A Systematic Review on Application of Machine Learning Techniques for Depression Detection

Sandeep K Tiwari[1], Dhramandra Sharma[2], Anand kumar Singh[3], Madhavi Singh[4]

[1,2,3]Vikrant University, Gwalior, India

[4]Penn State University, StateCollege, USA

Abstract: Depression is one of the leading causes of suicide worldwide and a complex clinical entity. Unfortunately, a huge percentage of depression patients goes undiagnosed and, therefore remains untreated. Therefore, it poses challenges for clinicians regarding both accurate diagnosis and effective timely treatment. Moreover, social networks have been developed as a great point for its users to communicate with their interested friends and share their opinions, photos, and videos reflecting their moods, feelings and sentiments. This creates an opportunity to analyze social network data for user's feelings and sentiments to investigate their moods and attitudes when they are communicating via these online tools. Thus, depression analysis of a person can be done using the textual data collected from an online public source. Messages and posts posted by individuals with major depressive disorder on social media platforms can be analyzed to predict if they are suffering, or likely to suffer, from depression. In this context, machine learning can turn out to be an efficient and scalable method to help improve the management of this disease. Thus, several text preprocessing and textual-based featuring methods along with machine learning classifiers can be used to propose a generalized approach for depression detection using social media texts. This article is a review paper that presents the state-of-the-art of machine learning techniques in the detection of depression.

Keywords: Information extraction, depression detection, data mining, machine learning, textual-based featuring

I. Introduction: Depression, a mental disorder is amongst one of the most frequent mental health issue and a chief donor to the whole load of diseases worldwide[1]. According to the World Health Organisation, depression is the most prevalent mental disorder, which affects more than 300 million people worldwide [2]. Depression is also the leading cause of more than two-thirds of suicides every year [3]. Depression could be effecting recurrently or could be long-lasting and impacting patient's overall quality of life. A popular clinical method of detecting and diagnosing depression was psychiatric interview in which a series of questions were asked regarding its symptoms, duration of issues and symptoms and their affect on patient's daily life [4]. As a result, this disease could be classified as mild, moderate or severe. Though a self-denial among a fraction of patients and deprived identification of this disease many a times makes depression stay undiagnosed or incurable. The lack of diagnosis and treatment can further aggravate the condition, which could lead to reduced quality of life and, in acute cases, an inability to maintain employment.

Recently the social media platforms are available where people freely share their thoughts and express their feelings could be a vital source for monitoring health issues and trends. Various posts on the social-media platforms like Instagram, Facebook, etc. helps researchers to examine and explore the multi-aspects of mental health, concerns and human behavior. Researches' focusing mental depression shows that tweet posted by persons with depressive disorder can be used for predicting cases and chances of depression in those individuals [5]. This allowed mental health researchers and clinicians to collect a vast range of data at a rapidly growing rate. Many social media datasets are available for depression detection from Twitter, Facebook, Reddit, we chat, an electronic diary, etc.

Most studies on the depression detection with use of social media posts and messages follows a featuring approach based on text data or a person's descriptive analysis. Textual-based featuring focuses on the linguistic features of the social media text, such as words, Part of Speech, n-gram, and other linguistic characteristics. Whereas, the descriptive approaches utilize descriptions like patients age, gender, job, salary, use of drugs, smoking, etc. All of these features are used as an input to the detection model.

An effective method that boomed a decade ago to analyse exemplary dataset is Machine Learning (ML). ML techniques utilize and evolves through set of advanced statistical and probabilistic approaches constructed in a way that pertain the ability to learn from the examples itself without human intervention. It helps to detect, identify, and determine various patterns in the data accurately to make more accurate predictions on the data generating source

(e.g. more accurate diagnosis and prognosis). These techniques have proven to be beneficial in various medical diagnosis such as COVID detection [6], cancer detection [7], acne detection [8], etc. for classification, object detection, regression and segmentation related tasks for extracting various vital knowledge from the datasets in order to achieve an automatic intelligent system. Within health fields such as bioinformatics, ML has led to significant advances by enabling speedy and scalable analysis of complex data [9]. Not only this, these investigative techniques are also being explored on the mental health data. This leads to a potential improvement in understanding patient outcomes as well getting their psychological state and its management. Fig. 1 shows the depression detection using the textual data.
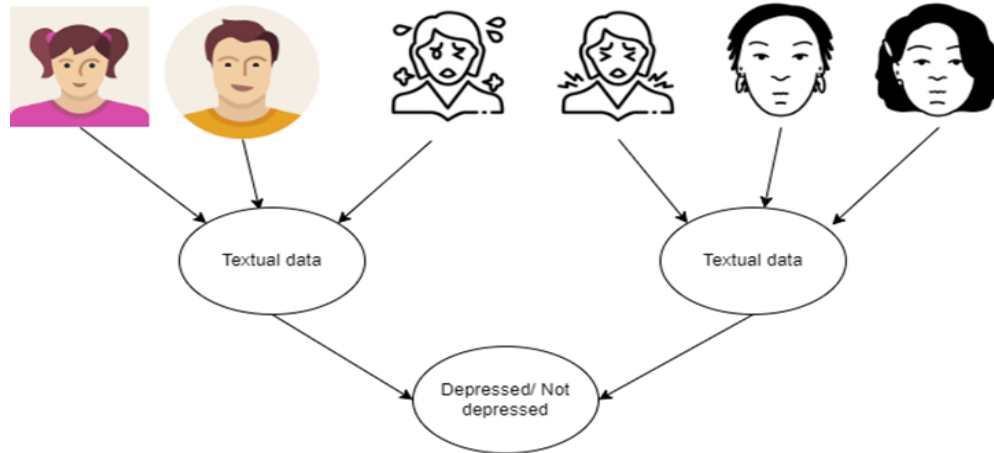

Figure1: Detecting Depression using the Textual Data.

Machine learning is a subset of artificial intelligence. It focuses mainly on the designing of system thereby allowing them to learn and make predictions based on some experience which is data in case of machines. Machine learning (ML) enables computer to act and make data driven decisions rather than being explicitly programmed to carry out certain task these programs are designed to learn and improve over time when exposed to new data [10]. The process starts with loading good quality data and then training machines by building machine learning models using the data and different algorithms. Before the application of machine learning techniques on social media texts for detecting depression several pre processing steps are performed such as Removal of punctuation, numbers, and stop words, word and spelling error corrections, etc.

After pre-processing the next step is to use machine learning models. It has been observed that machine learning techniques have proved to be give highly accurate results in such depression detection frameworks. Different models are present in literature such as linear regression, logistic regression, support vector machine, random forest, neural network etc. All these important methods are been explained in detail in the upcoming section. Most recent works find some keywords in the textual data such as 'depression', 'diagnosis' etc. in to order to determine the mental health. However less work is done in situations where such key works are absent.

In this survey paper, we have analysed the state of the art of various ML or Deep Learning (DL) schemes used in the literature. We have also discussed the various datasets available for the same. Moreover, at the end some of the key research gaps and the future research directions are also discussed to motivate the researchers and readers. The manuscript is organized as follows: section 2 gives an overview on the various machine learning techniques. Section 3 emphasize on the Literature review and organizes the different works done in this domain. Section 4 discusses the availability of the datasets used in this domain and Section 5 highlights the key research gaps and future research directions. The conclusion are drawn at the end of the manuscript.

II. Machine Learning Techniques:
ML techniques can be broadly classified into four categories namely, supervised, unsupervised, semi-supervised and reinforcement learning techniques [11]. Figure 2 shows the pictorial representation of the same.

(1) **Supervised Learning**- Supervised learning is that learning in which the model is trained on a labelled data set. A labelled dataset is having both the input features and the target or output values given. Supervised learning as the name indicates the presence of a supervisor as a teacher. Basically, supervised learning is a

learning in which train or test the machine using data. In supervised learning, data with known labels are used to train a model that can predict the label for unforeseen data or new data, such as classifying an email as spam based on previously seen labelled email examples.

Supervised learning algorithm consist of two types of technique- Classification and Regression

(a) **Classification:** It is a Supervised Learning task where output is having defined labels (discrete value). It can be a binary (yes or no, true or false, this or that) or a multi class classification problem. In a **binary** classification problem prediction is about either 0 or 1; yes or no but in case of **multi class** classification, model predicts more than one classes.

(b) **Regression**: It is a Supervised Learning task where output is having continuous value.

(2) **Unsupervised Learning**- It is a type of machine learning algorithm used to draw consequences from data set consisting of input data without output values. It is the training of machine using data that is neither classified nor labelled and algorithm to act on that data without any supervision. It means that teacher is not present in this type of learning. In contrast, unsupervised learning utilises mathematical techniques to cluster data in order to provide new insights, for example mapping topics of conversation in web forums. Unsupervised learning consists of two types of techniques Clustering and association rule mining.

(3) **Semisupervised Learning:** Moreover, there are semi-supervised learning techniques in which models are developed based on a combination of both labelled and unlabelled data. Such techniques are useful in enhancing supervised models through the use of unlabelled data, as labelled datasets may be scarce or expensive.

(4) **Reinforcement Learning**- It is a type of machine learning algorithm where an agent learns to behave in an environment by performing actions. In this learning agent decides what action is performed for a task, according to the action's agent got rewards by the environment. These rewards may be positive or negative.
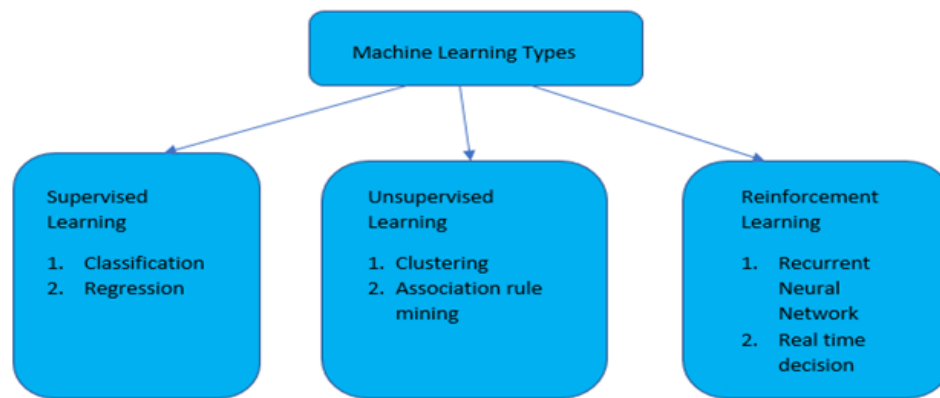
2. Figure 2: Machine learning Types

Machine Learning Algorithms-

There are various machine learning algorithms in which some important machine learning algorithms are as follows:

(1). Logistic Regression: Logistic Regression is one of the classification algorithms, used to predict a binary values in a given set of independent variables (1 / 0, Yes / No, True / False). To represent binary / categorical values, dummy variables are used. For the purpose of special case in the logistic regression is a linear regression, when the resulting variable is categorical then the log of odds are used for dependent variable and also it predicts the probability of occurrence of an event by fitting data to a logistic function.Ageneralised linear model, it was originally developed by Nelder and Wedderburn and further improved by Hastie and Tibshirani. Generalised linear models use non-normal dependent variables and thereby overcome a major limitation of linear models, i.e., continuous and normally distributed dependent variables. The dependent variables in LR are either unordered or ordered polytomous, and the independent predictor variables are either interval/ratio or dummy variables.

(2). Linear and multiple linear regression: Linear regression is a supervised learning algorithm; it performs a regression task. It describes a linear relationship between input variable and output variable. It is used for predictive analysis. The link between the criteria, or scalar response, and the several predictors, or explanatory factors, is modelled using a linear approach called linear regression. The conditional probability distribution of the response given the values of the predictors is the main focus of linear regression.

Multiple linear regression is a supervised learning algorithm which is used for regression. By applying an algorithm to the data, it describes the relationship between two or more independent variables and a single dependent variable.

(3). Polynomial Regression : Polynomial regression is a type of linear regression in which the relationship between the independent variable and dependent variable is describes as an mth degree of polynomial. It fits the nonlinear relationship between the dependent and independent variable. It is used for curvilinear data. Least squares is used to fit polynomial regression. Regression analysis's objective is to predict the value of a dependent variable y in relation to an independent variable x.

(4). Decision tree : The most effective and well-liked technique for categorization and prediction is the decision tree.Each internal node in a decision tree represents a test on an attribute, each branch a test result, and the tree's overall shape is similar to a flowchart and each leaf node (terminal node) a class label.Decision Tree is a tree structured framework. Decision Tree(DT) is a white box type of ML algorithm. Primarily, it is used for classification. However DT can also be used for regression. It works on the principle called "Decision-making logic". The selection tree is a distribution-loose or non-parametric method, which does now no longer depend on possibility distribution assumptions [12]. Figure 3 represents the branches, leafs and decision node in a decision tree.
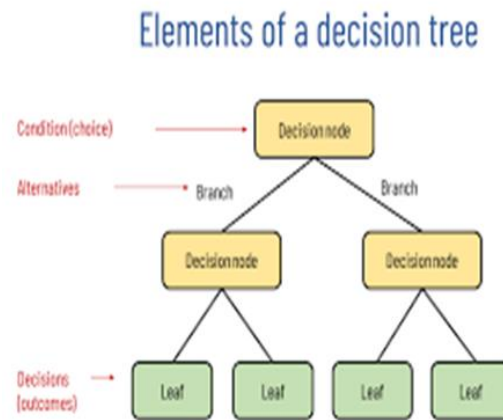

Figure 3: Representation of branches, leafs and decision node in a decision tree.

Decision trees will handle high dimensional knowledge with sensible accuracy.The DT was developed by Quinlan supported Hunt's algorithm, and may be a useful gizmo for exploring the cause-and-effect chain. It builds a tree-like call model for classification and prediction functions and is usually used as a base classifier for ensemble models.

(5). Random Forest: It belongs to the family of supervised learning approaches, appropriate for the classification and regression issues as well. Basic operating ideas behind this approach are multiple collections of tree-structured classifiers. Random forest is an ensemble learning method. it's used once size of dataset is massive and {also the} terribly sizable amount of input variables more or less lots of or thousands. The RF may be an ensemble of DT predictors wherever all call trees are trained severally victimization random vectors. The strength of trees and their correlation verify the error generalization of the RF. it's comparatively strong to outliers and noise.

(6). Support Vector Machine: In machine learning, support vector machines (SVMs, also support vector networks) are supervised learning models with associated learning algorithms that analyze knowledge used for classification and regression analysis.A Support Vector Machine (SVM) is a discriminative classifier formally outlined by a separating hyperplane. In different words, given labeled coaching knowledge (supervised learning), the algorithmic rule outputs AN optimum hyperplane that categorizes new examples. AN SVM model may be a illustration of the examples as points in space, mapped in order that the samples of the separate classes are divided by a transparent gap

Corresponding Author's E-mail ID: soet_sandeep@vikrantuniversity.ac.in

that's as wide as possible.

It works by separating totally different categories with a hyperplane, and so makes an attempt to maximise the separation distance from the hyperplane. Larger distances leads to lower errors and vise-versa by the classifiers. During this study, the SVM is employed along with the linear kernel (LSVM), that is mostly counseled for text classification. additionally to playacting linear classification, SVMs will with efficiency perform a non-linear classification, implicitly mapping their inputs into high-dimensional feature spaces.

(7). Artificial Neural Network: An artificial neural network (ANN) is comes from the idea and dealing of human brain and neurons. It simulate the functioning of a human brain. There are 3 layers Associate in Nursing input layer, one or additional hidden layers and an output layer as shown in Figure 4. there's net of neurons process units that frame an ANN, that successively contains inputs and outputs. For the provided inputs ANN learns to provide the required output. The coaching method is performed victimization backpropagation from which all the weights are learned. typically ANN are used for modelling non-linear problems.

They notice applications in several areas like health, classification, industries, and so on A feedforward artificial neural network that uses supervised learning, the MLP regularly computes and updates all the weights in its network to minimise error. In its 1st phase, referred to as the feedforward phase, the coaching information is forwarded to the output layer,

following that the distinction between this output and therefore the desired target (the error) is backpropagated to update the weights of the network within the second phase
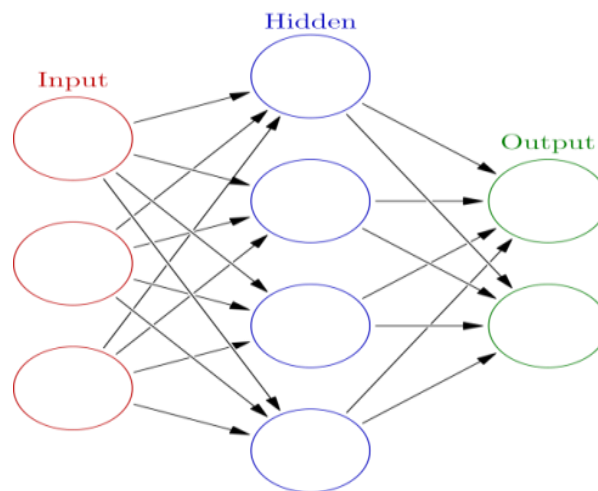
3. Figure 4: Artificial Neural Network.

III. Literature Review

Considering the feature extraction process, many methodologies have been adopted and fall into (a) language-based subgroups. Features of LIWC, articles, preambles, auxiliary verbs, adverbs, combinations, personal pronouns, impersonal pronouns, verbs and denials, semantics. (B) Network-based: Posts, replies, (c) Time-based: Emotional characteristics.

Much research on the automatic detection of depression symptoms has been done using artificial intelligence methods such as ML. Important areas of research include medical data such as fMRI signatures [13], results of depression questionnaires (such as DASS21 and DASS42), or clinical criteria for depression as defined by DSM-5 and ICD-10 [14]. Pain Analysis Interviews Data from clinical interviews using systems such as the Corpus-The Wizard of Oz (DAIC-WOZ) were also collected. DAIC-WOZ data includes video, speech, and text transcripts of participants, with or without stress. In recent years, researchers have also focused on detecting depression using text messages from social media platforms such as Twitter, Facebook, Reddit, and WeChat, even if the social media text is unaware of depression. Expected or rejected to help detect depression.

Shen et al.[15] Used descriptive features to determine the state of a normal or a depression patient and achieved

Corresponding Author's E-mail ID: soet_sandeep@vikrantuniversity.ac.in

an accuracy of approximately 85% whereas other authors such as Hassan et al. [16], Chen et al.[17], Islam et al.[18] and Fatima et al.[19] used textual features to determine the depression among humans. Someother used the images and textual data to determine the same among patients [20].

Most studies on the detection of depression using social media messages typically follow either a text-based functional approach or a person's description-based functional approach. Text-based features focus on the language features of social media text, such as words, part of speech, n-gram, and other language features. These features are then entered into the recognition model.

Choudhury et al. [21] claims depression to be a true test of individual and general well-being. A significant number of people experience the negative effects of depression, and only one department receives appropriate treatment each year. They also explored the possibility of using online networks to identify and analyze signs of serious depression in people. Through their postings on web-based social networks, they quantified behavioral credits that identify social involvement, emotions, dialect and semantic styles, self-system sensations, and antidepressant cues. The authors of [22] saw online networking as a promising public health tool and focused on using Twitter presentations to create predictive models of the imminent impact of childbirth on new mothers' behavior and propensity. We used Twitter posts to measure postpartum changes in 376 mothers using social engagement, emotional, informal community, and voice style measurements. Bachrachetal. [23] investigated how a user's activity on Facebook is distinguished from the user's identity as measured using the standard five-element model. They analyzed the relationship between the user's ID and the properties of the Facebook profile. For example, the size and thickness of your friendship network, the number of photos sent and attended, the number of meetings enrolled, and the number of times users have tagged your photos.

O'Deaetal. [24] examines that Twitter is increasingly being sought as a way to detect psychological well-being, such as depression and suicidal tendencies in the population. According to their research, both human programmers and programmed machine classifiers could be used to detect the level of concern for suicide tweets. Nguyen et al. [25] Using machine learning and statistical strategies, disposition, psycholinguistic techniques, and substance themes removed from posts created by individuals in these groups, in depression and control groups..

Ortigosa et al. [26] presented a new emotion testing strategy on Facebook. This suggests extracting data about user rating extremism (positive, bias, or negative) from user-written messages, as reflected in user-written messages. It then displays the user's default end edge and recognizes major passionate changes. These studies demonstrate that a person's behavior and mental state can be determined by social media sites. This has motivated and helped researchers move in this direction of mental health detection.

Rissola et al. 2019 [27] used an SVM model to measure the semantic proximity between text posts and a set of words currently associated with depression. Maxim Stankevich (2018) [28] also used SVM and CNN model to automatically detect depression signs from textual messages of Russian social network.
Wolohan et al. [29] with SVM ML model found Linguistic patterns of depressed Reddit users and Cacheda et al. (2019) [30] used the RF Model to detect depressed subjects and nondepressed subjects based on features defined from textual, semantic, and writing similarities. Guozheng Rao et al. (2020) [31] used both the sources of data Reddit and CLEF eRisk 2017 and applied different ML models. They used posts representation models for identifying depressed individuals, which was more accurate and efficient than general early depression detection models. Cultural and language barrier is alsways a considerable when dealing with large populations texual data. In different cultures, different languages are used by people to communicate with each other. As a result, people try to express their emotions, symptoms of disease and sensitive information through unique terms and expressions that belong to their own language and cultures.
Researchers performed a predominantly supervised learning-based approach after labeling the data as depressed or non-depressive. Classification algorithms such as Support Vector Machine (SVM), Logistic Regression (LR), Multilayer Perceptron (MLP), K-nearest Neighbor (KNN), and Random Forest (RF) are commonly used to classify depressed posts. Will be. In addition, some studies also use topic modeling and clustering algorithms. Both of these are unsupervised learning approaches that identify depression in textual data. Deep learning is a field of machine learning that makes it possible to carry out different types of research to identify depression. Deep learning techniques such as long short-term memory (LSTM) [18] and convolutional neural networks (CNN) [20] are also used.

Supervised learning-based algorithms and text mining approaches have been able to achieve good results. Most studies have conducted experiments based on depression-related vocabulary, syntax, and semantic characteristics collected from user contributions. According to some research, networking features provide additional support for identifying depressed users in online forums. In summary, computational models developed using supervised learning classifiers and text mining approaches were able to achieve 70-90% range

accuracy.

In addition, the supervised learning algorithm was the most commonly used method in previous studies of the classification of depression based on extracted features. Classification tasks and knowledge extraction processes can be extended based on more relevant characteristics such as demographics and personal health information without limiting and customizing common categories of social media. Although some studies are available to identify depression, deep learning mechanisms are rarely used in text functions.

Table 1: An overview of related work from 2017 to 2021.

Authors	Year of Publication	Features	Methodology	Maximum Accuracy
Shen et al. [15]	2017	Descriptive	MDL, MSNL, WDL, NB	85%
Hassan et al. [16]	2017	Textual	SVM, NB, ME	91%
Chen et al. [17]	2018	Textual	LSTM	Present the results in several graphs
Islam et al. [18]	2018	Textual	DT, KNN, SVM, Ensemble	NA
Fatima et al. [19]	2019	Textual	MLP, SVM, LR	91.63%
Rissola et al. [27]	2019	Textual	SVM	ERDE
Cacheda et al. [30]	2019	Textual, Semantic, and Writing similarities	RF	NA
Lin et al. [20]	2020	Visual and Textual	CNN	88.4%
Kim et al. [32]	2020	Textual	CNN, XGBoost	75.13
Alsagri et al. [33]	2020	Textual	SVM, NB, DT	82.5%

IV. Datasets and Data Sources

There are various sources of data set, especially the online sources. Here this section tells the key datasets available in literature and most popular among them. Many authors such as [15][20] used the public dataset from Twitter, whereas Kim et al. [32] used public data from Reddit. Some other authors have collected private data from Twitter [16] [33] , WeChat [17], Facebook [18], and Reddit [19]. Images [20] are also used as input for some depression schemes and others used the CLEF eRisk dataset, used by Rissola et al.[27]. This is publicly available corpus. CLEF eRisk data consists of a set of documents posted by Reddit users. This have two groups of users called as depressed and non-depressed (normal). Table 2 gives the summary of the eRisk dataset.

Table 2: Summary of eRisk 2017 dataset.

	Train		Train	
	Positive	Control	Positive	Control
# of Subjects	83	403	52	349
# of Documents	30, 851	264, 172	18, 706	217, 665
Avg. # of Documents/Subject	371.7	655.5	359.7	623.7

Corresponding Author's E-mail ID: soet_sandeep@vikrantuniversity.ac.in

V. Future Research Gaps and Research Directions:

This study aims to determine whether machine learning could be effectively used to detect signs of depression in social media users by analysing their social media posts. It will benefit the stigmatised people by identifying hidden textual cues in online social media platforms related to depression and to provide support and guidance to overcome depression.

A lot of previous work has been done by researchers for depression detection using texual data from social media sites and other sources. However still there many research gaps, that need to be adddressed for a best framework to determine the depression from texual data. There are various problems that researches might faced during the classification as follows:

1. There are few individual studies that have applied SVM, KNN, Decision Tree and Ensemble separately. There are no well-known studies that have combined all these techniques together at same dataset to investigate the variations in technique-based findings. Also, there is no signifcant study that has applied the abovementioned machine learning techniques on Facebook data for depression detection.

2. There is a lack of proper framework that include efforts to build a depression detection framework using textual features from social media posts using proper preprocessing [34][35], featuring and ML classification.

3. In research on depression detection using social media texts, custom datasets are frequently created but not made publicaly available. Therefore, there is a research gap that publicly available datasets from different sources should be used. This would help to generalise the approach for depression detection in social media texts, the datasets include only text messages and exclude any emoticons, emojis, pictures, videos and web links that are commonly part of social media messages.

4. Most authors have not addressed the issue of overfitting that generally arises when collecting depression data from social media messages. A model might perform poorly on datasets it was not trained on due to overfitting.

5. Moreover, efforts should be made to deal with the problem of imbalanced data samples, which can negatively impact the performance of classification models.

6. There are majority of studies focused on identifying depression-related posts from support forums, in which a high number were related to identifying self-harm or suicidality related behaviour. Many of these experiments were conducted on small sets of samples, which brings into question their robustness and ability to be generalised.

7. Most of the time, without having much knowledge on depression disorder, forum users are often confronted with misleading information or misdiagnose themselves. Therefore, to easily identify relevant symptoms from online forum data available as textual representations, we need a good quality analyser able to identify variations in behavioral and emotional patterns.

8. Semantics-based understanding of the post is required since forum posts are lengthy in nature, some consist of contradictory semantics, which misleads the extraction of the overall meaning of the post, making it hard for machines to understand depression disorder accurately.

9. Severity assessments: since depression disorder has different severity levels, treatments vary accordingly requiring the need to identify unique textual characteristics for each stage. Only after identifying the degree of severity can a forum moderator or clinician provide reliable recommendations.

VI. Conclusion

Depression, a mental disorder is amongst one of the most frequent mental health issue and a chief donor to the whole load of diseases worldwide. Recently the social media platforms are available where people freely share their thoughts and express their feelings, could be a vital source for monitoring health issues and trends. Machine learning plays a vital role in diagnosis of depression using texual data, images, posts, etc. Thus, this survey paper gices the state of the art of various ML or Deep Learning (DL) schemes used in the literature for depression detection. It also discuss various datasets available for the same. The key research gaps and the future research directions are also discussed to motivate the researchers and readers.

References:

[1] Dosani, S.; Harding, C.; Wilson, S. Online Groups and Patient Forums. *Curr. Psychiatry Rep.* 2014, *16*, 1–6.

[2] Priya, S. Garg, N.P. Tigga, Predicting anxiety, depression and stress in modern life using machine learning algorithms, Procedia Computer Science 167 (2020) 1258–1267.

[3] N.S. Srimadhur, S. Lalitha, An end-to-end model for detection and assessment of depression levels using speech, Procedia Computer Science 171 (2020) 12–21.

[4] Lloyd-Williams, M. Diffificulties in diagnosing and treating depression in the terminally ill cancer patient. *Postgrad. Med. J.* 2000, *76*, 555–558.

[5] H.S. Alsagri, M. Ykhlef, Machine learning-based approach for depression detection in Twitter using content and activity features, IEICE Transactions on Information.

[6] Roberts, M. Common pitfalls and recommendations for using machine learning to detect and prognosticate for COVID-19 using chest radiographs and CT scans. *Nature Machine Intelligence* 3, no. 3 (2021): 199-217.

[7] Saba, T. Recent advancement in cancer detection using machine learning: Systematic survey of decades, comparisons and challenges. *Journal of Infection and Public Health*, (2020), *13*(9), 1274-1289.

[8] Nain, G., & Gupta, A. (2022). Automatic selection algorithm for region of interest of acne face image compression. *Evolutionary Intelligence*, 1-7.

[9] Luo J, Wu M, Gopukumar D and Zhao Y (2016) Big data application in biomedical research and health care: a literature review. Biomedical Informatics Insights 8, 1–10.

[10] https://www.ibm.com/cloud/learn/machine-learning

[11] https://www.simplilearn.com/tutorials/machine-learning-tutorial/types-of-machine-learning

[12] https://en.wikipedia.org/wiki/Decision_tree

[13] J.R. Sato, J. Moll, S. Green, J.F. Deakin, C.E. Thomaz, R. Zahn, Machine learning algorithm accurately detects fMRI signature of vulnerability to major depression, Psychiatr. Res. 233 (2) (2015) 289–291.

[14] B. Ojeme, A. Mbogho, Selecting learning algorithms for simultaneous identification of depression and comorbid disorders, Procedia Computer Science 96 (2016) 1294–1303.

[15] G. Shen, J. Jia, L. Nie, F. Feng, C. Zhang, T. Hu, T.-S. Chua, W. Zhu, Depression detection via harvesting social media: a multimodal dictionary learning solution, in: Proceedings of the 26th International Joint Conference on Artificial Intelligence, Melbourne, Australia, 2017, pp. 3838–3844.

[16] Hassan, et al. Sentiment analysis of social networking sites (SNS) data using machine learning approach for the measurement of depression, in: Proceedings of the ICTC, South Korea, 2017.

[17] Y. Chen Sentiment analysis based on deep learning and its application in screening for perinatal depression, in: Proceedings of the IEEE 3rd International Conference on Data Science in Cyberspace (DSC), Guangzhou, China, 2018, pp. 451–456.

[18] Islam . Depression detection from social network data using machine learning techniques, Health Inf. Sci. Syst. 6 (1) (2018) 8, https://doi.org/10.1007/s13755-018-0046-0.

[19] Fatima , Prediction of postpartum depression using machine learning techniques from social media text, Expet Syst. 36 (4) (2019) e12409, https://doi.org/10.1111/exsy.12409.

[20] C. Lin, P. Hu, H. Su, S. Li, J. Mei, J. Zhou, H. Leung, SenseMood: depression detection on social media, in: Proceedings of the 2020 International Conference on Multimedia Retrieval, Dublin, Ireland, 2020, pp. 407–411.

[21] De Choudhury, M, Predicting depression via social Media. In: ICWSM, vol. 13. 2013. p. 1–10.

[22] Choudhury M, Counts S, Horvitz E. Predicting postpartum changes in emotion and behavior via social media. In: Proceedings of the SIGCHI conference on human factors in computing systems. New York: ACM; 2013.

[23] Bachrach Y, Personality and patterns of Facebook usage. In: Proceedings of the 4th annual ACM web science conference. 2012. New York: ACM.

[24] O'Dea B, . Detecting suicidality on Twitter. Internet Interv. 2015;2(2):183–8.

[25] Nguyen T, et al. Afective and content analysis of online depression communities. IEEE Trans AfectComput. 2014;5(3):217–26.

[26] Ortigosa A, Martín JM, Carro RM. Sentiment analysis in Facebook and its application to e-learning. Comput Hum Behav. 2014;31:527–41. Ríssola, E.A.; Losada, D.E.; Crestani, F. Discovering latent depression patterns in online social media. *CEUR Workshop Proc.* 2019, *2441*, 13–16.

[27] Ríssola, E.A.; Losada, D.E.; Crestani, F. Discovering latent depression patterns in online social media. *CEUR Workshop Proc.* 2019, *2441*, 13–16.

[28] Stankevich, M.; Latyshev, A.; Kuminskaya, E.; Smirnov, I.; Grigoriev, O. Depression detection from social media texts. *CEUR Workshop Proc.* 2019, *2523*, 279–289.

[29] Wolohan, J.T.; Hiraga, M.; Mukherjee, A.; Sayyed, Z.A. Detecting Linguistic Traces of Depression in Topic-Restricted Text: Attending to Self-Stigmatized Depression with NLP. In Proceedings of the First International Workshop on Language Cognition and Computational Model, Santa Fe, NM, USA, 20 August 2018; pp. 11–21.

[30] Cacheda, F.; Fernandez, D.; Novoa, F.J.; Carneiro, V. Early detection of depression: Social network analysis and random forest techniques. *J. Med. Internet Res.* 2019, *21*, e12554.

[31] Rao, G.; Zhang, Y.; Zhang, L.; Cong, Q.; Feng, Z. MGL-CNN: A Hierarchical Posts Representations Model for Identifying Depressed Individuals in Online Forums. *IEEE Access* 2020, *8*, 32395–32403.

[32] Kim , A deep learning model for detecting mental illness from user content on social media, Sci. Rep. 10 (1) (2020) 11846

[33] Alsagri . Machine learning-based approach for depression detection in Twitter using content and activity features, IEICE Transactions on Information and Systems E103.D (8) (2020) 1825–1832,

[34] P. Norvig, How to write a spelling corrector, 2016.

[35] W. Etaiwi, G. Naymat, The impact of applying different preprocessing steps on review spam detection, Procedia Computer Science 113 (2017) 273–279.