learning

It refers to those techniques that include clustering, mining association rules, and require no labelled data. These constitute techniques to find hidden patterns or clusters in data, which would then be explored for the presence of depressive symptoms based on (Guntuku et al., 2019).

2.2.1.6 Deep learning

It is sub-branch of machine learning, which involves artificial neural networks that are large in terms of number of layers. More specifically, models such as Convolutional Neural Networks and Recurrent Neural Networks have been applied for the extraction of features from text and image modalities in social media data as indicators for detecting sadness. It is in this class of models that one is able to automatically learn from the input features that drive performance at a very high power level within the highly constrained and extremely competitive domains (Shatte et al., 2019).

2.2.1.7 Existing Models and Techniques

Various models and methodologies have been suggested for identifying depression based on data from social media. These methodologies frequently integrate machine learning and natural language processing (NLP) techniques to examine textual and visual data.

2.2.1.8 Models Utilizing Machine Learning Techniques

De Choudhury et al. (2013): This study proposed a model for predicting depression based on linguistic and behavioral features extracted from Twitter data. The utilized model employed supervised learning methodologies and attained a commendable level of accuracy in discerning persons afflicted with sadness.

Moreno et al. (2011): This study analyzed Facebook posts of college students to identify depression-related disclosures. The researchers employed a blend of manual coding and computer text analysis approaches to identify indications of depression in the posts.

Reece et al. (2016): This study used image processing and machine learning techniques to analyze Instagram photos for signs of depression. The researchers discovered a correlation between depression and the tendency of individuals to share photographs with darker hues, reduced brightness, and lower degrees of saturation.

2.2.1.9 Integration of Natural Language Processing (NLP) Techniques

It involves essential NLP techniques applied in text data obtained from social media; their use would drastically improve the analysis of textual data to identify depression.

Sentiment Analysis: The sentiment analysis is the computational determination of whether text expresses whether the text expresses a good, negative, or neutral sentiment. Researchers will hence use sentiment assessment for posts by users in social media to establish people showing symptoms of depression (Becken et al., 2020).

Topic Modelling: Such methods like Latent Dirichlet Allocation — can be used to infer the latent subjects or themes from a text corpus. In identifying depression-related subjects, one gets to know the challenges and issues that characterize a person who is depressed (Blei et al. 2003).

Emotion Detection: This is establishing the expressed emotions in a text, whether it be joy, sadness, rage, and fear among others. Researchers have used this to analyse the kind of emotion expressed in micro-blog posts to find out people experiencing negative emotions characteristic of depression (Calvo et al., 2010).

2.2.1.10 Comparison of Techniques

Table 2.1 compares the approaches used in the studies listed below. It illustrates each approach's strengths and weaknesses, providing insights into their suitability for various forms of social media data.

Table 2.1: Comparison of Techniques

Techniques	Strengths	Weakness
Supervised Learning	Labelled data and level of accuracy is high	It needs large amount of labelled data
Unsupervised Learning	Can detect hidden patterns without labels	Results can be hard to interpret
Deep Learning	Can learn features from data	Needs intense computational power

NLP	Effective in text analysis	Requires high quality text data
-----	----------------------------	---------------------------------

2.2.2 Description of Related Studies

In mental health sector, investigation into depression detection on social media through machine learning techniques has become significant. The developing operation of these sites have resulted in an avalanche of information which have been used to establish whether individuals are suffering from psychological problems such as depression. The abundance of data available can be used either as a benefit or a challenge when it comes to detection of mental health issues. In this research, looking into machine learning methods that focus on detecting depression using social media data. Support Vector Machine (SVM), Recurrent Neural Network (RNN) and Naive Bayes (NB) models will be explored in order to know how effective and practical they are when utilized for depression detection from online platforms.

2.2.2.1 Summary Table of Related Research

This section displays a summary of related works in machine learning by several researchers. Table 2.2 highlights the techniques and algorithms used, the data observed by researchers and each purpose

Table2.2: Description of related studies

Reference	ML Algorithm	Datasets	Research Purpose	Best Performance
Babu, N. V., & Kanaga, E. G. M. (2022)	SVM, Naive Bayes, TAN CNN, LSTM	Kaggle (Twitter)	To transform raw data into actionable insights through systematic collection, cleaning, analysis, interpretation, and visualization, facilitating informed decision-making and strategic planning. algorithms, assessing their performance against traditional machine learning models for text sentiment classification with and without emoji cases.	CNN+LSTM
(Ahmed, A., Sultana, R., Ullas, M. T. R., Begom, M., Rahi, M. M. I., & Alam, M. A. (2020, December))	Convolutional Neural Network, Support vector machine	Two datasets of anxiety and depression	To develop an effective early-stage diagnostic model for anxiety and depression using psychological assessments and machine learning, achieving high accuracy in identifying mental disorders among Bangladeshi women.	CNN
(Aleem, S., Huda, N. U., Amin, R., Khalid, S., Alshamrani, S. S., & Alshehri, A. (2022))	RF, Synthetic Minority Oversampling Technique (SMOTE), SVM DCNN	604 individuals, including the sociodemographic and psychosocial data and the Burns Depression Checklist (BDC) data	To review the application of machine learning algorithms in diagnosing depression, present a general diagnostic model, categorize the ML algorithms into classification, deep learning, and ensemble methods, discuss the objectives	DCNN

Bokolo, B. G., & Liu, Q. (2023).	Naïve Bayes, RoBERTa, Random forest	Sentiment140 dataset	The objective of this study is to investigate the feasibility of detecting depression from user tweets using machine learning techniques	RoBERTa
Bieliński, A., Rojek, I., & Mikołajewski, D. (2023)	LBFGS Stochastic Dual Coordinate Ascent	Clinical dataset and compared based on criteria in the form of learning time	The study aims to evaluate the effectiveness of machine learning algorithms in diagnosing mental health conditions, specifically stress among active workers, and to develop a virtual mental health index.	Stochastic Dual Coordinate Ascent
Kuo, A. T., Chen, H., Kuo, Y. H., & Ku, W. S. (2023).	ContrastEgo, GNN	PsycheNetG	This research paper “Dynamic Graph Representation Learning for Depression Screening with Transformer” is mainly focused at developing a new technique through which depression can be detected by dynamic graph representation learning. Authors come up with a model called ContrastEgo, which used information from social media thus implement using advanced machine learning techniques to predict the existence of depression in a person based on their interactions on platforms like Twitter.	ContrastEgo
Haque, U. M., Kabir, E., & Khanam, R. (2021).	RF, XGBoost (XGB), Decision Tree (DT), and Gaussian Naive Bayes (GaussianNB)	In this study, a large dimensional dataset of 667 features and 6310 cases has been used, with the number of classes varying widely.	The purpose of this research, ‘Detection of child depression using machine learning methods’, by Haque, Kabir and Khanam (2021) is to build and assess machine learning algorithms for forecasting depression in kids and teenagers aged 4-17 years.” Young Minds Matter (YMM) survey data was utilized in this research.	RF

Zeberga, K., Attique, M., Shah, B., Ali, F., Jembre, Y Z., & Chung, T. (2022, March 3)	Bi-LSTM and BERT Model	Dataset from Reddit and Twitter	A Novel Text Mining Approach for Mental Health Prediction Using Bi-LSTM and BERT Model is an attempt to come up with a more efficient as well as effective method of predicting mental health problems such as depression and anxiety using the analysis of social media posts. The investigation of the social media posts makes use of the latest text mining tools, in particular Bi-LSTM (Bidirectional Long Short-Term Memory) and BERT (Bidirectional Encoder Representations from Transformers), for the purpose of keeping the meaning that is presented in the contexts used for posting on this platform.	BERT
HosseiniFard, B., Moradi, M. H., & Rostami, R. (2013).	KNN, LDA and LR	EEG of 45 depressed patients and 45 normal subjects by power of EEG bands and four nonlinear features	This study shows that nonlinear analysis of EEG can be a useful method for discriminating depressed patients and normal subjects..	LR
Smys, S., & Raj, J. S. (2021).	Decision Tree, Support Vector Machine, Random Forest, Naïve Bayes Method	This research article tested 2500 sentences for emotion prediction from a Twitter dataset	The Proposed Hybrid Method Provides Better Results For Early Detection And Retained Good Sensitivity And Better Accuracy Of Existing Methods. The Study From Results Can Help To Develop A New Idea To Develop A Early Prediction Of Various Emotions Of People Present In Social Media.	The Fusion Of Support Vector Machine And Naïve Bayes Theorem Is Used To Acquire High Accuracy And Sensitive Metrics

Choudhury, M D., Gamon, M., Counts, S., & Horvitz, E. (2021, August 3)	PCA,SVM,RA,Cross validation.	Final dataset of 476 users	In the work "Predicting Depression via Social Media" by Munmun De Choudhury, Michael Gamon, Scott Counts, and Eric Horvitz seek to use social media data to test and diagnose major depressive disorder. The authors collected a group of Twitter users who had diagnosed themselves with clinical depression using crowdsourcing. During a stepwise process as they connected with different subgroups people gathered to self-manage their depression, they found that individuals who used negative and hesitant forms of communication most often posted to social media that were related to depression.	The best performing classifier was found to be a Support Vector Machine classifier with a radial-basis function (RBF)
Sudha, K., Sreemathi, S., Nathiya, B., & RahiniPriya, D. (2020).	K-Nearest Neighbors, Naïve Bayes, Decision Tree and Random Forest have been used for the detection model	The text dataset is collected from people who were clinically diagnosed with depression.	The objective of the research paper "Depression Detection Using Machine Learning" by Sudha, Sreemathi, Nathiya, and RahiniPriya (2020) was to create a depression detection system based on machine learning. The study is looking at the written text data collected from people with depression to find patterns and characteristics that are diagnostic of the psychological state.	In the proposed system, the Decision Tree Algorithm is the most efficient algorithm with an accuracy of 98.24%, Naïve Bayes Algorithm is the fastest algorithm with 7.6second.

According to Babu and Kanaga (2022), sentiment analysis will provide a fresh perspective on how people feel about the things that are going on in their lives. In order to determine if a person's relatively few feelings will indicate that they are depressed or anxious, Ahmed et al. (2020) suggested a diagnostic of depression and anxiety using the supervised learning model. It was made clear that these illnesses seriously impair a person's capacity to function in daily life, and that developing nations like Bangladesh—which have the greatest rate of anxiety disorders—should prioritize preventing these mental illnesses. A primary focus of the model is its effectiveness, which highlights the cost and time savings associated with early intervention for the management of mental diseases (Kanaga et al., 2022).

Aleem et al. (2022) highlight the significance of the psychological effects of contemporary lifestyles on human body and health as well as the rapid advancement of medical technology based on healthcare data analysis while addressing its application to allow a computer to identify depressed affects. The benefit of using a machine learning model to analyze healthcare data was summarized in the research (Aleem et al., 2022).

Bokolo and Liu's (2023) study compared several deep learning-based methods for identifying depression from social media content and ranked them against various human evaluators. Models including logistic regression, Bernoulli Naive Bayes, random forests, and transformer models like Distil BERT, Squeeze BERT, DE BERTA, and Roberta were used by Bokolo and Liu (2023) in a research using a dataset of 632,000 tweets. This strategy demonstrates how easily mental health concerns should be identified and tracked by social media analysis in the early stages (Bokolo et al., 2023).

Bieliński, Rojek, and Mikołajewski (2023) differentiate selected machine learning algorithms in analysing mental health indicators, focusing on the creation of a virtual mental health index. Their findings demonstrate the effectiveness of these

algorithms in treating mental health issues and illustrate the expanding use of machine learning in clinical settings for diagnosis, treatment, and care (Bieliński et al., 2023)

Kuo, A. T., Chen, H., Kuo, Y. H., & Ku, W. S. (2023). This research primarily focuses on improving depression screening via the use of transformer architecture in conjunction with the dynamic graph representation learning technique. This method has the benefit of being able to anticipate things more precisely and in depth as it can capture both the temporal dynamics and the connections between the individuals in the network. Contrast Ego's model verification demonstrates that it outperforms the most recent rule-based model by up to 4%, demonstrating the technique's validity and the model's suitability as a useful tool for psycho-emotional healthcare (Kuo et al., 2023)

Haque, U. M., Kabir, E., & Khanam, R. (2021). This study particularly stands out in its emphasis on using artificial intelligence models to identify depression in kids. The drawback of such measures is the use of drugs and other forms of treatment to avoid early symptoms. With the use of computational techniques like random forests, decision trees, and logistic regression, the research demonstrates great validity and efficiency, accurately depicting children with mental illnesses. The use of several algorithms has made this tool very precise and successful, making it a valuable resource for mental health professionals who deal with children. (Haque et al., 2021)

Zeberga, K., Attique, M., Shah, B., Ali, F., Jembre, Y. Z., & Chung, T. (2022). This research is a standout by employing Bi-LSTM and BERT models jointly, thereby creating an efficient text mining technique for mental health prediction. This method's primary benefit is that it uses cutting-edge natural language processing techniques to increase diagnostic accuracy. The model surpasses traditional models and demonstrates its present use in mental state monitoring and diagnostics in the real world by providing significant augmentation via the use of social media text data (Zeberga et al., 2022).

Hosseinfard, B., Moradi, M. H., & Rostami, R. (2013). This study is novel in its application of nonlinear features from the EEG signals for depression patient and healthy subject classification. This method's primary benefit is its non-evasiveness, which functions as an excellent diagnostic tool without requiring subjective evaluations yet being quite accurate. The study's outcome is the consequence of applying extreme machine learning techniques, which are logically at the highest level of accuracy. This further establishes the validity of EEG as a real predictor for the diagnosis and classification of depression and enhances the accuracy of mental health assessment results (Hosseinfard et al., 2013).

Smys, S., & Raj, J. S. (2021). The distinction of this comparative study lies in its ability to evaluate various deep learning techniques in depression early detection using social media data. This investigation's strength is that it thoroughly examined models like CNNs, RNNs, and LSTMs, concluding that each is adept at handling lengthy data sequences. The research concludes that deep learning models are a method to demonstrate efficiency in this domain since they provide a significant improvement in prediction accuracy (Smys et al., 2021).

Choudhury, M. D., Gamon, M., Counts, S., & Horvitz, E. (2021). The paper succeeds in using social media data in order to differentiate depression, which is a new method in mental health monitoring. Its capacity to closely examine vocabulary and social conversation to create prediction models that will identify potentially dangerous people accounts for its use. This study highlights the use of social media data in conjunction with machine learning, which results in an effective way to monitor mental health and implement early intervention techniques (Choudhury et al., 2021).

2.3 Machine Learning Approaches for Depression Detection

In order to detect depression, Researchers have been looking for approaches to effectively identify depression. The techniques of machine learning are being investigated for identification approaches effective in identifying depression. Besides, the paper discusses several methods used to detect emotion and mood in people. Machine learning examines how facial expressions, images, chatbots portraying emotions, and texting on social network platforms could effectively identify emotions and, thus, depression. Some of the ML techniques that recognize emotion from text are the Support Vector Machines (SVM), Recurrent Neural Networks (RNN), and Naive Bayes (NB). Machine learning algorithms can detect emotion from text, and to detect emotion from tweets, sentiment analysis can be performed using NLP techniques (Kanaga et al., 2022).

2.3.1 Support Vector Machine

SVM is a supervised learning technique that helps with regression prediction and classification by helping to analyze and identify patterns in data. It automatically transforms nominal data into numerical for categorization. The SVM method makes it possible to choose lines, often referred to as separators, in a way that closely resembles the underlying function in the target space.

These are ham and spam, two hyperplanes. These might be straight, curved, or polygonal lines. By using this method, a prediction model that categorizes data into several groups is created. For instance, estimating the likelihood that a student would fail. There are many classes, either groupings or categories, in the dataset that this classifier uses.

2.3.2 Artificial Neural Network

An artificial neural network (ANN) is a framework that connects nodes and various neuronal layers. The neurons are connected by weighted connections. ANN is a method that suggests studying the nervous system and brain, as shown in Figure 2.1 (Facundo et al., 2017). Be aware that not every weight is shown.

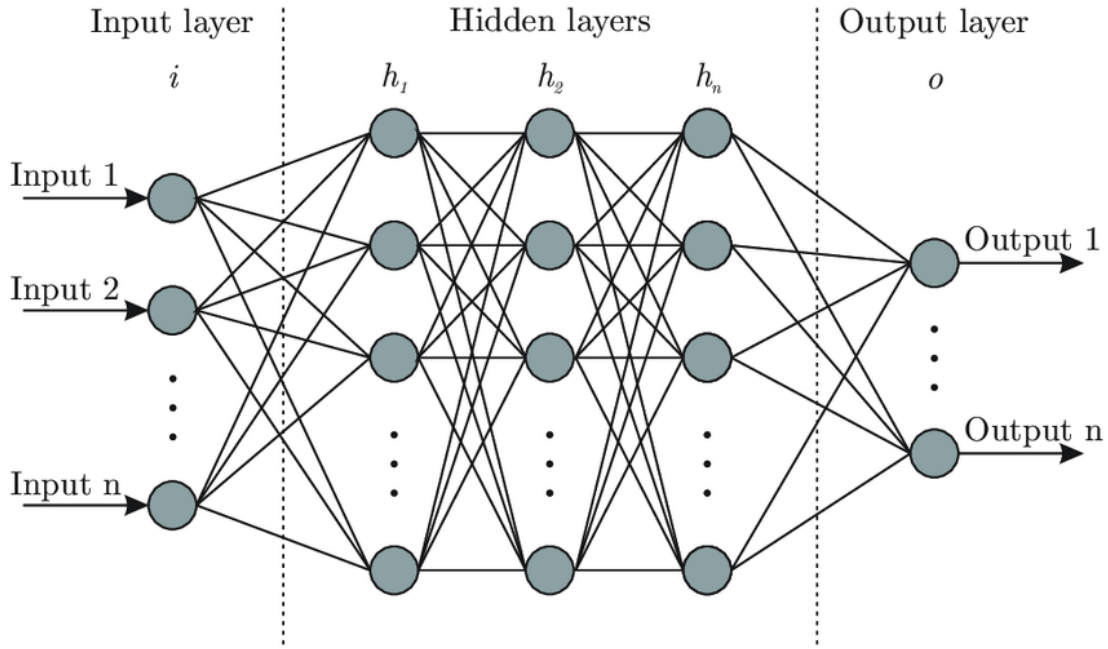


Figure 2.1: Simple artificial neural network architecture (Facundo & Juan, 2017)

An ANN includes three layers: the input layer, the hidden layer, and the output layer. Based on Figure 2.1, some values are for hidden layers; some values are for output layers; and the input value is for the input layer. First of all, the data is sent to the input layer. For one characteristic, each input is an independent variable. Note that the variables have to be normalized or

standardization. It proceeds to the secret layer after that. An operation known as the weighted sum and activation function is carried out in this layer. The total weight of synapses plus the input value is called a weighted sum. The connections between the input layer and the hidden layer are known as synapses. The neuron's weighted sum is converted by the activation function. After that, the neuron chooses

whether or not to forward the signal to the next layer. Lastly, output values will be generated in the output layer. The output result will be binary, continuous, or categorical.

2.3.3 Naive Bayes

For most textual classification tasks, Naive Bayes is a noisy but effective solution. Naive Bayes is a friendly classifier, working well especially when the data volume is large. It copes efficiently with noisy and high-dimensional data, like what commonly happens with Social Media posts. Figure 2.2 illustrates the overall architecture of the Naive Bayes approach.

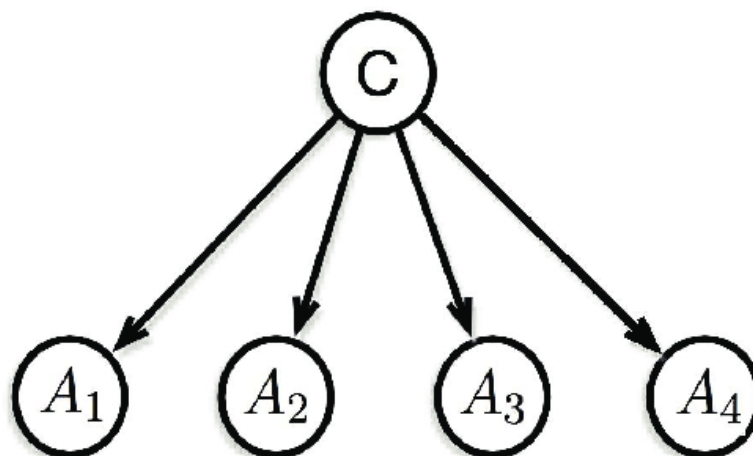


Figure 2.2: Basic structure of Naïve Bayes

2.4 Future Directions and Challenges

Even while research on the use of NLP and machine learning for depression identification has advanced significantly, there are still a number of obstacles to overcome. It is crucial to protect data privacy and take ethical issues seriously. Furthermore, the use of big labeled datasets in supervised learning methods is a problem due to the difficulty and resource-intensive nature of obtaining such data.

Future studies should concentrate on creating more effective techniques for labeling and gathering data, enhancing the models' capacity to generalize to various populations, and tackling privacy and ethical concerns (Saha et al., 2020).

2.5 Chapter Summary

This chapter comprehensively explores the use of machine learning techniques for detecting depression from social media data, emphasizing Support Vector Machines (SVM), Recurrent Neural Networks (RNN), and Natural Language Processing (NLP). As a wealth of information where people freely express their ideas and feelings, social media is far more useful for tracking mental health than traditional diagnostic techniques. By filling in the gaps and overcoming obstacles in the detection of sadness, contemporary machine learning techniques assist in identifying subtle signs in social media data. The chapter examines a number of methods, stressing the benefits and drawbacks of each, such as supervised learning, unsupervised learning, deep learning, and natural language processing. The study indicates that although machine learning, particularly in conjunction with natural language processing (NLP), exhibits potential for detecting signs of depression, issues including data privacy, ethical considerations, and the need for large labelled datasets still need to be addressed. To improve mental health diagnoses, future research should concentrate on improving model accuracy and applicability across a range of groups.

CHAPTER 3

RESEARCH METHODOLOGY

3.1 Introduction

This chapter presents the overview of the methodologies adopted for the comparison of multiple machine learning models in diagnosing depression on social media platforms like Facebook, Instagram, and other platforms. The text offers details related to techniques, process, and criteria used in support vector machines, recurrent neural network, and Naive Bayes model selection. This chapter concludes with a synopsis of the measurement tools and their approaches to performance. The methods and procedures outlined in this chapter are those that have direct links with the conceptual framework of Chapter 2, Literature Review, and their implementations are aimed at achieving the objectives laid out in Chapter 1, Introduction. The chapter finally concludes with a summary.

3.2 Operational Framework/Research Workflow

The detection of depression through social media goes through four stages: literature review, data collection, algorithm model development, and testing or outcome analysis. These machine learning algorithms for textual data and sentiment analysis furnish opportunities for mental health monitoring and interventions (De Choudhury et al., 2013; Guntuku et al., 2017)

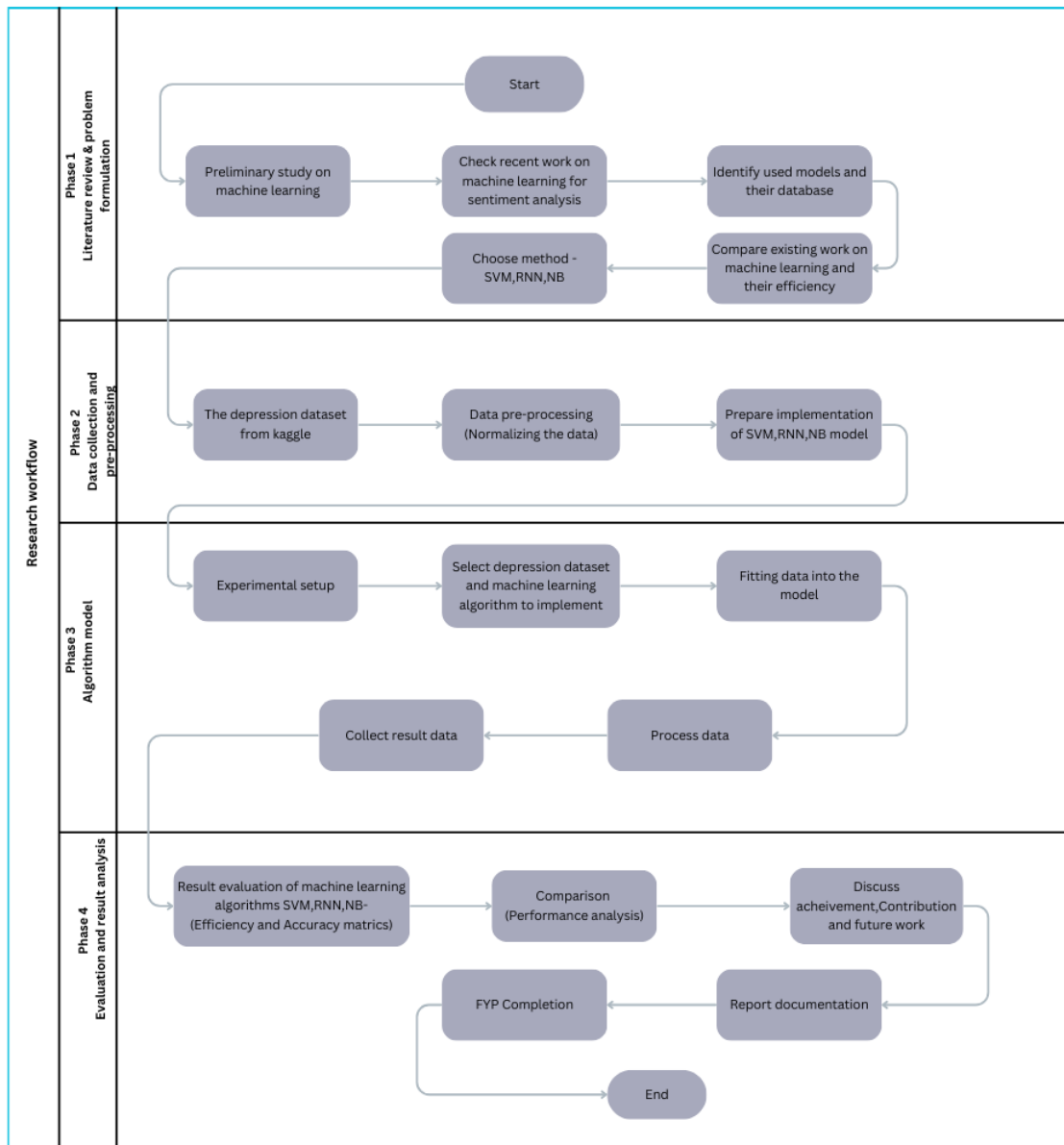


Figure 3.1: Operational Workflow

3.2.1 Phase1: Literature review

The literature review includes studies, papers, and articles that examine the latest detection methods in mental health. These methods utilize machine learning techniques applied to content from social media. This phase delineates advancements in methodology and highlights issues within the discipline, organizing current approaches and datasets into categories. This classification facilitates comprehension of the benefits, drawbacks, and suitability of different machine learning models in identifying depression (Ahuja & Banga, 2019; Tsugawa et al., 2015).

Various developed methodologies apply classifiers such as SVM, RNN, and Naive Bayes to detect depression. This categorizes the already developed methodologies and the available data set, creating a proper understanding of the research area. Such categorization is going to be done to identify the advantages, disadvantages, and applicability of the various machine-learning-based models in detecting depression. The strings of research get more specific based on what the literature reviewed in the domain has to say, to discover potentials and limitations of machine learning approaches in the detection of depression in social media.

3.2.2 Phase 2: Data collection

Data collection involves the acquisition of data sets from numerous places on social media networks including Facebook, Instagram, and Kaggle among other places. Assuming that these data sets are to be fed into the machines for learning, there is a need for standardization, such as annotating the text data as well as any other forms of relevant textual information. At this stage, it also sets up a computational environment meant for training as well as testing the machine learning models. These include proper software installation as well as computational libraries as indicated herein, assisting in the prevention or reduction of mental illnesses such as stress and depression through early detection with the use of data mining techniques, mainly textual analysis, which have been described (Coppersmith et al., 2014; Reece & Danforth, 2017).

- **Text Pre-processing:** This step cleans up the text data by removing stop words, punctuation, and irrelevant content. Standardization of text data employs a tokenization, lemmatization, and stemming process (Aggarwal & Zhai, 2012).
- **Annotation and Labelling:** It ensures that the data has labels indicating depressive content and non-depressive content. This is an important stage for supervised learning models, emphasis added by Guntuku et al. (2017).

3.2.3 Phase 3: Algorithm Model

Designing a context-specific test plan for the chosen machine learning models (SVM, RNN, and Naive Bayes) incorporates data and computation management through pipelines. The models allow uniform datasets, which guarantees accuracy. The best hyperparameters relate to learning rate, batch size, and the number of layers. Experimental adjustments along these lines are done to find the best combinations that can give the ideal elements; iteratively, this makes the model most accurate and efficient (Calvo et al., 2017; Orabi et al., 2018).

- **Feature Extraction:** This stage involves the identification from the textual data of such features as sentiment scores, frequency of depressive keywords, and linguistic styles (Tausczik & Pennebaker, 2010).
- **Model Training:** A process of applying training data to models to modify parameters that minimize errors, hence making better predictions.

Hyperparameters need to be fine-tuned to perform optimization in the best possible way. The instances are critical components of the architecture, which include parameters such as learning rate, batch size, number of layers, and so forth. Following that, experimental modifications are conducted to determine the best combinations of factors. Machine learning models are then trained using pre-prepared data and iteratively improved to improve accuracy and efficiency.

3.2.4 Phase 4: Evaluation and Result Analysis

Outcome information is developed and analysed to measure the models' performance. Effectiveness and accuracy are evaluated concerning parameters such as accuracy, computational efficiency, and stability. This phase interprets the data to determine the best approach for detecting depression on social media and identifies areas needing improvement (Shatte et al., 2019; Yates et al., 2017).

- **Model Performance:** Evaluate the effectiveness and accuracy of trained models with respect to parameters such as accuracy, efficiency in computations, and stability. In other words, this is the section where I get to explain just how well these models execute their tasks and the accuracy at which conclusions are drawn (Cortes et al., 1995; Hochreiter et al., 1997).
- **Cross-Validation:** It ensures the models generalize well to unseen data by splitting it into several test and train sets. (Kohavi et al., 1995).
- **Statistical tests:** Several statistical tests are conducted to compare the performances of these models and validate the results (Friedman et al., 2001).
- **Evaluation and Interpretation:** Based on results, establish how this machine learning model works toward detecting depression on social media. It involves interpreting the data to determine if the approach to settle on is best and areas where additional improvements need to be bought in (Guntuku et al., 2017; Shatte et al., 2019).

3.3 Justification

For this study, the researchers choose to use Support Vector Machines (SVM), Recurrent Neural Networks (RNN), and Naive Bayes models due to their distinctive architectures and demonstrated effectiveness in various tasks related to text and pattern recognition.

3.3.1 Support Vector Machines (SVM)

Support Vector Machines (SVMs) are extensively employed for constructing classification models, especially in domains with a large number of dimensions. Due to the intricate and diverse characteristics of social media data, Support Vector Machines (SVMs) are successful in identifying distinctions between depressive and non-depressive content by analysing textual aspects (Cortes & Vapnik, 1995; Wang et al., 2013).

3.3.2 Recurrent Neural Networks (RNN)

Concerning ordered data, recurrent neural networks turn out to be ideal for that very function and, above all, very beneficial in the assessment of social media posts over time. They capture dependencies and contextual information in sequences, hence unveiling the evolution of a user's post. Attention-augmented RNNs focus on important segments of the sequence, hence improving the precision of depression identification. Indeed, this approach has been supported by works spearheaded by (Hochreiter & Schmidhuber, 1997; Tadesse et al., 2019).

Some more sophisticated forms of attention-augmented RNNs focus on relevant parts of the sequence, making it a framework for depression detection with high accuracy.

3.3.3 The Naive Bayes:

Naive Bayes is a straightforward and efficient method for text categorization, especially when dealing with noisy and high-dimensional data supplied from social media. The method is highly efficient and yields easily interpretable findings, making it a valuable tool for early screening when compared to more complex models (McCallum & Nigam, 1998; Nahar et al., 2012).

Thus, the reason why its model was included in the ensemble is that Naive Bayes is simple in approach and noisy in effectiveness in most cases of text classification tasks. Though simple to use, Naive Bayes is one of the classifiers that works well enough, especially in a vast amount of data; it is efficient sufficient with noisy and high-dimensional data as obtained through social media content. Naive Bayes results are, therefore, easy to interpret, providing the screening tool in early comparison with the more sophisticated models.

3.4 Performance measurement

This part tests machine learning models in preserving their efficiency to detect depression from social media data. To evaluate the performance of the model in classifying depression signals, accuracy, precision, recall, f1 score and cross validation are used for the robustness of the depressive signals.

3.4.1 Accuracy

The primary metrics of evaluating our models for machine learning are accuracy measures. These are proportions that depict true positives and true negatives upon capturing the cases related to depression out of the totals. The accuracy obtained will reveal an easy and straightforward idea of the classification of the models about social media posts.

$$\text{Accuracy} = \frac{\text{Number of correct prediction}}{\text{Total number of prediction}}$$

3.4.2 Precision

Precision refers to the accuracy of the optimistic predictions of the model. It represents the ratio of true positives relative to the sum of true and false positives. In that sense, it bears enormous importance here to minimize false identifications of depression.

$$\text{Precision} = \frac{\text{True positives}}{\text{True positives} + \text{False positives}}$$

3.4.3 Recall (Sensitivity)

Recall is the model's ability to extract, with reasonable magnitude, all the relevant instances of the class in question. It finds use in calculation as the number of true positives divided by the sum of true positives and false negatives. Very high recall ensures that people with expressions of depression are not missed at any cost.

$$\text{Recall} = \frac{\text{True positives}}{\text{True positives} + \text{False positives}}$$

3.4.4 F1 Score

The F1 score is the harmonic mean of the precision and recall of a test combined into one single measure. It should thus be a good metric in case of imbalanced class distributions because it would give equal weight to precision and recall.

$$\text{F1 Score} = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}}$$

3.4.4 Cross validation

Cross-validation ensures the developed models are generalized and robust. It involves partitioning the dataset into several folds and training/testing each fold on the rest of the data, helping to assess model consistency and reliability (Kohavi et al. ,1995).Cross-validation involves partitioning the data set into several folds and, consequently, model training based on testing for each of the folds on the rest of the data. This would help understand the consistency of performance or reliability of the model.

3.5 Chapter Summary

This chapter offered an overview of the use of machine learning models to diagnose depression via social media. It comprised a disciplined framework covering data collecting, preprocessing, model selection, training, and evaluation. The study made advantage of a large dataset including several aspects of social media engagement. Models like Naive Bayes, RNNs, and SVMs were assessed depending on their strengths and applicability for several facets of the work. Model efficacy in identifying depression from social media data was evaluated using performance criteria including accuracy, precision, recall, and F1 score.

CHAPTER 4

RESEARCH DESIGN AND IMPLEMENTATION

4.1 Introduction

Chapter 4 addresses the useful exercises and thorough methods applied using machine learning models to identify depression on social media. Ensuring the researcher understands the procedure and can thus efficiently do the research, this chapter describes the stages required for data gathering, pre-processing, model training, evaluation, and implementation. The techniques and exercises discussed here complement past parts and help to support the general research agenda.

This chapter elaborates on the testing phase, where some performance metrics—accuracy, precision, recall, F1 score, have been run over training and test sets. There will be a detailed and adequate description of techniques used for the implementation and tools for performance evaluation of acquired models, including Support Vector Machine (SVM), Recurrent Neural Networks (RNN), and Naive Bayes (NB). It will allow for a complete comparison of the models of early detection of depression through a structured approach to effectiveness analysis. The background will then be laid for practical applications and further research.

4.2 Proposed Solution

The proposed method combines Support Vector Machines, Recurrent Neural Networks, and Naive Bayes for depression detection in social media text. A detailed overview of the methodology and benefits of these models used to analyse emotional expression and contextual signals in social media writing are reported in Chapter 3. This is a hybrid system that can inherit the strengths of SVM in linear classification, RNN for long-term dependencies, and NB for text-related data. It offers a way that will identify early signs of depression in social media posts all-rounded. Social media posts.

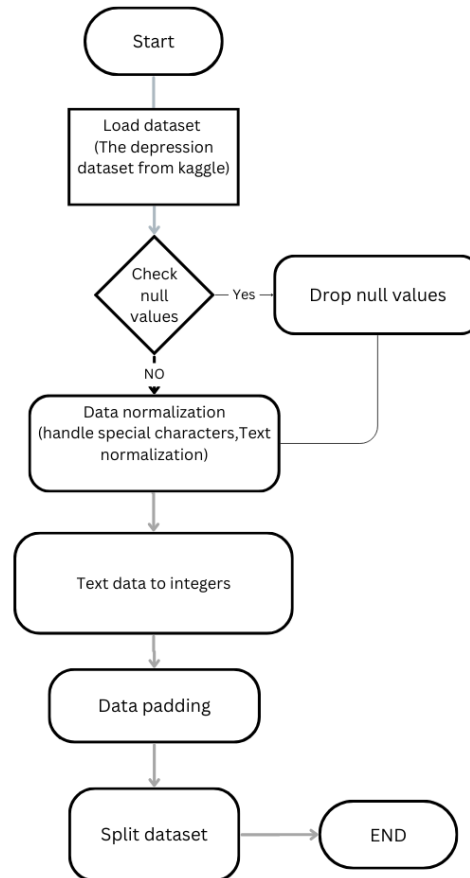


Figure 4.1: Data processing workflow

First, load the depression dataset on Kaggle. The second thing that should be done is to check if one of the data contains null values. In the incidence where such null values do exist, then this is the time they should be removed to ensure the completeness and integrity of the data.

It is with this regard that data normalization also handles special characters and text normalization as part of the data cleaning process. Data should have all this noise removed, and the text so formed standardized so that the data will be in a form that is ready to undergo the iterating process.

The process of converting textual data into numerical form is called tokenization. Thus prepared, this data is converted to become a kind of integer token that can be operated inside machine learning algorithms.

The following process, following tokenization, is padding the data so that the data is of uniform length. This is for consistency and, more than anything, to allow the model's training to be efficient. Further, the sets are separated into a training dataset and a testing dataset. This is to ensure that the robust model formed from an established pre-processing framework will already be taken into account before the performance assessment. The pre-processing framework, in this case will lay solid grounds for sourcing depression signs from the machine learning models that will have their basis and give very high-fidelity predictions to come up with a quality dataset for analysis.

4.3 Experiment Design

This experiment design primarily concerns five major phases: data collection, data pre-processing, model training and implementation, Cross-Validation and statistical tests and content analysis. A diverse dataset collection and machine learning model analysis to detect signals of depression within the text of social media data. Evaluation metrics will be used to assess model effectiveness, and further, the strategy for mental health monitoring will be improved.

4.3.1 Data Collection

The "Mental Health and Social Media" dataset will be downloaded from Kaggle:<<https://www.kaggle.com/datasets/infamouscoder/mental-health-social-media>>.

Since specimens of social media posts that are indicative of mental health conditions are integral to the training and testing of models, this dataset will form the basis of the study. The integrity of the data will be core to this, entailing either the removal or imputation of null values.

4.3.2 Data Pre-Processing

Data pre-processing involves the following steps:

- **Normalization:** Converting all the letters into case lower, followed by removing special letters and punctuations. This creates uniformity and decreases the amount of noise within the dataset.

- **Tokenization:** It insists that each textual data must be converted into numerical forms so that machine learning algorithms can process them. This step is most crucial for feeding data to algorithms.
- **Padding:** Padding tokens ensure that all sequences are of the same length. Again, the idea is that for efficiency in training, it is very important to have a fixed length of sequence.
- **Data Splitting:** Divide the dataset into two parts—training and testing—to make sure that the model implemented works efficiently. It's usually an 80/20 or 70-30 division.

Figure 4.2: Data Pre-Processing

```

Sample data from the dataset:
  Unnamed: 0      post_id      post_created \
0      0  637894677824413696  Sun Aug 30 07:48:37 +0000 2015
1      1  637890384576778240  Sun Aug 30 07:31:33 +0000 2015
2      2  637749345908051968  Sat Aug 29 22:11:07 +0000 2015
3      3  637696421077123073  Sat Aug 29 18:40:49 +0000 2015
4      4  637696327485366272  Sat Aug 29 18:40:26 +0000 2015

      post_text      user_id  followers \
0  It's just over 2 years since I was diagnosed w...  1013187241      84
1  It's Sunday, I need a break, so I'm planning t...  1013187241      84
2  Awake but tired. I need to sleep but my brain ...  1013187241      84
3  RT @SewHQ: #Retro bears make perfect gifts and...  1013187241      84
4  It's hard to say whether packing lists are mak...  1013187241      84

      friends  favourites  statuses  retweets  label
0      211      251      837      0      1
1      211      251      837      1      1
2      211      251      837      0      1
3      211      251      837      2      1
4      211      251      837      1      1
Total number of samples: 20000
Vocabulary size: 34822
Max sequence length: 34
Shape of X_train: (16000, 34), Shape of X_test: (4000, 34)
Sample padded sequence: [ 0  0  0  0  0  0  0  0  35  28  134  148  180  265
 2  36  427  19  301  7  32  102  20  545  4 1296  3 5844
15  40  445 105 182 265]
Sample label: 1

```

4.3.3 Model Training and Implementation

Support Vector Machines (SVMs):

- **Kernel Selection:** A proper kernel has to be chosen based on characteristics of the data among linear, polynomial, or RBF.
- **Hyperparameter Tuning:** Modify the parameters, such as parameters of regularization (C) and the coefficient related to the kernel (gamma), by grid search or random search.
- **Training:** Train the SVM model on the pre-processed training dataset.
- **Evaluation:** Test the model on this test dataset. Among many, the parameters to be estimated are accuracy, precision, recall, and F1 score.

Figure 4.3: SVM Model Evaluation

SVM Model Evaluation					
Accuracy: 0.8635					
Precision: 0.8658718330849479					
Recall: 0.8632986627043091					
F1 Score: 0.8645833333333334					
	precision	recall	f1-score	support	
0	0.86	0.86	0.86	1981	
1	0.87	0.86	0.86	2019	
accuracy			0.86	4000	
macro avg	0.86	0.86	0.86	4000	
weighted avg	0.86	0.86	0.86	4000	

Recurrent Neural Networks (RNN):

- **Architecture Design:** Finally design an RNN architecture with regard to the number of layers, number of neurons per layer, and activation functions that will be used.
- **Sequence-handling:** Take into consideration that it needs to deal with sequential data and, most importantly, temporal dependencies.
- **Training:** Train the RNN model on the training dataset. One can use BPTT or any other standard method for training neural networks.
- **Evaluation:** Run the model on the test dataset and evaluate using relevant evaluation metrics.

Figure 4.4: RNN Model Evaluation

RNN Model Evaluation					
Accuracy: 0.8675					
Precision: 0.838871187983614					
Recall: 0.9128281327389797					
F1 Score: 0.8742884250474383					
	precision	recall	f1-score	support	
0	0.90	0.82	0.86	1981	
1	0.84	0.91	0.87	2019	
accuracy			0.87	4000	
macro avg	0.87	0.87	0.87	4000	
weighted avg	0.87	0.87	0.87	4000	

Naive Bayes (NB):

- **Feature extraction:** It is a process of extracting features from text into a feature vector by using either the TF-IDF or Bag-of-Words approach.
- **Model Training:** Train the Naive Bayes classifier on the training dataset.
- **Evaluation:** Test the model with the test dataset and use the measures of accuracy, precision, recall, and F1 score to gauge the model's performance.

Figure 4.5: NB Evaluation Model

Naive Bayes Model Evaluation				
Accuracy: 0.875				
Precision: 0.8488745980707395				
Recall: 0.9153046062407132				
F1 Score: 0.880838894184938				
	precision	recall	f1-score	support
0	0.91	0.83	0.87	1981
1	0.85	0.92	0.88	2019
accuracy			0.88	4000
macro avg	0.88	0.87	0.87	4000
weighted avg	0.88	0.88	0.87	4000

4.3.4 Cross-Validation and Statistical Tests

- **Cross-validation:** Use k-fold cross-validation to confirm that the models generalize successfully to new data (Kohavi et al., 1995). By dividing the dataset into k subgroups, this strategy entails rotating the procedure k times to train the model on k-1 subsets and validate it on the remaining subset.
- **Statistical test:** Conduct statistical tests to assess model performance and validate results (Friedman et al., 2001). These tests help to establish whether the observed changes in performance measures are statistically significant.

Figure 4.6: Cross-Validation and statistical Test

```
SVM Model Evaluation
Accuracy: 0.88025
Precision: 0.8800592300098716
Recall: 0.8831104507181773
F1 Score: 0.8815822002472188
```

	precision	recall	f1-score	support
0	0.88	0.88	0.88	1981
1	0.88	0.88	0.88	2019
accuracy			0.88	4000
macro avg	0.88	0.88	0.88	4000
weighted avg	0.88	0.88	0.88	4000

```
SVM Cross-Validation Scores: [0.68225 0.71425 0.7145 0.70775 0.7225 ]
SVM Cross-Validation Mean Score: 0.70825

Naive Bayes Model Evaluation
Accuracy: 0.875
Precision: 0.8488745980707395
Recall: 0.9153046062407132
F1 Score: 0.880838894184938
```

	precision	recall	f1-score	support
0	0.91	0.83	0.87	1981
...				

```
Naive Bayes Cross-Validation Scores: [0.66025 0.70575 0.69425 0.70075 0.621 ]
Naive Bayes Cross-Validation Mean Score: 0.6764
```

4.3.5 Content Analysis and Discussion

The detection of depression from social media data was done using the performances of the Support Vector Machine, Recurrent Neural Network, and the Naïve Bayes machine learning model. Therefore, the analysis metrics will be cross-validated against each other and tested for speed of translation, errors, and processing. This evaluation will give insights into the effectiveness of detecting depression using the models.

4.4 Parameter and Testing Method

Parameters and Testing Method: This part of the explanation considers parameters within the machine learning model and the measures to be taken for testing the parameters. This shall include methods used to tune and test parameters for cross-validation.

4.4.1 Parameters

A) Model Size

Larger models need more memory and processing capacity to capture more intricate patterns. Based on the resources at hand and the unique needs of the study, the model's size and complexity should be balanced.

B) Batch Size

Simply put, the "batch size" sets the number of training samples that will be processed in one go. This has several effects: it can influence the rate of convergence, memory consumption, and the stability of the training process. Smaller batches generally induce more randomness, while larger batches typically provide more stable gradient estimation. Transformers work well with a batch size of up to 4,096 tokens.

For example, in the Depression Detection dataset, the model performance is tuned with a batch size of 2048 tokens,

4.4.2 Test Method

Mainly, there are a couple of key stages where structured approach to the testing of the models' performance are applied:

Data Splitting: The sample data is partitioned into two sets, one for training and the other for testing. A training set is used for fitting the model, and a validation set is used for testing the model.

Model Training: After application, each machine model—SVM, RNN, Naive Bayes—needs to be trained piecewise according to the training dataset with some parameters set. Measurement Metrics The developed models will be evaluated and compared on the following.

Measurement metrics: accuracy, precision, recall, F1 score. Therefore, the term is all in a set of metrics that will be understood based on the efficiency of depression detection through social media data.

Performance Comparison: The evaluation metrics will compare model performances and identify the best performance model for diagnosing depression from social media.

4.5 Implementation and Results

Model Implementation Uses libraries like Scikit-learn, TensorFlow, or Keras to implement the models in a computer language like Python. To ensure repeatability and future updates, make sure the code is modular and thoroughly documented.

Analyse the outcomes that were acquired from the model assessments. Outline the results in a logical fashion, stressing the advantages and disadvantages of each model.

Practical Application talks about how the study's conclusions might be used in real-world situations. Considering potential problems and solutions, describe how the models might be applied in real-world situations to identify sadness using social media data.

4.6 Chapter Summary

Chapter 4 focused on the concept of research and implementation strategy for detecting depression in social media data using integrated machine learning. This was done by collecting data from Kaggle and other sources as a depression dataset, handling missing values, data normalisation, tokenization, and guaranteeing consistency in the analysis. It provides details of the training and testing phases for some machine learning models, including Support Vector Machine, Recurrent Neural Network, Naive Bayes, model parameter sets, batch sizes, and how model size impacts performance. The chapter also addresses performance assessment techniques applied to review the efficiency of such models in depression diagnosis, thus setting a base for the proper comparison and analysis of data in future chapters.

REFERENCES

- Ahuja, R., & Banga, S. (2019). Sentiment analysis in the prediction of mental health. *International Journal of Engineering and Advanced Technology*, 8(3), 138-141.
- Aleem, S., Huda, N.u., Amin, R., Khalid, S., Alshamrani, S.S., & Alshehri, A. (2022). Machine Learning Algorithms for Depression: Diagnosis, Insights, and Research Directions. *Electronics*, 11, 1111. <https://doi.org/10.3390/electronics11071111>
- Babu, N.V., & Kanaga, E.G.M. (2022). Sentiment Analysis in Social Media Data for Depression Detection Using Artificial Intelligence: A Review. *SN COMPUT. SCI.* 3, 74.
- Becken, S., & Gnoth, J. (2020). Sentiment analysis in tourism: A case study of New Zealand. *Tourism Management*, 81, 104-125.
- Bieliński, A., Rojek, I., & Mikołajewski, D. (2023). Comparison of Selected Machine Learning Algorithms in the Analysis of Mental Health Indicators. *Electronics*, 12(21), 4407. <https://www.mdpi.com/2079-9292/12/21/4407>
- Blei, D. M., Ng, A. Y., & Jordan, M. I. (2003). Latent Dirichlet Allocation. *Journal of Machine Learning Research*, 3, 993-1022.
- Bokolo, B. G., & Liu, Q. (2023). Deep Learning-Based Depression Detection from Social Media: Comparative Evaluation of ML and Transformer Techniques. *Electronics*, 12, 4396. <https://www.mdpi.com/2079-9292/12/21/4396>
- Calvo, R. A., & D'Mello, S. K. (2010). Affect detection: An interdisciplinary review of models, methods, and their applications. *IEEE Transactions on Affective Computing*, 1(1), 18-37.
- Calvo, R. A., Milne, D. N., Hussain, M. S., & Christensen, H. (2017). Natural language processing for mental health applications using non-clinical texts. *Natural Language Engineering*, 23(5), 649-685.
- Choudhury, M. D., Gamon, M., Counts, S., & Horvitz, E. (2013). Predicting depression via social media. *Proceedings of the Seventh International AAAI Conference on Weblogs and Social Media*, 128-137.

- Coppersmith, G., Dredze, M., Harman, C., & Hollingshead, K. (2014). From ADHD to SAD: Analyzing the language of mental health on Twitter through self-reported diagnoses. *Proceedings of the 2nd Workshop on Computational Linguistics and Clinical Psychology: From Linguistic Signal to Clinical Reality*, 1-10.
- De Choudhury, M., Gamon, M., Counts, S., & Horvitz, E. (2013). Predicting postpartum changes in emotion and behaviour via social media. *Proceedings of SIGCHI Conference on Human Factors in Computing Systems*, 3267–3276.
- Depression: Twitter Dataset + Feature Extraction Retrieve from: <https://www.kaggle.com/datasets/infamouscoder/mental-health-social-media>
- Friedman, J., Hastie, T., & Tibshirani, R. (2001). The elements of statistical learning (Vol. 1, No. 10). New York: Springer series in statistics.
- Guntuku, S. C., Ramsay, J. R., Merchant, R. M., & Ungar, L. H. (2019). Language of ADHD in adults on social media. *Journal of Attention Disorders*, 23(12), 1475-1485.
- Guntuku, S. C., Yaden, D. B., Kern, M. L., Ungar, L. H., & Eichstaedt, J. C. (2017). Detecting Depression and Mental Illness on Social Media: An Integrative Review. *Current Opinion in Psychology*, 23, 43–49.
- Haque, U. M., Kabir, E., & Khanam, R. (2021). Detection of child depression using machine learning methods. *PLoS One*, 16(12), e0261131.
- Hochreiter, S., & Schmidhuber, J. (1997). Long short-term memory. *Neural Computation*, 9(8), 1735-1780.
- HosseiniFard, B., Moradi, M. H., & Rostami, R. (2013). Classifying depression patients and normal subjects using machine learning techniques and nonlinear features from EEG signal. *Computer Methods and Programs in Biomedicine*, 109(3), 339-345. <https://doi.org/10.1016/j.cmpb.2012.10.008>
- Joshi, M. L., & Kanoongo, N. (2022). Depression detection using emotional artificial intelligence and machine learning: A closer review. *Materials Today: Proceedings*, 58, 217-226.
- Kohavi, R. (1995). A study of cross-validation and bootstrap for accuracy estimation and model selection. In *Ijcai* (Vol. 14, No. 2, pp. 1137-1145).

- Kuo, A. T., Chen, H., Kuo, Y. H., & Ku, W. S. (2023). Dynamic Graph Representation Learning for Depression Screening with Transformer. *arXiv preprint arXiv:2305.06447*.
- Kohavi, R., & Provost, F. (1998). Glossary of terms. *Machine Learning*, 30(2-3), 271-274.
- McCallum, A., & Nigam, K. (1998). A comparison of event models for Naive Bayes text classification. In *AAAI-98 workshop on learning for text categorization* (Vol. 752, pp. 41-48).
- Moreno, M. A., Jelenchick, L. A., Egan, K. G., Cox, E., Young, H., Gannon, K. E., & Becker, T. (2011). Feeling bad on Facebook: Depression disclosures by college students on a social networking site. *Depression and Anxiety*, 28(6), 447-455.
- Nahar, V., Hossain, M. S., & Azam, S. (2012). Association rule mining to detect factors which contribute to heart disease in males and females. *Expert Systems with Applications*, 40(4), 1086-1093.
- Reece, A. G., & Danforth, C. M. (2016). Instagram photos reveal predictive markers of depression. *EPJ Data Science*, 5, 1-12.
- Saha, K., Torous, J., Ernala, S. K., Rizuto, C., Stafford, A., & De Choudhury, M. (2020). A computational study of mental health awareness campaigns on social media. *Translational Behavioral Medicine*, 10(1), 11-20.
- Shatte, A. B. R., 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.
- Smys, S., & Raj, J. S. (2021). Analysis of deep learning techniques for early detection of depression on social media network-a comparative study. *Journal of trends in Computer Science and Smart technology (TCSST)*, 3(01), 24-39.
- Tariq, Q., Daniels, J., Schwartz, J. N., Washington, P., Kalantarian, H., & Wall, D. P. (2020). Mobile detection of autism through machine learning on home video: A development and prospective validation study. *PLoS Medicine*, 17(11), e1003271.
- Zeberga, K., Attique, M., Shah, B., Ali, F., Jembre, Y. Z., & Chung, T. (2022). A Novel Text Mining Approach for Mental Health Prediction Using Bi-LSTM

and BERT Model. *Hindawi Publishing Corporation*, 2022, 1-18.
<https://doi.org/10.1155/2022/7893775>