Available online at www.sciencedirect.com

ScienceDirect

Procedia Computer Science 230 (2023) 138–149

Procedia Computer Science

www.elsevier.com/locate/procedia

3rd International Conference on Evolutionary Computing and Mobile Sustainable Networks (ICECMSN 2023)

# FFO-ABC DepressioGuard: A Hybrid Classification Framework for Social Media Depression Detection

U Ananthanagu [a]* , Ebin P M [b], Nivedita Manohar Mathkunti [c]

[a,b,c]*Assistant Professor, Alliance University, Bengaluru 562106, India*

**nagu.anantha@gmail.com, pmebin195@gmail.com, niveditagg@gmail.com*

**Abstract**

In the phase of digital communication, social media sites have developed into a major centre for individuals to express their ideas, emotions, and views. Amidst this influx of user-generated content, mental health conditions, notably depression, have garnered increasing attention due to their pervasive impact on individuals and societies. Early detection and intervention are crucial in managing and preventing its adverse effects. As a consequence an innovative machine learning (ML) textual data classification framework is designed to detect depression through social media streams, employing a Firefly-Optimized Support Vector Machine (FFO-SVM) and Artificial Bee Colony (ABC) Optimized SVM classifiers. Initially, data collection and preprocessing are performed, followed by feature extraction using Time Frequency-Inverse Document Frequency (TF-IDF). After extraction of features classification is performed using FFO-SVM and ABC-SVM classifiers. To tune the parameters of SVM to work efficiently, FFO and ABC are employed. The proposed framework combines the power of ML with the optimization capabilities of the ABC and FFO Algorithm to enhance classification accuracy. Through extensive experimentation and analysis, the framework's performance is evaluated using relevant metrics. Results indicated that proposed classification techniques outperformed conventional methods, showcasing its effectiveness in handling the complexity of depression detection from social media data.



*Keywords:* Depression detection, Machine learning, Firefly-Optimization, Artificial Bee Colony Optimization, Time Frequency-Inverse Document Frequency.


10.1016/j.procs.2023.12.069

## 1. Introduction

Social media platforms have experienced enormous development in the past few years in terms of information exchange and have integrated themselves into daily life. Such live data depicts a person's everyday activities by revealing their personal space, behavioural traits, and emotions. Due to the extensive availability of private data on social media, it is crucial for the procedure of identifying mental illness in individuals [1]. Depression is such a common kind of mental disease that experts estimate that between 6 and 21% of people worldwide suffer from it. Particularly in undeveloped and impoverished nations with limited access to healthcare, the situation for mental diseases is significantly worse. Worldwide estimates place the prevalence of depression in the young population at 5%, with 20% of cases being classified as moderate depression, partial symptoms, or probable depression [2, 3]. Since mental health services are not often available, depression and other mental problems frequently go untreated. Sometimes the illness is not even recognized or given a diagnosis.

Finding depression symptoms in patients is the most crucial step in the treatment of any illnesses, especially mental disorders like depression. The chance of reducing depression symptoms and underlying disorders can be greatly increased by identifying depressive symptoms early on before they are assessed and treated. Additionally, it can lessen detrimental effects on social, economic, and academic life as well as health and well-being. Potential diagnostic methods need the patient to report to a mental health care professional and require patients or carers to describe the symptoms. These methods are sometimes expensive for the underprivileged populations of impoverished nations. This component of diagnostic approaches frequently results in patients' depression not being recognised and leaves them untreated. A low or no-cost diagnostic technique has to be created without the patient being physically present [4-6].

Social media has been a widely utilised tool for the exchange of ideas and information during the past 10 years. A tiny passage of text might reflect a person's mental state. Because of this, practitioners may infer a lot about a person's mental health and wellbeing from their Facebook, Twitter, or Instagram posts. Numerous books have been written about the anatomy of social interactions in relation to social media, including breakups in relationships, quitting smoking and drinking, celebrating suicide, sexual harassment, and mental diseases. The possibility to utilise social media as a tool for depression diagnosis is made possible by the steadily rising use of social media worldwide [6-8].

Numerous research have experimented with various methodologies and put forward numerous frameworks with regard to the diagnostic procedures necessary for the identification of depression. Most research focuses on explicitly identifying depression and making a precise diagnosis of the condition. Additionally, the temporal component of social media activity has been taken into account with explicit detection. Patients with depression commonly have an insomniac propensity, which causes them to use social media more frequently and remain up late. SVM and CNNs have both performed well in this challenge, with SVM achieving the maximum accuracy of 85% and 88.5%, respectively [9–11]. Our objective is to identify a sarcastic tone or implicit comment about the condition rather than an open declaration of sadness or the passage of time, which are not the best indicators to rely on. Even if they are not expressing it explicitly, a person may still be suicidal.

Deep learning, a kind of machine learning, may provide remarkable results. Deep learning techniques are very useful since they work from start to finish and can instantly recognize feature representations from raw data. The availability of high-quality data is frequently required for the development of high-performing, accurate ML applications. Data replacement, data modification, and data deletion are other names for the process of modifying or removing inaccurate data from a data file [12]. Processing incorrect numbers, missing values, and data reasonableness detection are all included. The type of encoding strategy employed has a big impact on how well the classification process works. One-hot encoding is a common technique for categorical data processing. For machine learning (ML) models to be more efficient in detecting instances of insider data leakage, categorical variables required to be translated into a format. It just draws attention to the variables included in the features in order to prevent improper interpretation of the correlations between independent variables [13, 14]. The process of proportionately scaling the data such that all values fall within the required range is known as data normalization.

Data is separated into a variety of categories using the M-SVM non-probabilistic binary classifier. Classification is achieved by the binary division SVM by categorizing input data. Text, images, audio, and other forms of data, as well as those that are asymmetrically distributed, may all be classified using SVM [15]. The foundation of M-SVMs is the idea that two data classes can be divided by a margin on each side of a plane. A learning model based on supervised learning is constructed using an M-SVM, which trains the model to classify training data using prelabeled labels.

In this paper, an innovative ML textual data classification framework is designed to detect depression through social media streams, employing a FFO-SVM and ABC Optimized SVM classifiers. Initially, data collection and preprocessing are performed, followed by feature extraction using TF-IDF.

## 2. Proposed System

Details on the dataset, its annotation, and the study methods are provided in this section. The three components of our research are data processing, feature extraction, and data categorization. The data were gathered, manually labelled as binary and ternary labels, and then pre-processed in the first module to get the data ready for classification. In the second module, we used the TF-IDF feature extraction approach on the produced data and the binary and ternary data alone were subjected to ML classification algorithms. Finally, a hybrid paradigm for classifying depression that combines FF0-SVM and ABC-SVM is suggested. Our research modules are shown in Fig 1.

A collection of label data is used to train SVM. SVM's primary benefit is that it can be applied to regression as well as classification problems. For the purpose of trying to distinguish between or categorize any two classes, SVM creates a decision limit, which is a hyperplane. SVM is also utilized in image classification and identifying objects.

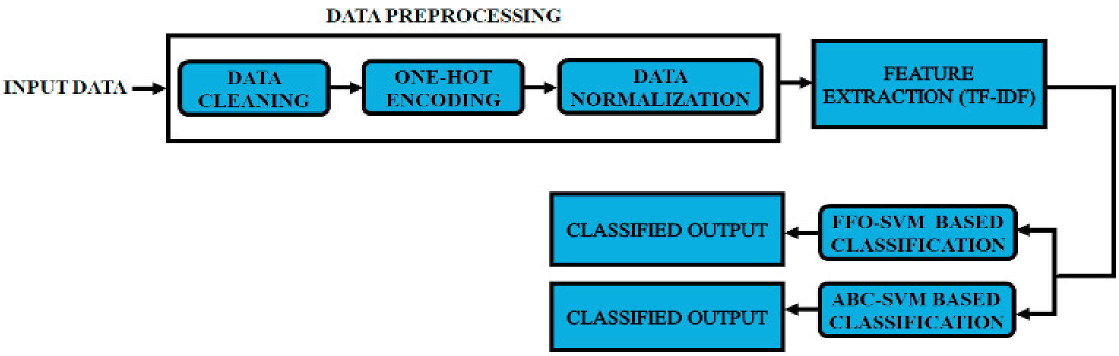


Fig. 1. Proposed System

### 2.1. Data Cleaning

Data cleaning is the first step in every ML project and crucial step in making that the dataset is free of incorrect or deceptive data. Whether manually using data manipulation tools or programmatically with computer software is possible. Data cleaning is a broad category of tasks that includes preparing data for analysis. An increasing number of people are interested in creating effective and productive data cleansing solutions due to their importance across several sectors. Yet, data is rarely reliable in practice because of inaccurate inputs from manual data curation or unintentional mistakes from robotic data collection or generation operations.

Real-world datasets typically feature gaps and inconsistencies caused, for example, by broken sensors or human error, which might affect machine learning systems based on those datasets. Schema is typically used to clean structured data at scale that must adhere to integrity standards, denial limits, and functional dependencies. The accuracy of the machine learning pipeline and data validation techniques have recently been improved. These approaches do not, however, address the pressing issues of system fairness and system robustness against adversarial input.

Since mistakes and discrepancies in the data used for training may prevent models from finding patterns, cleaning up data to guarantee the integrity of the learning data is a crucial step for sustaining the model's efficacy. Resources can be lost, productivity can be reduced, and marketing money can be lost due to ineffective data management. In order to accomplish effective filtering, the aforementioned method is utilized to segment images that have been warped by high density impulsive noise.

Processing incorrect values, missing values, and data rationality detection are all parts of data cleaning, which is often referred to as data replacement, data alteration, and data elimination. Additionally, inaccurate data must be fixed or deleted from the data file. The following is the main data cleansing technique, which is shown in Fig 2.

Step 1: Find the missing data in step one. On rare occasions, errors or human errors during the data collection process might leave one or more data elements with a blank value. By altering the positioning condition to a null value, the positioning is accomplished.

Step 2: Since there aren't many missing values in the sample data set utilized for this study, the method of instantly deleting this row of data is employed to deal with them.

Step 3: Even when missing numbers are eliminated, there will still be some obvious mistakes in the data. The whole row of data implicated in the issue will be removed since the information in these two columns truly has to be preserved in numerical form.

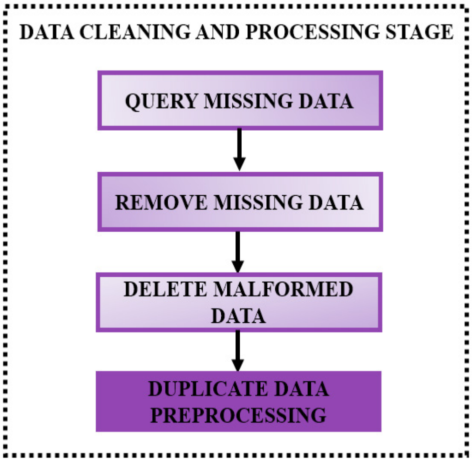


Fig. 2. Data Cleaning Process

Data cleaning tasks include finding missing values and content errors in addition to actions like eliminating duplicate records. The model's performance will be impacted by any logical errors or excessive data repetition. After searching, no indication of a logical error is found, and there is little duplication of the material. Therefore, no corrective action is taken. However, encoding the classified data is required for further classification and to stop data leakage. The correction or deletion of inaccurate, damaged, improperly organized, duplicate, or missing data from a dataset are achieved by data cleaning.

### *2.2. One hot Encoding*

The type of encoding strategy employed has a big impact on how well the classification process performs. One-hot encoding is often used in categorical data management. Categorical variables need to be formatted in order for ML models to discover instances of insider data breaches more effectively. To avoid wrong interpretation of correlations between independent variables, it solely focuses attention to the factors that make up features. A good encoding method for classification issues is one-hot encoding. The label encoding technique that was previously employed is less complex than the one-hot encoding strategy, however ordering issues might arise since some ML algorithms might not comprehend specific integer values. These ordering issues are resolved using the one-hot encoding technique. In one-hot encoding, each category value is converted into a new column, and the label values are transformed to either a digital representation of (1 or 0). ML systems cannot directly process textual data. Numbers must be present in the data. For the purpose of this study's data preparation, the email text was encoded as one-hot vectors. One-hot encoding, which uses a sparse vector with one member set to 1 and all other elements set to 0, is a common technique for recording strings with a finite number of values. When utilizing one-hot encoding, high cardinality will provide high dimensional feature vectors. One-hot encoding, however, is a well-liked encoding method due to its simplicity.

One-hot encoding is often employed with models that have excellent smoothing properties and works well for tweets or sentences with few repeated components. One-hot encoding is often used in neural networks, which need input to be in the discrete range of [0,1] or [-1,1]. With the exception of a single 1 in a cell that is used to precisely identify a word, one-hot vectors are 1 N matrices (vectors) that are entirely composed of 0s. One hot encoding allows for a more expressive representation of categorical data. An example word range of [good, good, horrible] might be represented in 3 such encodings [0, 0, 1], as shown in Table 1.

Table 1. One-hot encoding

| | |
|---|---|
| 1 | 0 |
| 1 | 0 |
| 0 | 1 |

The information that the IoT gathers contains a lot of discrete data. It is viewed as transforming the labels in string format because its data is in string format. Without first converting the aforementioned data from labels to numbers, it cannot be used in the design. One-Hot Encoding is the answer to this problem. N states must be encoded during the procedure. One number represents each state; hence n digits are needed to represent n states. The other numbers are 0, and the state's corresponding number is 1, when that state of the result is present. It is also obvious that if a feature has m possible values, it will undergo one-hot encoding and turn into m binary features. Additionally, only one of these qualities may be present at once, and they are all exclusive of one another. As a result, data will become limited. One-hot encoding not only addresses the issue of data characteristics but also adds the necessary dimensions to the experimental data set.

*2.3 Data Normalization*

The process of proportionately scaling the data such that all values fall within the required range is known as data normalisation. The two main advantages of normalising data cleanup rapid encoding Data normalisation following preparation of the data set Examples of data test outcomes Create an ID network model.

Additionally, because the maximum value will alter when new data is uploaded, it will need to be updated. This experiment use the standard deviation normalisation technique to accomplish the normalisation of the data's standard deviation. The calculation for normalising standard deviation is shown in formula (1),

$$\chi_{norm} = \frac{\chi - \mu}{\sigma} \tag{1}$$

If the initial sample in the data set for the IoT is x, the standard variation is, and the median value is. The distribution of data ought to be typically close to the Gaussian statistical to have a favorable normalizing impact when using standard deviation standardization. In the normalization process, rules are applied to a set of data. Every one of these rules converts the data into a particular arrangement known as a normal form.

*2.4 Feature extraction by TF-IDF*

The technique of turning raw data into numerical characteristics that can be handled while keeping the details in the original data set is known as feature extraction. Comparing to using ML on the raw data directly, it produces superior outcomes. For the separation of ham and spam in SMS text messages, feature extraction and selection are crucial. TFIDF will be employed for this phase. The Vector Space Model frequently uses TFIDF, notably in the IR domain, which includes text mining. A word's importance to the entire corpus of the document can be determined statistically. The term frequency is merely determined as a function of how many times a word occurs in the text, and it is often normalised in the positive quadrant between 0 and 1 to remove bias towards long publications. Punctuation is eliminated during tokenization, and all text is converted to lowercase in order to create the TFIDF word index. If a term appears more frequently in a text, its importance is indicated by the first two letters, TF, or term frequency. As a result, the greater TF indicates that the phrase is more likely to be used frequently in the corresponding papers.

Inverse Document Frequency, or IDF, also determines how seldom a phrase or term appears in the documents. The concept behind IDF is that a word is not seen to be an appropriate choice for representing the text if it appears often throughout the whole dataset since it may be stop words or prevalent, generic phrases. Therefore, just a small number of words not the full dataset are pertinent to those papers. The TF-IDF examines the significance of words in document databases and corpora in addition to the relevance of words in individual texts. In this way, the word frequency in the text will proportionately raise the weight of the words but will later be countered by the word frequency in the corpus. This essential TF-IDF feature presupposes that some words occur more frequently than others in the document as a whole. As a result, Eq. 2 illustrates the relevance of a word to a document.

$$F-IDF=\frac{Term\ Frequency}{Document\ Frequency} \quad (2)$$

## 2.5. *FFO Based SVM Algorithm*

The creation of the SVM model is greatly influenced by the choice and optimization of model parameters. One of the key elements in ensuring the correctness and dependability of the SVM model's inversion results is the use of suitable model parameters. The SVM model comprises two parameters: a fundamental parameter and a parameter associated with the kernel function. Finding the appropriate parameters is challenging since earlier techniques of parameter selection mostly relied on subjective experience and grid search.

In recent years, researchers have successfully selected SVM variables using a globally optimized adaptive clever search technique capability. As a result, in this research shown in fig. 3, the firefly method is utilized to determine the SVM model's ideal parameters. When an SVM model is not the best parameter, this work uses the firefly method to optimize the SVM model and then addresses the issues of "overlearning" and "underlearning."

The following are the precise steps used by the firefly method to optimize the SVM model parameters:

(1) Setting the beginning location of each firefly using a range of its initial characteristics and giving the initial values for every firefly's characteristic $C_{min}\sim C_{max}$ , $\sigma_{min}\sim\sigma_{max}$ , the maximum number of repetitions
(2) Computing brightness value as a brightness update.
(3) Acquiring each firefly's domain set and picking at random firefly i from the set as its movement direction.
(4) Determining out the SVM train value following location change. The position will be replaced if the revised SVM's training value is superior to the prior one; else, nothing will change.
(5) Modifying the decision domain's dynamic; Evaluating the cycle of the aforementioned procedure. The cycle procedure will be stopped once the maximum search times are reached. Up until the end condition is met, the SVM will be trained using the process' optimum solution.

## *2.6. ABC Optimized SVM Method*

The candidate approaches to the feature selection challenge are expressed by bit vectors, contrasting the candidate solutions to optimization problems, which can be represented by vectors with actual values. A bit vector of size N, where N is the total number of characteristics, is connected to each food source. The amount of characteristics that need to be assessed is represented by a location in the vector. If the feature is included in the subset to be assessed, it means that the value at the corresponding location is 1. If the value is 0, on the other hand, it means that the feature does not belong to the subset that is being evaluated. Each food source also records its quality (fitness), which is determined by the classifier's precision when utilizing the bit vector-indicated feature subset (Fig. 4) The penalty factor C and the kernel parameter are the main variables that influence how well the SVM classifier performs. For that, the following crucial parameters will be optimized using the ABC method:

Step 1: Preprocessing of raw data is step one. To accomplish the data process, follow the stages of the data preparation stated above.

Step 2: Initializing the ABC algorithm.

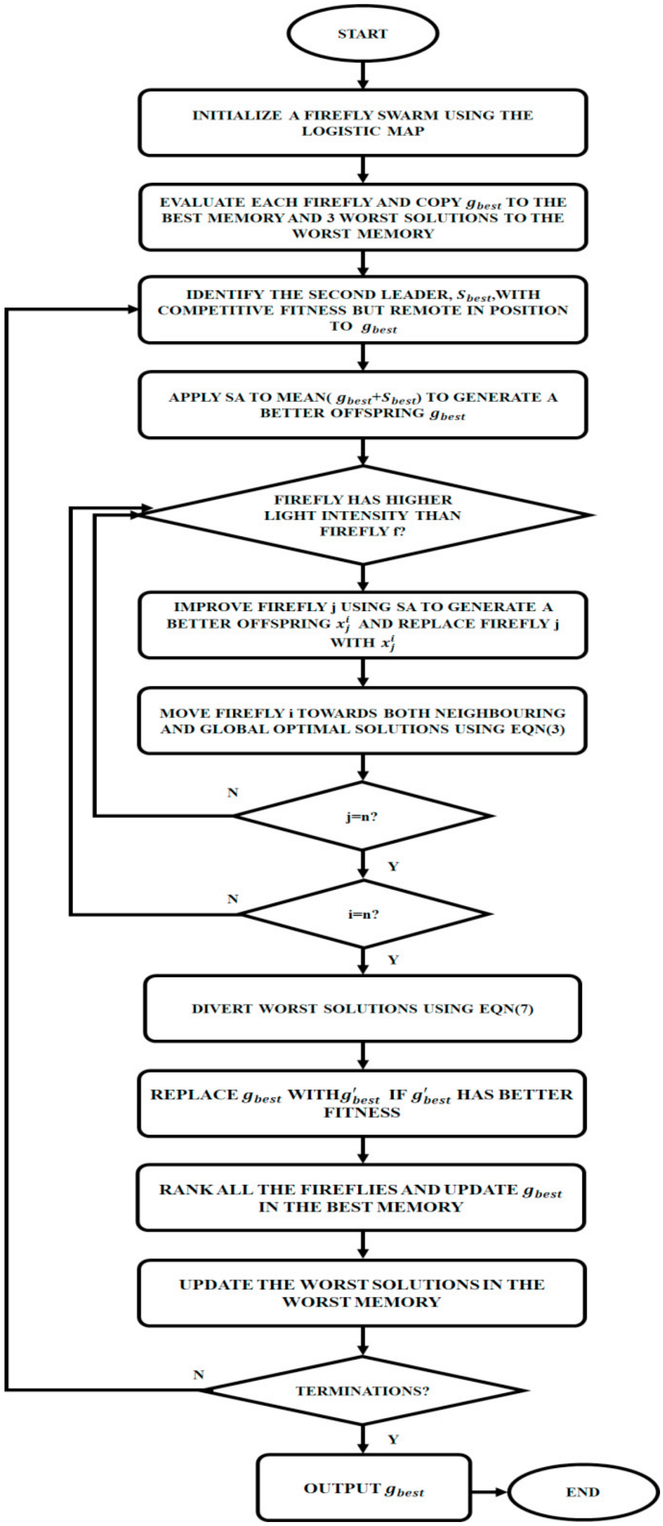


Fig. 3. FFO Flowchart

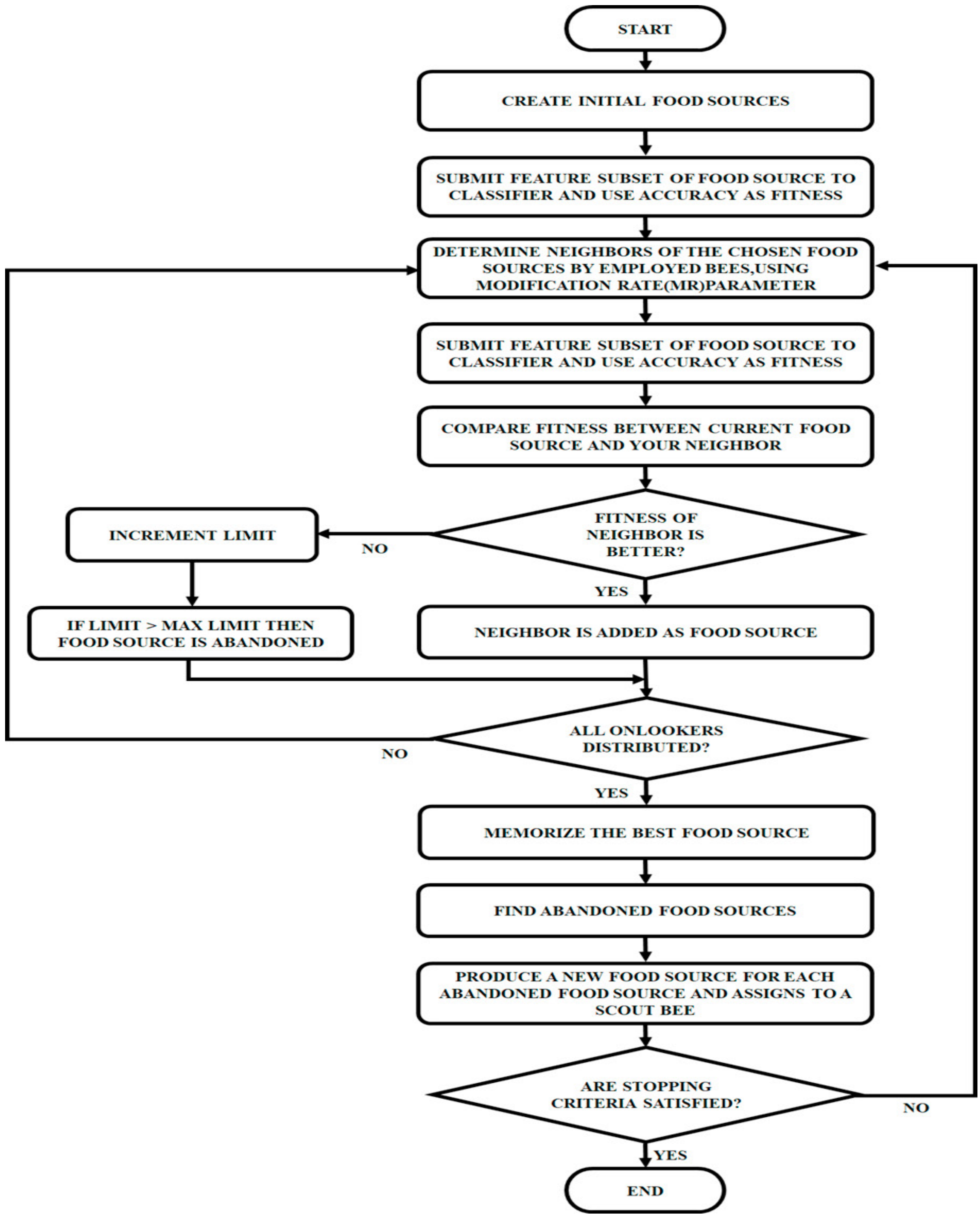


Fig. 4 ABC Flowchart

Step 3: Step 3: Defining the goal function. The processed development sub-dataset (70%) will be utilized as input information during the training phase of the model.

$$t = f_t(v_{acc}) = 1 - v_{acc} \qquad (3)$$

where t is the goal value of the optimization process and $v_{acc}$ range [0, +1], with 0 being the best.

Step 4: Training the model and optimizing its parameters. Use the ABC technique to improve the C and SVM model parameters. The nectar source fitness (3) states that all potential solutions can be optimized by the three honeybees' individual seeking efforts.

Step 5: Performance evaluation of the model categorization. The global ideal nectar source, which was determined using the optimal parameters C, in the preceding stage, is utilised to build the SVM model. The model's generalizability will next be assessed using the testing (30%) sub-database.

## 3. Results And Discussion

The following subsection analyses the outcomes of the suggested ML hybrid approach and contrasts it with different methods that were previously used to accomplish the objective using Python software. With regard to binary and ternary labelled data, experimental examination of deep learning models was reported. As previously stated in the parts previously, the depressive issue of classification was separated into two categories: FFO and ABC categorization. Each challenge was explicitly investigated with using a variety of feature extraction techniques and machine learning algorithms. FFO and SVM had the highest accuracy for the classification challenge out of all the classified approaches. The dataset is taken from Beck Depression Inventory (BDI) in which database of 20,000 patient records and 11 features are included in this paper.

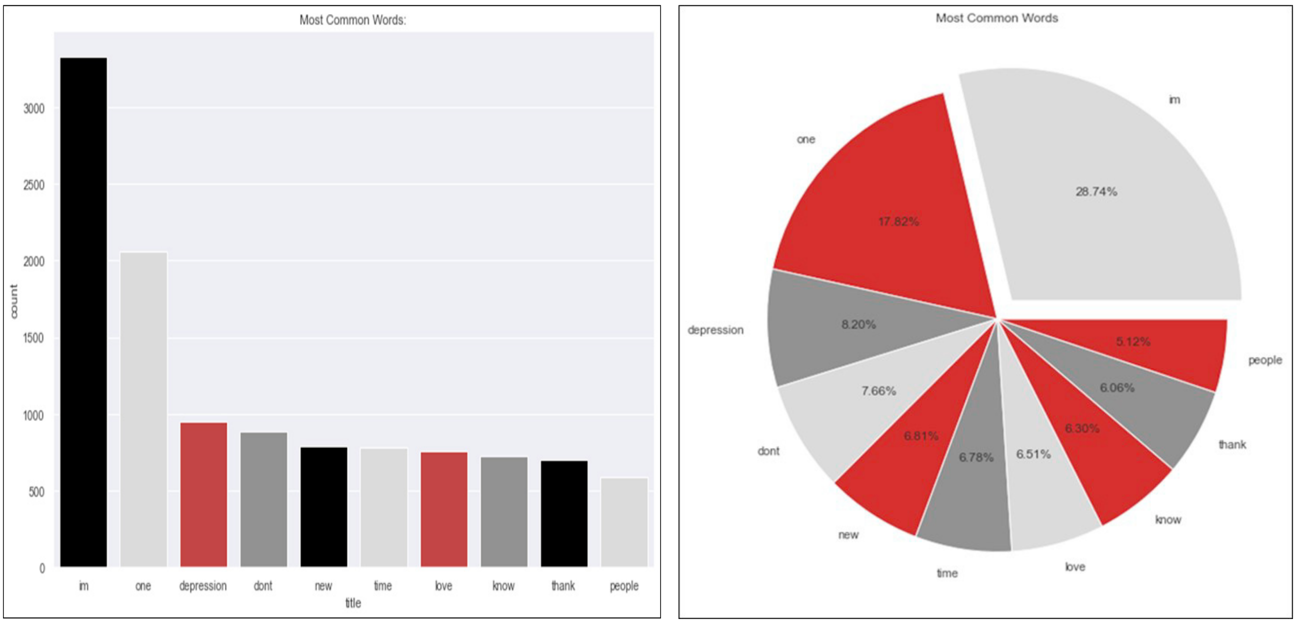


Fig.5. Display of most common words

As shown in Fig.5, there are twice as many single FA as hybrid FA utilized to address optimization issues. It demonstrated that very few academics had combined hybrid FA with another technique to address optimization issues. As a result, the hybrid algorithms outperform the single FA. Applications of hybrid FA appear to be superior than single FA since the other approaches enhance the FA's performance and processing speed. The single FA may need a lot of time at times to obtain the ideal parameter values. FA should thus be used with other approaches to speed up processing.

The study also aimed to improve the rate of true positive and true negative detection while lowering the rate of false positives and false negatives by training the machine learning classifiers. The TF-IDF confusion matrix is displayed in Fig. 6

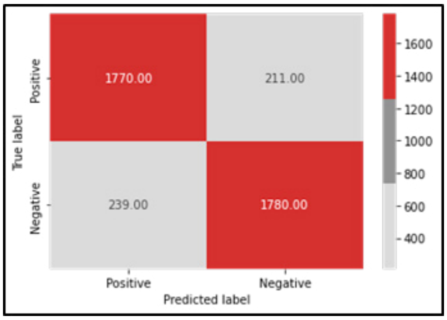


Fig. 6a). Confusion matrix (a) ABC (b) FFO

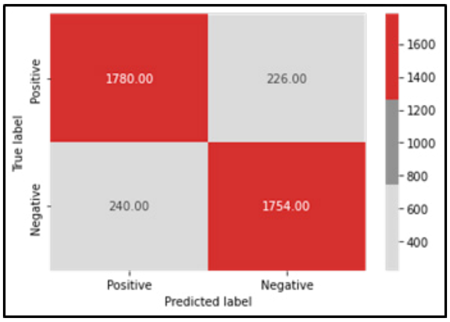


Fig. 6b). Confusion matrix (b) FFO

The fact that FFO-SVM functions effectively in highly dimensional domains and memory effective is one of its key features. By converting it to a space with greater dimensions where it is separated by linearity utilising the kernel method, it is also capable of handling irregularly recoverable data.

Table 2. Depression Detection based on TF-IDF

| Features | Precision | Recall | F1-Score | Accuracy |
|---|---|---|---|---|
| TF-IDF | 0.908 | 0.893 | 0.897 | 0.901 |

One inference from this is that the TF-IDF textual feature can contribute most satisfactorily to the detection of depressive tweets, while other modalities can offer further help. In this study, several feature sets for the objective of detecting depression in social media are investigated. The content of social media user's messages is used and researched how bag-of-words, embedding words, and bigrams function with data in order to categorize social media users.

To determine how good the algorithm is in classifying in compared to a one-hot encoding strategy, the ROC is calculated. The more top-left the curve is, the larger the area and the higher the ROC value. The average ROC curve values with one-hot encoding are shown in Fig. 7. It reveals that the ABC-SVM approach outperforms other ML algorithms, with a ROC value of 0.88 and FFO-SVM value is 0.87.

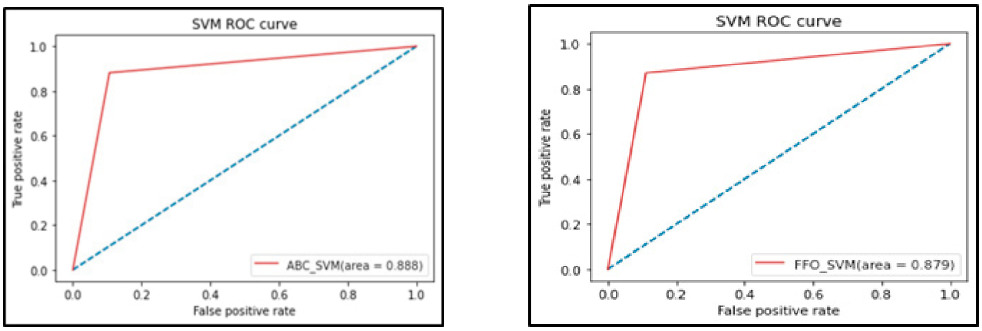


Fig. 7. ROC Curve

From the Fig. 8, it is clear that FFO-SVM shows higher accuracy than other classifiers like SVM [16-18], MRF [17] and DT [19].

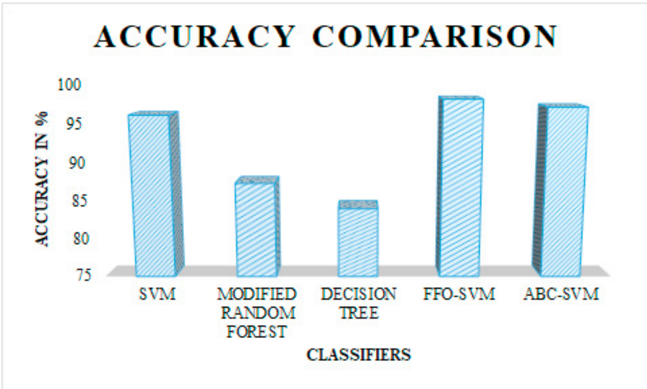


Fig. 8. Comparison of Accuracy

## 4. Conclusion

One of the psychological disorders is depression that has spread widely in the modern world, where mental health disorders have become quite common. According to WHO data, depression is the second most common factor contributing to the world's illness burden. Social media has shown to be a fantastic medium for individuals to express themselves in the emergence of such problems. Social media accounts may therefore reveal a lot about a user's emotional condition and mental health. This research proposes a unique framework for depression diagnosis from textual data using machine learning approaches, taking into account the high prevalence of the disorder. The effective feature approach known as TF-IDF is used to locate essential or, more specifically, uncommon terms in text data. Nearly all text data uses, including data mining from texts, systems for retrieving data, and classification. The domain experts manually annotated a dataset made up of social media to capture the implicit and explicit context of depression for this aim. By detecting the classifiers with an ROC of 0.88 for ABC-SVM and 0.87 for FFO-SVM, the FFO-SVM provides the best results.

## References


[1] Ananthanagu, U., Agarwal, P. (2023). A Systematic Review and Future Perspective of Mental Illness Detection Using Artificial Intelligence on Multimodal Digital Media. In: Nagar, A.K., Singh Jat, D., Mishra, D.K., Joshi, A. (eds) Intelligent Sustainable Systems. Lecture Notes in Networks and Systems, vol 578. Springer, Singapore. https://doi.org/10.1007/978-981-19-7660-5_4

[2] S. Md, H. Sathish, K. S. Koulini, A. Inamdar and U. Ananthanagu, "A Radical Approach To Depression Detection," 2022 IEEE 7th International conference for Convergence in Technology (I2CT), Mumbai, India, 2022, pp. 1-6, doi: 10.1109/I2CT54291.2022.9825306.

[3] J. Zhu et al., "An Improved Classification Model for Depression Detection Using EEG and Eye Tracking Data," in IEEE Transactions on NanoBioscience, vol. 19, no. 3, pp. 527-537, July 2020.

[4] S. Hurayn, A. M. Shetty, S. V. Deshpande, V. Gupta, U. Ananthanagu and A. Professor, "Analysis of Mental Illness using Twitter Data," 2021 International Conference on Forensics, Analytics, Big Data, Security (FABS), Bengaluru, India, 2021, pp. 1-4, doi: 10.1109/FABS52071.2021.9702565.

[5] Y. Xing et al., "Task-State Heart Rate Variability Parameter- Based Depression Detection Model and Effect of Therapy on the Parameters," in IEEE Access, vol. 7, pp. 105701-105709, 2019.

[6] H. Dibeklioğlu, Z. Hammal and J. F. Cohn, "Dynamic Multimodal Measurement of Depression Severity Using Deep Autoencoding," in IEEE Journal of Biomedical and Health Informatics, vol. 22, no. 2, pp. 525-536, March 2018.

[7] M. Trotzek, S. Koitka and C. M. Friedrich, "Utilizing Neural Networks and Linguistic Metadata for Early Detection of Depression Indications in Text Sequences," in IEEE Transactions on Knowledge and Data Engineering, vol. 32, no. 3, pp. 588-601, 1 March 2020.

[8] M. M. Tadesse, H. Lin, B. Xu and L. Yang, "Detection of Depression-Related Posts in Reddit Social Media Forum," in IEEE Access, vol. 7, pp. 44883-44893, 2019.

[9] X. Zhang, J. Shen, Z. u. Din, J. Liu, G. Wang and B. Hu, "Multimodal Depression Detection: Fusion of Electroencephalography and Paralinguistic Behaviors Using a Novel Strategy for Classifier Ensemble," in IEEE Journal of Biomedical and Health Informatics, vol. 23, no. 6, pp. 2265-2275, Nov. 2019.

[10] Z. Huang, J. Epps, D. Joachim and V. Sethu, "Natural Language Processing Methods for Acoustic and Landmark Event-Based Features in Speech-Based Depression Detection," in IEEE Journal of Selected Topics in Signal Processing, vol. 14, no. 2, pp. 435-448, Feb. 2020.

[11] Y. Ding, X. Chen, Q. Fu and S. Zhong, "A Depression Recognition Method for College Students Using Deep Integrated Support Vector Algorithm," in IEEE Access, vol. 8, pp. 75616-75629, 2020.

[12] G. Rao, Y. Zhang, L. Zhang, Q. Cong and Z. Feng, "MGL-CNN: A Hierarchical Posts Representations Model for Identifying Depressed Individuals in Online Forums," in IEEE Access, vol. 8, pp. 32395-32403, 2020.

[13] H. Cai, X. Zhang, Y. Zhang, Z. Wang and B. Hu, "A Case-Based Reasoning Model for Depression Based on Three-Electrode EEG Data," in IEEE Transactions on Affective Computing, vol. 11, no. 3, pp. 383-392, 1 July-Sept. 2020.

[14] Y. Li, B. Hu, X. Zheng and X. Li, "EEG-Based Mild Depressive Detection Using Differential Evolution," in IEEE Access, vol. 7, pp. 7814-7822, 2019.

[15] S. Venugopal, S. Y. Panale, M. Agarwal, R. Kashyap and U. Ananthanagu, "Detection of Malicious URLs through an Ensemble of Machine Learning Techniques," 2021 IEEE Asia-Pacific Conference on Computer Science and Data Engineering (CSDE), Brisbane, Australia, 2021, pp. 1-6, doi: 10.1109/CSDE53843.2021.9718370

[16] S. Tariq et al., "A Novel Co-Training-Based Approach for the Classification of Mental Illnesses Using Social Media Posts," in IEEE Access, vol. 7, pp. 166165-166172, 2019.

[17] K. -Y. Huang, C. -H. Wu, M. -H. Su and Y. -T. Kuo, "Detecting Unipolar and Bipolar Depressive Disorders from Elicited Speech Responses Using Latent Affective Structure Model," in IEEE Transactions on Affective Computing, vol. 11, no. 3, pp. 393-404, 1 July-Sept. 2020.

[18] C. Venkatesan, P. Karthigaikumar, A. Paul, S. Satheeskumaran and R. Kumar, "ECG Signal Preprocessing and SVM Classifier-Based Abnormality Detection in Remote Healthcare Applications," in IEEE Access, vol. 6, pp. 9767-9773, 2018.