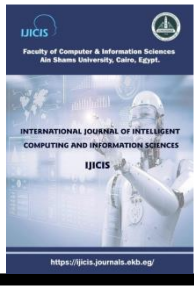
IJICIS, Vol.24, No.3, 58-69
DOI :10.21608/ijicis.2024.319767.1354

International Journal of Intelligent Computing and Information Sciences

https://ijicis.journals.ekb.eg/



AUTOMATED DEPRESSION SCREENING OF CLINICAL TRANSCRIPT OPTIMIZED BY GREY WOLF OPTIMIZER

Alwan Atta*
Scientific Computing Department, Faculty of Computers and information sciences, Ain Shams University & Galala University, Cairo, Egypt
alwan.atta@cis.asu.edu.eg

Dina El Sayad
Scientific Computing Department, Faculty of Computers and information sciences, Ain Shams University, Cairo, Egypt
dina.elsayad@cis.asu.edu.eg

Doaa Ezzat
Scientific Computing Department, Faculty of Computers and information sciences, Ain Shams University, Cairo, Egypt
doaaezzat@cis.asu.edu.eg

Safaa Amin
Scientific Computing Department, Faculty of Computers and information sciences, Ain Shams University, Cairo, Egypt
safaa.ameen@cis.asu.edu.eg

Mahmoud El Gamal
Scientific Computing Department, Faculty of Computers and information sciences, Ain Shams University, Caieo, Egypt
mahmoud.elgamal@cis.asu.edu.eg

Received 2024-09-10; Revised 2024-09-10; Accepted 2024-09-30

Abstract: *Depression diagnosis depends on the transcripts obtained from clinical interviews during mental health assessment. This hybrid model proposes Decision Tree (DT) and Grey Wolf Optimizer (GWO) for feature selection to enhance depression detection. The proposed model makes use of the Decision Tree, thereby effectively grasping the temporal context embedded in clinical interview transcripts. It can detect significant linguistic features of depressive symptoms. This model will also include the integration of GWO to optimize feature selection for added strength and efficiency. In this way, a hybrid model featuring a high F-score macro of 0.83 could be derived, which proved its effectiveness in detecting depression with accuracy. This will contribute to the literature by enhancing computational methods for measuring mental health, and this model has immense potential to apply in clinical practice. Beyond academic research, the applications of such a hybrid model are huge and promising in a clinical setting. The model automates the diagnosis of depression to assist healthcare professionals in informed and timely intervention.*

Keywords: *Swarm Intelligence, Decision Tree, Clinical Interviews, Automated Depression Diagnosis.*

1. Introduction

***Corresponding Author**: Alwan Atta
Scientific Computing Department, Faculty of Computer and Information Science, Ain Shams University & Galala University, Cairo, Egypt
Email address: alwan.atta@cis.asu.edu.eg

In recent years, ML and DL algorithms are increasingly adopted within the healthcare sector. First of all by the remarkable number of research works on improving diagnosis and treatment regarding mental health disorders. For example, since 2017, a mental health model has been trained in England. This program was designed and implemented by the National Institute of Mental Health (NIMHE) [1] as part of a national public mental health strategy. However, Australia was the first country to implement these programs in the year 2000. Currently, there are licensed programs in 25 countries with over three million persons trained worldwide.

While most chronic disease diagnoses rely heavily on laboratory tests and measures, diagnoses of mental health generally depend on self-reporting questionnaires designed to identify a pattern of emotions or social interactions, or through private interviews with psychologists. Major depression is the most prevalent mental health condition and is estimated to affect 3.8% [2] of the global population or approximately 280 million people worldwide. Depression is more than a sad mood or a transient emotional response to a problem in living; it is distinct from the ups and downs experienced by everybody. At worst, depression can be a relatively frequent and seriously debilitating health condition that may disrupt performance at work and school, damage family life, and even result in suicide-the fourth leading cause of death [3] among people between the ages of 15 and 29 years.

Current methods of diagnosing depression mainly depend on patient self-reports or symptom severity clinical assessments. There are major challenges faced by the existing methods of diagnosis, including denial of illness by the patients or low sensitivity of the tests. Subjective bias and inaccuracies make diagnosis of depression a time-consuming process. Recently, ML approaches have been applied to biosignals for detecting depression. However, there is still a gap between classification accuracy and practical application scenarios.

2. Related Work

A number of research works have been conducted with the aim of predicting or monitoring mental illness through natural language analysis to retrieve key features and indicators that show the best performance. As one of the most common mental disorders, a variety of data collection methods have been conducted in several literatures, including self-assessment questionnaires, self-reported diagnoses, participation in specialized online forums, and posts manually annotated.

Xiuzhuang Z. et al., [4] proposed a deep residual network with a Global Average Pooling (GAP) layer for training a deep regression model on visual depression data. This model predicted depression severity with accuracy while locating the most informative regions of the input facial image by generating a depression activation map. They also proposed multi-region DepressNet which learns jointly multiple local depth regression models on different facial regions to improve overall recognition performance.

On the other hand, the work of [5] Zhao, Z. et al. presents a new hierarchical attentionbased model that leverages the strengths of unsupervised learning and attention transfer in the estimation of the severity of depression from speech. An attention transfer process was used to estimate the severity of depression at the frame and sentence levels across tasks. They proposed an attention mechanism to train hierarchical autoencoders that can represent speech from persons with depression in an unsupervised manner.

Xiuzhuang Z. et al., [6] proposed a model that uses content-aware weights for the aggregation of the input video frames in order to provide a discriminative deep depression representation. This pooling approach for features effectively empowers or suppresses face images with varying poses and imaging conditions and has been considered superior compared to several state-of-the-art methods.

Salma A. et al. [7] designed a predictive model that can identify whether a user's tweet portrays depression; the authors used Random Forest, Naïve Bayes, AdaBoostM1, and Lib-linear classifiers for Arabic sentiment analysis. The performance result of the model showed an accuracy of 0.88 using the Lib-linear classifier.

The random forest classifier proposed by Rouzbeh R. et al. [8] reached an out-of-sample balanced accuracy of 0.768, increasing to 0.811 after the inclusion of participants' age and gender. Mental health mobile applications can continuously screen for depressive symptoms by collecting mobile usage statistics, in this respect, to aid in the initial diagnosis or monitor ongoing treatments.

Anupama R. et al. [9] proposed an early fusion network based on multi-level attention, and audio, video, and text patterns were combined into one with higher weights toward the audio pattern in order to predict the severity of depression. That means that the video data, which captures electrophysiological signals, is less important in this framework when trying to create better diagnostics.

Such is the case with Zhu J. et al. [10], where two technologies, EEG and eye-tracking, are compared; it is observed that EEG is more accurate than eye-tracking, though more expensive. CBEM also performed significantly better than the conventional classifiers using biosignals, achieving 82.5% accuracy for the eye-tracking dataset and 92.65% for the resting-state EEG dataset.

Norah A. et al. [11] proposed a model to predict symptoms of depression using ArabDep, which is a lexicon designed for the Arab community. Here, the idea integrates a semi-supervised approach-lexicon-based-with a supervised approach that is machine-learning-based for advice seeking about online forums on mental disorders. The lexicon-based performed at 78% accuracy, while the highest performing machine-learning classifier achieved an accuracy of 72%.

Xezonaki D. et al. [12] proposed the HAN model for depression detection, which introduced psycholinguistic information into the model to improve its performance and robustness. The HAN+L was ranked first according to the experiments in this paper on the DIAC-WOZ dataset, with an F-macro score of 0.69 and a UAR of 0.72.

Zhenxing X. et al. [13] have proposed a model that proposes subphenotypes of depression using machine learning methods from EHR. This model analyzes features around demographics, comorbidities, and medications from 11275 patients.

Alice O. et al. [14] presented a deep neural network for depression and emotion features recognition, called EmoAudioNet, using Convolutional Neural Networks. The network learns from the time-frequency representation of audio signals and the visual representation of their frequency spectrum.

The problem of class imbalance in datasets was utilized by Amita S. et al. [15] The authors employed different resampling strategies; therefore, the model studied biomarkers and self-reported depression through the XGBoost technique of machine learning; hence, its result was a balanced accuracy, precision, recall, and F1 scores over 0.90 with oversampling and over-under sampling.

Takashi N. et al. [16] proposed the method of reducing site bias and improving classification accuracy in the identification of patients with MDD. By taking part of the healthy control group to regress out the site bias, the SVM was the best for the classification of melancholic depressed patients versus healthy controls, with a classification accuracy of 73.3%.

Composite measures were applied to improve MDD detection performance. Predictive factors for diagnosis of MDD are identified as key variables such as disability, family history of mental illness, and stressful events by Yiye Z. et al. [17].

Kyoung-Sae N. et al. [18] demonstrated that future depression can be predicted according to survey data. The purpose of the predictive model will be early identification of subjects at risk who will develop depression, thus enabling early intervention.

Thalia R. et al. [19] emphasized that psychiatric diagnostics should depend on behavioral measures combined with methods of machine learning. Having developed a bagged decision tree classifier, they report an accuracy of 68%, thus showing that tools of machine learning could support diagnostic confidence.

The paper by Latif et al. [20] considered the task of separating the depression-relevant features from speaker-specific characteristics for speech-based depression detection. The authors introduced a new framework called speaker-disentanglement that is aimed at enhancing model robustness across different speakers.

Amir et al. [21] presented research into how the inclusion of therapist prompts can affect automatic depression detection from clinical interviews. Their research focused on validating this source of information in the area of depression detection. The authors managed to show that an increase in performance can be achieved by including therapist prompts.

Latif et al. [22] proposed a deep learning architecture called Speechformer-CTC that could optimally employ the strengths of Transformers and Connectionist Temporal Classification together for speech-based depression detection. The effectiveness of their proposed model was evident in the temporal dependencies of speech data that obtained state-of-the-art performance in detecting depression.

Yin et al. [23] introduced a deep learning-based approach that integrated the parallel CNN and transformer for depression detection from speech segments. The parallel-CNN module leverages local knowledge, while the transformer module does the task of extracting temporal sequential information through linear attention mechanisms. The results achieved on the Distress Analysis Interview Corpus-Wizard of OZ and Multi-modal Open Dataset for Mentaldisorder Analysis demonstrated superiority over competitive state-of-the-art approaches.

In this regard, Lorenzoni et al. [24] have done an experimental case study on the performance evaluation of various ML classifiers and NLP techniques for depression detection. More emphasis in their work has been given to data cleaning, pre-processing, feature selection, parameter tuning, and model choices. Experiments with the DAIC-WOZ dataset have given accuracy of approximately 84% with Random Forest and XGBoost models, outperforming state-of-the-art results in the literature.

Automatic clinical depression recognition system using speech features was done by Rejaibi et al. [25], in which MFCCs from speech signals were used as features, and depression recognition was made using a recurrent neural network.

Rai et al. [26] investigated different ways in which depression may be diagnosed using different modalities: voice, text, and facial expressions. Their system, MANOBAL, combined features in audio recordings, text, and facial expressions to predict depression and its severity. By utilizing transfer characteristics and increasing training labels, they had impressive F1 scores on the DAIC-WOZ dataset.

W. Zhang et al., [27] proposed a depression detection and assessment system using a multimodal fusion model. The applied modality text used sentence embeddings pre-trained and BiLSTM, whereas the audio modality relied on PCA and SVM, and the video modality relied on XGBoost. Their model outperforms baselines in both depression detection and assessment tasks.

Aghaei et al. [28] conducted passive detection of depressive states from participants by analyzing data in the form of videos, audios, and text from interviews. The method described here has much potential for the improvement of assessments on mental health and may open a better direction toward treatments in light of the growing mental health crisis.

One of the major challenges in depression detection is spotting the symptoms of mental illness from clinical interviews, as the symptoms of different mental illnesses are mostly overlapped. Previous work [29] has explored various ways to validate the legitimacy of claims for diagnosis with depression. Another major challenge for the researchers [30] is the lack of availability of legalized clinical data. While a lot of data from mental care warehouses are present, there is a major concern regarding patient privacy. For this reason, many experiments have been conducted in the DAIC-WOZ dataset, whose methods and results can easily be compared.

Until now, no work has explored the concept of swarm intelligence for depression detection. Swarm Intelligence techniques have gained wide acceptance in several engineering applications because of their simple structure, limited requirements of operators, quick convergence speed, and a good trade-off between exploration and exploitation phases. Due to optimal performance and efficiency, swarm intelligence algorithms have been one of the multidisciplinary approaches in recent years.

3. Methodology

This section addresses the applied tools and methodology for the proposed model before and after optimization, to predict the diagnosis state of each patient.

3.1. Text Preprocessing

Text cleaning and preprocessing [31] are the very first steps of NLP models, where improvement in quality and precision of text analysis can be enhanced. The following techniques are applied in the proposed model for text cleaning and preprocessing:

- Lowercasing: All texts should be converted into lower case to maintain uniformity.
- Tokenization: Breaking down text into smaller units such as words or sentences.
- Removing Punctuation: Most of the time, punctuation marks do not convey any meaningful information; therefore, they can be removed.
- Removing Stop Words: Filtering out common words that do not carry significant meaning.
- Stemming and Lemmatization: Reducing the words to their base or root form.
- Removing Extra Whitespace: Trim extra whitespace to normalize text.

3.2. Feature Extraction

To convert the text after cleaning into features that helps in the classification task we used Bag of Words (BOW) [32] model to create a vocabulary of all unique words available in the corpus, then represents each document as a vector of word occurrences. In this vector, each dimension corresponds to one of these words, and the value inside each dimension will be how many times that particular word occurred in the document. This, in turn, makes text data amenable to structured data conversion.

3.3. Grey Wolf Optimizer (GWO)

GWO [33] basically takes inspiration from the leadership hierarchy in nature and hunting strategies of grey wolves. It considers three fundamental hunting processes, including searching for prey, encircling prey, and attacking preys. This algorithm simulates the leadership hierarchy by utilizing four types of grey wolves, including alpha, beta, delta, and omega. Besides, GWO solves three classical engineering design problems and one from the optical engineering field to demonstrate its practical applicability. The results obtained for these classical engineering design problems and real-world applications show that the proposed approach has effectiveness in complex problems with unknown search spaces.

As grey wolves encircle prey during the hunt. In order to mathematically model encircling behavior the following equations are proposed:

$$\vec{D} = |\vec{C} . \vec{X}_p(t) - \vec{X}(t)| \quad (1)$$

$$\vec{X}(t + 1) = \vec{X}_p (t) - \vec{A} . \vec{D} \quad (2)$$

Where t indicates the current iteration, $\vec{A}$ and $\vec{C}$ are coefficient vectors, $\vec{X}_p$ is the position vector of the prey, and $\vec{X}$ indicates the position vector of a grey wolf.

The vectors $\vec{A}$ and $\vec{C}$ are calculated as follows:

$$\vec{A} = 2\vec{a} . \vec{r}_1 - \vec{a} \quad (3)$$

$$\vec{C} = 2 . \vec{r}_2 \quad (4)$$

Where components of $\vec{a}$ are linearly decreased from 2 to 0 over the course of iterations and r_1, r_2 are random vectors in [0, 1].

3.4. Decision Tree

The decision tree [34] is a supervised machine learning algorithm used in both classification and regression tasks. They basically are flowchart-like structures in which every internal node represents a test on an attribute; with the various branches emerging from it, representing the outcome of the test; and each leaf node denotes a class label or a predicted value. And it follows the following steps:

- Partitioning the data: This algorithm starts off by choosing the best attribute for partitioning the data according to a chosen metric measure.
- Node Creation: Based on the attribute's value, subsets of the data are made and new nodes are created.
- Recursive Process: In that respect, such a process is recursively repeated for each new node until a stopping criterion is met; for example, by maximum depth or minimum number of samples.
- Leaf Nodes: The nodes at the very end hold the predicted class or value for that subset of data.

4. Experimental Setup

To evaluate the effectiveness of the proposed hybrid model, we conducted experiments on benchmark dataset. We compared the performance of the hybrid model with standalone Decision Tree models and other previous models. Evaluation metrics such as accuracy and F1-score macro are used to assess the model performance.

4.1.DAIC-WOZ Dataset

The Distress Analysis Interview Corpus (DIAC) [35] is an extensive dataset comprising clinical interviews designed to aid in diagnosing psychological stress conditions such as anxiety, depression, and post-traumatic stress disorders. These interviews were part of a significant initiative to develop computer agents capable of identifying linguistic and non-verbal indicators of mental illness [36]. The dataset includes audio and video recordings, along with comprehensive survey responses. A notable component of this corpus is the Wizard of Oz interviews, conducted by an animated virtual interviewer named Ellie, who is controlled by a human interviewer located in another room. The data have been transcribed and annotated for various linguistic and non-linguistic features. This dataset includes 189 interactive sessions, each containing a transcript of the interaction.

"Ellie: hi i'm ellie thanks for coming in today"
"Ellie: i'm here to learn about people and would love to learn about you"
"Ellie: i'll ask a few questions to get us started and please feel free to tell me anything your answers are totally confidential"
"Ellie: how are you doing today?"
"Participant: good"

Data are split into a train set of 107 sessions, a development set of 35 sessions, and 47 sessions as the test set in CSV files with label for each session as well as the participant is depressed or not.

4.2. Evaluation Criteria

The most discussed indicator in the binary classification problem of diagnosing whether a patient has the disorder or not are the following:

Accuracy: The most commonly used metric to judge a model and is actually not a clear indicator of the performance. The worse happens when classes are imbalanced.

$$\text{Accuracy} = \frac{TP+TN}{TP+FP+FN+TN} \tag{5}$$

F1 score: It is the harmonic mean of precision and recall. This takes into consideration the contribution of both, so higher the F1 score, the better. We can see that due to the product in the numerator if one decreases, the final F1 score decreases significantly.

$$\text{F1 Score} = \frac{2*(Recall * Precision)}{(Recall + Precision)} \tag{6}$$

4.3.GWO-DT Model

Interest has grown in hybridizing meta-heuristic algorithms with machine learning techniques for the purpose of boosting performance and solving complex optimization problems. Grey Wolf Optimizer has emerged as another promising tool in this regard; it is a meta-heuristic inspired by grey wolves' hunting behavior. This will improve the performance and interpretability of a model, especially when combined with some classifier like Decision Trees by these steps:

- Feature Subset Generation: GWO will be used for the generation of different subsets of features from the original dataset. Every wolf is considered to be a potential solution or, in other words, a feature subset.
- Fitness Evaluation: In this phase, for every generated subset of features, a decision tree model will be trained and then evaluated with a suitable performance metric such as accuracy or the F1-score. The fitness value of any wolf will then be obtained from the performance of its corresponding decision tree.
- Update Wolf Positions: GWO updates the positions of wolves (feature subsets) based on fitness values to move towards the optimal solution.
- Termination: The algorithm shall terminate when a designated stopping criterion has been met, such as the maximum number of iterations or convergence.

In our hybrid model as shown in Figure 1, we leverage the exploration and exploitation capabilities of the GWO algorithm in the optimization of Decision Tree model hyper-parameters or features. The GWO algorithm searches for optimal hyper-parameters to enhance the performance of a Decision Tree toward better accuracy and generalization on the dataset.

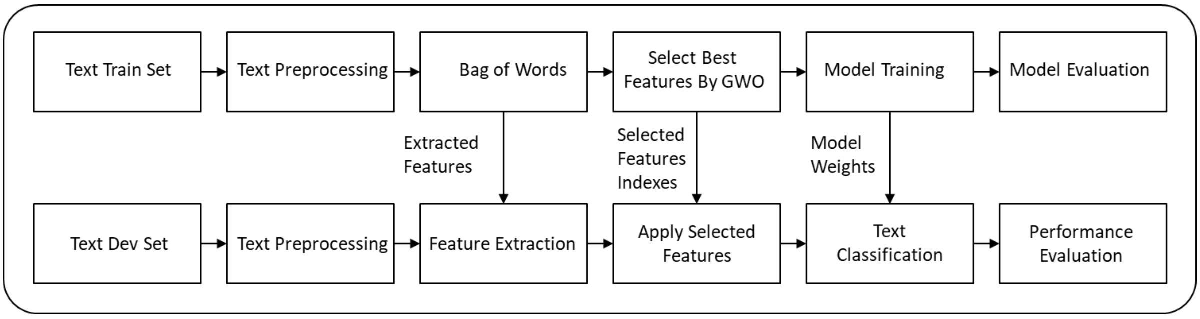


Figure 1 GWO-DT Model Workflow

5. Results Discussion

Our experimental results show that the proposed hybrid model with Grey Wolf Optimizer-enhanced Decision Tree outperformed standalone Decision Tree models and presented very competitive performance compared to other optimization techniques. Integrating GWO into the decision tree improves not only the accuracy of classification but also makes the developed model more interpretable. The hybrid model as shown in Table 1 has achieved 0.83 F-score macro to improve the efficiency compared with the other models.

Table 1 The Evalution Performance Comparison

Model	**Accuracy**	**F1-Score-Macro**
Zhijun D. et al [37]	-	0.67
Danai X. et al [38]	-	0.69
NGUMIMI K. et al [39]	-	0.66
Kirill M. et al [40]	-	0.79
Proposed Decision tree	0.77	0.76
Proposed GWO-DT	0.83	0.83

6. Conclusion

The combination of the Decision Tree method and the GWO is a major development in machine learning optimization. The proposed hybrid model demonstrates a viable strategy that leverages Decision Trees' intrinsic simplicity and interpretability as well as GWO's unique advantages in optimization techniques. By combining several techniques, the integration of hybrid model provided more reliable and effective solution of automated depression diagnosis. Using the optimization capabilities of GWO to improve the hyper-parameters and feature selection of decision trees can lead to enhancement in model interpretability and prediction performance. Our experimental assessments on the DIAC-WOZ dataset confirm the hybrid model's effectiveness in surpassing stand-alone Decision Tree models and demonstrating competitive performance compared to other optimization methods. In addition to improving classification accuracy featuring a high F-score macro of 0.83, combining GWO with Decision Trees enhances the model's interpretability and transparency, making it ideal for real-world applications where decision-making processes must be accurate and explicable. Finally, the hybridization of Decision Trees with the Grey Wolf Optimizer offers a viable path forward for machine learning optimization. Through the combined use of GWO and Decision Trees, the hybrid model becomes a more powerful and adaptable solution that can diagnose depression from clinical transcripts.

References

[1]. M. Shim, H. Hwang, D. Kim, S. Lee, C. Hwan Im,Machine-learning-based diagnosis of schizophrenia using combined sensor-level and source-level EEG features,Schizophrenia Research,Volume 176, Issues 2–3,2016,Pages 314-319,ISSN 0920-9964,https://doi.org/10.1016/j.schres.2016.05.007.

[2]. Institute of health Metrics and Evaluation (IHME). Global Health Data Exchange (GHDx). http://ghdx.healthdata.org/gbd-results-tool?params=gbd-api-2019permalink/27a7644e8ad28e739382d31e77589dd7 (Accessed 2 September 2024).

[3]. B. Runciman. (2020). It's Easier to Talk About Diarrhoea Than Depression. ITNOW. 62. 8-11. 10.1093/itnow/bwaa003 .

[4]. Zhou, Xiuzhuang & Kai, Jin & Shang, Yuanyuan & Guo, Guodong. (2018). Visually Interpretable Representation Learning for Depression Recognition from Facial Images. IEEE Transactions on Affective Computing. PP. 1-1. 10.1109/TAFFC.2018.2828819.

[5]. Zhao, Ziping & Bao, Zhongtian & Zhang, Zixing & Deng, Jun & Cummins, Nicholas & Wang, Haishuai & Tao, Jianhua & Schuller, Björn. (2019). Automatic Assessment of Depression From Speech via a Hierarchical Attention Transfer Network and Attention Autoencoders. IEEE Journal of Selected Topics in Signal Processing. PP. 1-1. 10.1109/JSTSP.2019.2955012.

[6]. Zhou, Xiuzhuang & Huang, Peng & Liu, Haoming & Niu, Sihua. (2019). Learning content-adaptive feature pooling for facial depression recognition in videos. Electronics Letters. 55. 10.1049/el.2019.0443.

[7]. Almouzini, Salma & Khemakhem, Maher & Alageel, Asem. (2019). Detecting Arabic Depressed Users from Twitter Data. Procedia Computer Science. 163. 257-265. 10.1016/j.procs.2019.12.107.

[8]. Razavi, Rouzbeh & Gharipour, Amin & Gharipour, Mojgan. (2020). Depression screening using mobile phone usage metadata: a machine learning approach. Journal of the American Medical Informatics Association : JAMIA. 27. 10.1093/jamia/ocz221.

[9]. Ray, Anupama & Kumar, Siddharth & Reddy, Rutvik & Mukherjee, Prerana & Garg, Ritu. (2019). Multi-level Attention Network using Text, Audio and Video for Depression Prediction. 81-88. 10.1145/3347320.3357697.

[10]. Jing, Zhu & Zihan, Wang & Gong, Tao & Shuai, Zeng & Li, Xiaowei & Hu, Bin & Li, Jianxiu & Shuting, Sun & Zhang, Lan. (2020). An Improved Classification Model for Depression Detection Using EEG and EyeTracking Data. IEEE Transactions on NanoBioscience. PP. 1-1. 10.1109/TNB.2020.2990690.

[11]. Alghamdi, Norah & Mahmoud, Hanan & Abraham, Ajith & Alanazi, Samar & Garcia-Hernandez, Laura. (2020). Predicting Depression Symptoms In An Arabic Psychological Forum. IEEE Access. PP. 1-1. 10.1109/ACCESS.2020.2981834.

[12]. Xezonaki, Danai & Paraskevopoulos, Georgios & Potamianos, Alexandros & Narayanan, Shrikanth. (2020). Affective Conditioning on Hierarchical Attention Networks Applied to Depression Detection from Transcribed Clinical Interviews. 4556-4560. 10.21437/Interspeech.2020-2819.

[13]. Xu, Zhenxing & Wang, Fei & Adekkanattu, Prakash & Bose, Budhaditya & Vekaria, Veer & Brandt, Pascal & Jiang, Guoqian & Kiefer, Richard & Luo, Yuan & Pacheco, Jennifer & Rasmussen, Luke & Xu, Jie & Alexopoulos, George & Pathak, Jyotishman. (2020). Subphenotyping depression using machine learning and electronic health records. Learning Health Systems.

[14]. Othmani, Alice & Kadoch, Daoud & Bentounes, Kamil & Rejaibi, Emna & Alfred, Romain & Hadid, Abdenour. (2021). Towards Robust Deep Neural Networks for Affect and Depression Recognition from Speech. 10.1007/978-3-030-68790-8_1.

[15]. Sharma, Amita & Verbeke, Willem. (2020). Improving Diagnosis of Depression With XGBOOST Machine Learning Model and a Large Biomarkers Dutch Dataset (n = 11,081). Frontiers in Big Data. 3. 10.3389/fdata.2020.00015.

[16]. Nakano, Takashi & Takamura, Masahiro & Ichikawa, Naho & Okada, Go & Okamoto, Yasumasa & Yamada, Makiko & Suhara, Tetsuya & Yamawaki, Shigeto & Yoshimoto, Junichiro. (2020). Enhancing Multi-Center Generalization of Machine Learning-Based Depression Diagnosis From Resting-State fMRI. Frontiers in Psychiatry. 11. 400. 10.3389/fpsyt.2020.00400.

[17]. Zhang, Yiye & Wang, Shuojia & Hermann, Alison & Joly, Rochelle & Pathak, Jyotishman. (2021). Development and validation of a machine learning algorithm for predicting the risk of postpartum depression among pregnant women. Journal of Affective Disorders. 279. 1-8. 10.1016/j.jad.2020.09.113.

[18]. Na, Kyoung-Sae & Cho, Seo-Eun & Geem, Zong Woo & Kim, Yong-Ku. (2020). Predicting future onset of depression among community dwelling adults in the Republic of Korea using a machine learning algorithm. Neuroscience Letters. 721. 134804. 10.1016/j.neulet.2020.134804.

[19]. Richter, Thalia & Fishbain, Barak & Markus, Andrey & Richter-Levin, Gal & Okon-Singer, Hadas. (2020). Using machine learning-based analysis for behavioral differentiation between anxiety and depression. Scientific Reports. 10. 10.1038/s41598-020-72289-9.

[20]. Latif, S., Usman, M., Irfan, R., \& Qadir, J. (2020). Promoting Independence of Depression and Speaker Features for Speaker Disentanglement in Speech-Based Depression Detection. IEEE Journal of Biomedical and Health Informatics, 24(9), 2664-2673.

[21]. Amir, O., Cohen-Kalka, S., \& Zigdon, A. (2020). DAIC-WOZ: On the Validity of Using the Therapist's prompts in Automatic Depression Detection from Clinical Interviews. In Proceedings of the Sixth Workshop on Computational Linguistics and Clinical Psychology (pp. 91-101).

[22]. Latif, S., Qadir, J., Farooq, H., \& Imran, M. A. (2021). Speechformer-Ctc: Sequential Modeling of Depression Detection with Speech Temporal Classification. IEEE Journal of Biomedical and Health Informatics, 25(7), 2525-2535.

[23]. Yin, F., Du, J., Xu, X., \& Zhao, L. (2023). Depression Detection in Speech Using Transformer and Parallel Convolutional Neural Networks. Electronics, 12(2), 328.

[24]. Lorenzoni, Giuliano and Tavares, Cristina and Nascimento, Nathalia and Alencar, Paulo and Cowan, Donald, 2024 , Assessing ML Classification Algorithms and NLP Techniques for Depression Detection: An Experimental Case Study

[25]. Rejaibi, E., Komaty, A., Meriaudeau, F., Agrebi, S., \& Othmani, A. (2023). MFCC-based recurrent neural network for automatic clinical depression recognition and assessment from speech. Psychological Medicine, 1–11. DOI: 10.1017/S0033291720002718

[26]. Rai, B.K., Jain, I., Tiwari, B. et al. Multimodal mental state analysis. Health Serv Outcomes Res Method (2024). https://doi.org/10.1007/s10742-024-00329-2

[27]. Zhang, W., Mao, K. \& Chen, J. A Multimodal Approach for Detection and Assessment of Depression Using Text, Audio and Video. Phenomics (2024).

[28]. Afzal Aghaei, A., \& Khodaei, N. (2023). Automated Depression Recognition Using Multimodal Machine Learning: A Study on the DAIC-WOZ Dataset. Computational Mathematics and Computer Modeling with Applications (CMCMA), 2(1), 45-53. doi: 10.48308/CMCMA.2.1.45

[29]. Hamed, G., Marey, M., Amin, S., & Tolba, M. (2021). Comparative Study and Analysis of Recent Computer Aided Diagnosis Systems for Masses Detection in Mammograms. *International Journal of Intelligent Computing and Information Sciences, 21*(1), 33-48. doi: 10.21608/ijicis.2021.56425.1050

[30]. Ibrahim, A., Alfonse, M., & Aref, M. (2023). A SYSTEMATIC REVIEW ON TEXT SUMMARIZATION OF MEDICAL RESEARCH ARTICLES. *International Journal of Intelligent Computing and Information Sciences*, *23*(2), 50-61. doi: 10.21608/ijicis.2023.190004.1252

[31]. Daniel Jurafsky and James H. Martin. 2024. Speech and Language Processing: An Introduction to Natural Language Processing, Computational Linguistics, and Speech Recognition with Language Models, 3rd edition. Online manuscript released August 20, 2024. https://web.stanford.edu/~jurafsky/slp3.

[32]. Amin, A. (2020). A Face Recognition System Based on Deep Learning (FRDLS) to Support the Entry and Supervision Procedures on Electronic Exams. International Journal of Intelligent Computing and Information Sciences, 20(1), 59-75. doi: 10.21608/ijicis.2020.23149.1015

[33]. Alkady, W., ElBahnasy, K., & Gad, W. (2022). Comparative Study on Feature Selection Methods for Protein. International Journal of Intelligent Computing and Information Sciences, 22(3), 109-123. doi: 10.21608/ijicis.2022.144051.1190

[34]. Ghaleb, M., Moushier, H., Shedeed, H., & Tolba, M. (2022). Weather Classification using Fusion Of Convolutional Neural Networks and Traditional Classification Methods. International Journal of Intelligent Computing and Information Sciences, 22(2), 84-96. doi: 10.21608/ijicis.2022.117060.1156.

[35]. J. Gratch ,R. Artstein, GM. Lucas, G. Stratou, S. Scherer et al. The Distress Analysis Interview Corpus of Human and Computer Interviews. In Proceedings of LREC 2014 May (pp. 3123-3128).

[36]. D. DeVault, R. Artstein, G. Benn, T. Dey, E. Fast, A. Gainer et al. (2014). "SimSensei kiosk: A virtual human interviewer for healthcare decision support." In Proceedings of the 13th International Conference on Autonomous Agents and Multiagent Systems (AAMAS'14), Paris.

[37]. Dai, Zhijun Zhou, Heng Ba, Qingfang Zhou, Yang Wang, Lifeng Li, Guochen. (2021). Improving Depression Prediction Using a Novel Feature Selection Algorithm Coupled with Context-Aware Analysis. Journal of Affective Disorders. 295. 10.1016/j.jad.2021.09.001.

[38]. Xezonaki, Danai Paraskevopoulos, Georgios Potamianos, Alexandros Narayanan, Shrikanth. (2020). Affective Conditioning on Hierarchical Attention Networks Applied to Depression Detection from Transcribed Clinical Interviews. 4556-4560. 10.21437/Interspeech.2020-2819.

[39]. N. K. Iyortsuun, S. -H. Kim, H. -J. Yang, S. -W. Kim and M. Jhon, "Additive Cross-Modal Attention Network (ACMA) for Depression Detection Based on Audio and Textual Features," in *IEEE Access*, vol. 12, pp. 20479-20489, 2024, doi: 10.1109/ACCESS.2024.3362233.

[40]. Milintsevich, Kirill & Sirts, Kairit & Dias, Gaël. (2023). Towards automatic text-based estimation of depression through symptom prediction. Brain informatics. 10. 4. 10.1186/s40708-023-00185-9.