A Comprehensive Survey on Depression Detection Techniques

Yash Kandalkar, Sanket Kulkarni, Abdurrahman Yusufi, Raghav Dodla, Manjiri Ranjanikar

Pimpri Chinchwad College of Engineering Nigdi, Pune, India

Corresponding author: Yash Kandalkar, Email: yashkandalkar555@gmail.com

Depression and mental diseases are significant problems in today's culture. It may result in suicidal thoughts.In this study, we analyze several earlier studies that identified depression using deep learning (DL), machine learning (ML), and artificial intelligence (AI).This study explores how emotions may be precisely identified, and how depressive symptoms can be communicated via social media posts, images, audio, and facial expressions. Linear Support Vector, Logistic Regression, Long Term Short Memory (LSTM), Naive-Bayes, Support Vector Machines (SVM), etc. are some of the ML approaches used to recognize and classify emotions from the data. Artificial neural networks (ANN) are used to classify and extract features from photographs to identify emotions from facial expressions. The purpose of this work is to present an overview of several AI, DL, and ML approaches that aid in the identification and study of emotion, and thereby depression, as well as the research concerns, they raise.

Keywords: Deep Learning, Image Processing, Machine Learning, Depression Detection, Audio Processing, Text Processing

2023. In Deepak Bhatia & Himanshu Arora (eds.), *SCRS Proceedings of International Conference of Undergraduate Students 2023*, 25–34. Computing & Intelligent Systems, SCRS, India. https://doi.org/10.56155/978-81-955020-4-2-3

1. Introduction:

All ages can experience depression. It can be extremely dangerous and result in conditions like anxiety attacks, heart attack mortality, and issues with blood pressure and diabetes. Many surveys have shown that there is an increase in depression among adults. In addition, one of the main factors contributing to suicide is depression. Therefore, it is crucial to identify it and its causes so that proper treatment may be found. Additionally, social network mental disorder detection can be used to assist in de-stigmatize depression and mental health because there is a need to do so. Tests based on different artificial intelligence, computer vision, and machine learning algorithms can be done in several scenarios to detect emotional imbalance. As technology advances, several AI-based methods are being developed to make machines emotionally aware so they can recognize human emotions. Text-based emotion recognition, such as sentiment analysis of tweets[6] and postings on different social media platforms, can assist in identifying the user's emotions and mood as well as assist in predicting the user's suicidal thoughts and preventing suicide by alerting the user or their loved ones. This can be done using a variety of machine learning methods, such as Support Vector Machines (SVM), Naive-Bayes, Random Forest, and others. The results may be assessed using a confusion matrix. A powerful algorithm will properly predict sentiment, which can be either positive or negative, and will have a high precision score. The study[22] gives a brief about the various deep learning methodologies.

Not only text, but using other biomarkers like audio, video, image, Heart Rate Variability (HRV), and EEG can be used for detecting depression among people. Audio recordings of the subjects can be used for the detection of depression as many features can be extracted from the audio speech. Coming to biological factors are found to be very accurate in the prediction of depression detection. Apart from these HRV data can also be used to predict depression. Much research has been made in this area and has shown promising outcomes overall.

In many pieces of research [17], it has been observed that the problem of dataset generation or dataset creation takes a huge effort. Since the person's medical history is not public property, we cannot have access to this sensitive data. The research has mainly focused on existing available datasets like AVEC2013, AVEC2014, and its successors. Many factors are considered while detecting depression including the facial features like temporal areas, nose, chin, areas below the eyes, etc. In this paper, we will go through the prominent ways of depression detection, their techniques, and a brief overview of each of them.

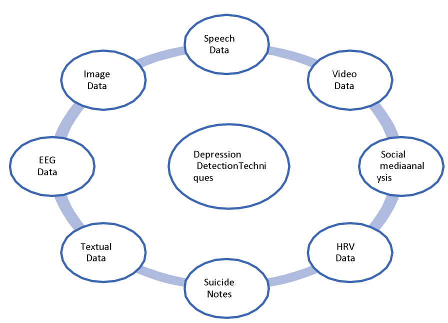

Fig 1. Types of data for depression detection

2. Literature review

The research in this domain is very abundant. This section consists of various methods and techniques which are used for depression detection. Each of the individual techniques is discussed in detail to get a good idea of the technologies and methodologies. Also, consider the following figures which describe the year-wise and publication-wise distribution of the research paper published in the depression detection domain.

Figure 2 shows the number of published documents broken down by document type. The articles and conference papers have each had the most publications on this subject. Numerous high-caliber articles on current research trends in this field are anticipated to be published during the coming years.

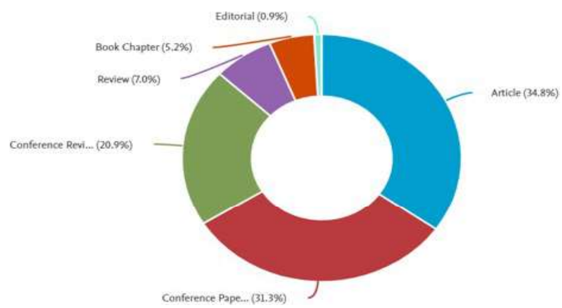

Fig. 2 Documents by Type
Source:http://www.scopus.com(accessedon20thSeptember2022)

Figure 3 shows the data from 2018 to 2022 from the documents relating to depression detection methods utilizing deep learning algorithms for the last five years.

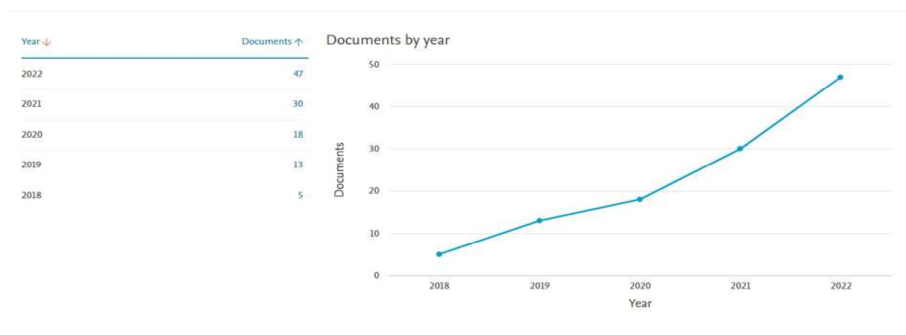

Fig.3 DocumentsbyYear
Source:http://www.scopus.com(accessedon20thSeptember2022)

2.1 Depression detection using audio data

Speech samples are used in [18] to detect people with clinical depression. Two models, an end-to-end convolutional neural network model, and a spectrogram-based model were trained to accomplish this. Parameter tweaking has been done by choosing a few convolutional neural network sub-parameters. The models are tested using speech samples from the DAIC-Woz dataset (AVEC), which was originally from the Audio-Visual Emotion Challenge 2016. NCH wave pad sound editor is used for segmenting data into 7 seconds of a sample. In the spectrogram-based CNN model, 3 convolutional layers were used. Also, max pool layers were used after each convolutional layer for dimensionality reduction and to reduce redundancy. In the end-to-end CNN model also, three convolutional layers were used but with different kernel sizes. The output from both was sent to the flattening layer, which transformed it into a 1D array, and it was then given to fully linked layers, which used the loss function to calculate the error. Additionally, SoftMax performed the classification based on the estimated class's computed probability. The Spyder tool was used to carry out the approach. It was discovered that the end-to-end model outperformed the baseline model and the spectrogram-based convolution neural network.

In [25], a method was proposed to extract full vocal tract coordination features using CNN. A correlation matrix is given to the four convolutional layers as input. The output of which is then combined and passed through two more convolutional layers. Further, the output is then flattened and then given to two fully connected layers each having 32 ReLUs. The experiment was done using two datasets, which are DAIC-WOZ and SH2-FS. There were 4 sets of acoustic feature formants, spectral centroid frequencies (SCF), Mel-frequency cepstral coefficients (MFCC), and delta MFCC. An MFCC-based voice activity detection was applied on SH2-FS while for DAIC-WOZ, transcription-based voice activity detection was applied. Adam optimizer algorithm was used for optimization. The method showed performance improvement compared to the previous similar method which used vocal tract coordination features and not only that it also overcame the limitations of previous work. The results had high accuracy on the DAIC-WOZ dataset than in the SH2-FS dataset.

In [7], the method proposed contains three steps. In the first step called deep features extraction, the spectrogram of the audio signal is given as input for extracting two deep features which are Speech Emotion Recognition (SER) and Speaker Recognition (SR). SR model was used for the extraction of vocal differences among different candidates. While SER was used for emotional differences of the same candidate. For SER, happy, nervous, sad, and peaceful, these 4 emotional classes were used. Both features were extracted using a pre-trained ResNet-based model. FVCM algorithm was proposed for extracting the coordination features. These coordination features were obtained by calculating the covariance and correlation coefficients of time-delayed multi-channel variations. Also, a hierarchical model with multiple classifiers was proposed for depression detection. First deep SER and SR features were given as input to the multilayer perceptron. The regression interval scaling algorithm was used for calculating the regression interval which was given as a constraint to the MLP regressor. The datasets used were AVEC2013 and AVEC2014. The results of this method were comparable to other video or multi-model-based methods for the detection of depression. Combining the two features improved the performance of the model.

In [10], a method was implemented by recording 33 interviews of both depressed and non-depressed groups. Tascam was used for recording the audio during the interview. Audio software Audacity was used for audio editing. Unwanted voice signals were removed. Open-source GNU Octave software was used for signal processing and feature extraction. Using the audio, 33 features such as skewness, kurtosis, etc. were extracted, and then different well-known classifiers like Multi-Layer Perceptron, Logistic Regression, SVM, Bayes Net, Random Forest, and Naïve Bayes were used for the classification of the candidates. Machine learning software Weka was used for classification. The outcome was that the Random Forest with 100 trees achieved the highest performance in classification.

In [14], a hierarchical context-aware graph (HCAG) attention model is proposed for depression identification. This method uses the DAIC-WOZ dataset which contains the audio data in interview format. It uses a sequential context encoder and subject-level context encoder for encoding question and answers pairs in the dataset. Audio features were extracted by using OpenSmile. Adam optimizer is used to train the model. The model performed well, earning an F1 score of 0.92 and scores for RMSE and MAE of 3.80 and 2.93, respectively. The F1 Score, Precision, Recall, MAE, and RMSE were the performance metrics. On every parameter, the model fared better than the baseline models.

In [15], 59 recorded interviews were taken. Of these, 29 were depressed and the remaining 30 were non-depressed. It uses both the audio and the text transcript of the interview. Two separate unimodal were proposed for text and speech. Theproposed method first extracts the features from the dataset of 59 recorded interviews. The features extracted are then fed to bidirectional Long Short Term Memory networks, the output of which is then passed to a softmax layer for classification.With two separate unimodal speech and text, a multimodal method was also proposed for the same. The multimodal used the methodologies like Sum Rule, Joint Representation, and Gated Multimodal Unit (GMU). The results found that the multi-modal method was found to give better performance than the unimodal approaches.

In [2], an audio-based framework is proposed using a deep learning technique for extracting highly relevant features for depression detection. It uses the technique of Convolution Neural Network based Autoencoder (CNN AE). The method uses the DIAC-WOZ dataset. PyAudioAnalysis library was used for removing silence and removing the interviewer's voice. Python library Librosa was used for downsampling. Further, data augmentation techniques were applied. The cluster centroids under-sampling technique was applied to avoid the overfitting problem. The data then was provided to CNN AE for feature extraction which containsconvolutional, max-pooling, and dropout layers. The output is further provided to flatten and dense layer which generates a 1D vector of 700 features. In the decoding phase, the data goes through reshaped layer, a dense layer, and four convolutional layers to reconstruct the input. Further, classification techniques like SVM, Random Forest, MLP, and GB were used. The results showed a 7% improvement when compared to other audio-based models.

2.2 Depression detection using audio and video data

Facial expression analysis is one of the ways to detect depression in patients. [9] extracts feature using an LBP (Local Binary Descriptor), and it classifies depression levels using an SVM classifier. Here the dataset used was the Cohn- Kanade dataset with a few images from the internet also added to make the dataset more inclusive. The Viola-Jones method was used to extract the face from the photos, and the face was then cropped from the relevant area of the image. The SVM classifier is applied to these images and the final output is taken. By comparing the number of happy features with features of disgust and unhappiness the final output is predicted.

In the review [8], depression is analyzed using computer vision. The Amsterdam Dynamic Look Set (ADFES) was demonstrated to the group as a motivator. Their responses were captured, and a CNN model was created using them. Convolutional Neural Network (CNN) was used to distinguish facial features, while the Dlib AI framework was used to recognize members' facial feelings. (Conv 5×5, Maxpool 2×1, Dropout, Conv2 3×3, Maxpool2 2×1, Dropout, Conv3 3×3, maxpool3 2×1, Thick, SoftMax) is the design of the CNN model that was used. The research of vocal reactions as well as other biometrics like eye staring, electrocardiography, heart rate, skin conductance, or electroencephalography is not considered in this test.

The research [16] takes a unique approach to analyze eye movements and cognitive bias. The study focuses on tracking of eye movements of the patients. It uses images on the screen which are positive, negative, or neutral. The three images appear on the screen. The subject needs to look at the image whose border turns green. The databases used in this research are OASIS and ThuPis. SVM is used to classify depressed and non-depressed patients. It has been observed that a person with depressionspends a long time looking at the negative image and has very slow eye movement. The researchers in this study have gotten 77.6% accuracy.

The study in [11] utilizes two floods of CNNs. The model is prepared on the AVEC 2014 dataset. The utilization of a consecutive combination system with two separate CNNs on crude pictures and the optical stream can enormously improve depression acknowledgment. This showed that consecutive combination gives a reasonable probabilistic viewpoint. It comes up short on examination of the private-share model for multimodal sadness.

In the research [24], the creators have proposed an original fleeting pooling technique to catch and encode the Spatiotemporal dynamic of video cuts into a picture map. This execution of theirs has given a preferred option over the 2D and 3D CNNs and diminished the requirement for enormous preparation datasets. The model comprises a two-stream model which takes in spatial and worldly elements and applies the idea of scores in each stream. The stream has a huge likelihood/most extreme score picked and the rest is disregarded. This cycle is called score-level combination. For the outcomes part, this model delivered an RMSE of 7.9 on the AVEC-2013 dataset, and the MAE of the AVEC-2013 dataset is 5.96. Additionally, the model generated an RMSE of 7.94 and an MAE of 6.20 on the AVEC-2014 dataset.

2.3 Depression Detection using images and video data

Three steps are shown in the pipeline for the proposed methodology in [19] the extraction of a facial region, feature extraction, and the machine learning stage. A face region can be extracted by combining the outcomes of OpenFace 2.0 for facial alignment, Convolutional Experts Constrained Local Model (CE-CLM) for locating facial landmarks, and Point Distribution Model (PDM) to regularize the form. To extract features, Local Binary Patterns (LBP) and Histogram of Oriented Gradients(HOG), two appearance-based descriptors, are used.The Visual Geometry Group (VGG), University of Oxford, developed a pre-trained Convolutional Neural Network (CNN) architecture that was previously put to the test by the successful 2014 ImageNet Large Scale Visual Recognition Challenge (ILSVRC). Principal Component Analysis(PCA) is used to reduce dimensionality. Support Vector Machines (SVM) are used for binary classification, while Support Vector Regression (SVR) is used to predict the result score.

The key symptom of depression is mood swings or low mood. Using facial expressions emotions of people can be identified. According to a study in [21], facial behavioral recognition can identify patients with mental illnesses more accurately than the questionaries method since the former fails to recognize the patient's many facial emotions. The proposed architecture contains three main stages Face data alignment and keyframe face detection, generating per-frame basis features using pre-trained VGG-face and feeding them to Long Short-Term Memory (LSTM) model, and at last, combining sentence-level predictions by committee fusion. This study on a sample of a multimodal behavioral dataset of Chinese university students gives high accuracy for multiple fused systems compared to single classifiers.

The paper [26] reviews different machine-learning techniques for depression detection on audio and image data. The National Institutes of Health's database is used to diagnose depression in patients. The three main types of depression indicators used to identify depression are visual, speech, and social indicators. To classify human facial developments according to how they appear on the face, we require the Facial Action Coding Framework (FACS) framework. Voice Activity Detection (VAD) is crucial

during the pre-processing step when building a model of speech indicators. The social indicator shows that a depressed person exhibits weak body language and undesirable social behavior. It might help with depression detection. General machine learning architecture involves pre-processing, feature extraction, classifier training, and algorithm like SVM. The non-verbal and verbal behaviors of a depressed and non-depressed individual differ significantly, according to this paper's conclusion.

The largest issue in developing an efficient and reliable model for depression detection is giving multimodal data (text, picture, video, audio, etc.) simultaneously to a model. In the study [27], a solution is put out using a graph attention model that incorporates multimodal knowledge. The DAIC-WOZ dataset is utilized in this instance to compare the results with other state-of-the-art approaches. Multimodal Attention mechanism with Temporal Convolutional Network (TCN) provides the solution to the above problem. In this method, the multimodal data is fed into a TCN which gives multimodal sentence vectors (Fnetwork). From Fnetwork and attention embedding vectors, we get the knowledge-attention representation vector (f) which is used to predict information on labels, i.e., ☐ˆp. The proposed model follows the following steps: Inter-Modality-Level Attention → Intra-Modality-Level Attention → Cross-Modality Shared Attention. The proposed model gives 95.4% and 94.6% F1 measures respectively for dilated and causal temporal convolutional networks which are greater than all other methods.

The study [12] makes use of the Maximization and Differentiation Network (MDN). The dataset used in the study are AVEC2013 and AVEC 2014. It focuses on capturing the spatiotemporal features from the given frames. MDN engineering is a convolutional network that uses expansion and distinction structures to examine spatiotemporal variations.

2.4 HRVdata for depression detection

The method for stress, anxiety, depression, and other mental health measures can be detected using HRV data from wearables [20]. The three forms of data inputs used for the prediction are HRV measurements, time domain data, and frequency domain data. The 652 participants provided a sample of data, they were fromthe University of Surrey, England. Data capturing was done using the Biobeam band (BioBeats Group Ltd., London). The data sample was further split into day and night data. The Long Short-Term Memory (LSTM) network along with the sigmoid activation function was employed for the prediction of depression levels. For two and five-minute HRV data, up to 73% and 83% categorization accuracy were attained respectively.

2.5 EEG data techniques for depression detection

The study in [23] shows that using EEG data we can predict depression. It mainly focuses on analyzing the EEG data and predicting the mental state of the patient. A combination of LSTM and CNN is used to predict depression in patients. It uses CNN to extract the feature and these extracted features are fed to LSTM. The dataset used in this research is obtained from "PATIENT REPOSITORY FOR EEG DATA + COMPUTATIONAL TOOLS". The major flaw of this techniqueis the cost of the EEG machine and its maintenance. Though it is very accurate from other depression diagnosis techniques it is notcost-effective. The method showed a 20% classification error as well as 98.84% accuracy from the left hemisphere and 99.07% accuracy from the right hemisphere, respectively.

The study in [1] describes the use of deep learning to identify depression using EEG data. Studies using RNN, RNN plus SVM, LR, and other techniques have produced positive outcomes.

The research in [13] aims at using EEG data and facial features together. Here the biological data parameters are used and combined with facial features to get the depression score. The method uses a BiLSTM for the prediction in which two streams are combined to get the result. The proposed methodology achieved an accuracy of more than 95% which speaks a lot about the accuracy of the model.

2.6 Text-based and social media-based depression detectiontechniques

In the study [4] depression is detected using the suicide notes of the subjects. There is a very high possibility that a person who is depressed can show suicidal tendencies. The research aims at using the corpus of suicide notes in English, CEASE, and somewherearound two thousand additional lines have been added to make the dataset more inclusive. It aims at building multitask frameworkwhich can perform emotion recognition, sentiment classification,and depression detection. The researchers in this study obtained the highest cross-validation MR of 56.47 %. The study also suggests that all three tasks mentioned above if performed simultaneously will result in better learning rather than if they are done alone.

The research [5] prepares a hybrid model, Multi-Aspect Depression Detection Hierarchical Attention Network MDHANto boost the classification of depressed users using multi-aspect features and word embedding features. a novel explainable depression detection framework using deep learning of the textual, behavioral, temporal, and semantic aspect features of social media. In comparison to SVM, NB, and BiGRU, the MDL model performs better and achieves higher accuracy. Since this model is relatively new and was created specifically to find depressed people, it has successfully captured the subtleties of the dataset and correctly learned its characteristics, producing better results. Since the timeline of the user post is an essential factor in depression detection, the proposed method takes into consideration this aspect as well.

The study [3] focuses on using ensemble learning. Here the authors are using NLP techniques to detect depression from the text-based data. The datasets used in this study areCLPsych 2015 Shared Task, Reddit, and eRisk dataset. The highest accuracy of 75% was observed on the Reddit dataset, which contained social media data of depressed patients. CLPsych and eRisk showed an accuracy of 65% and 75% respectively.

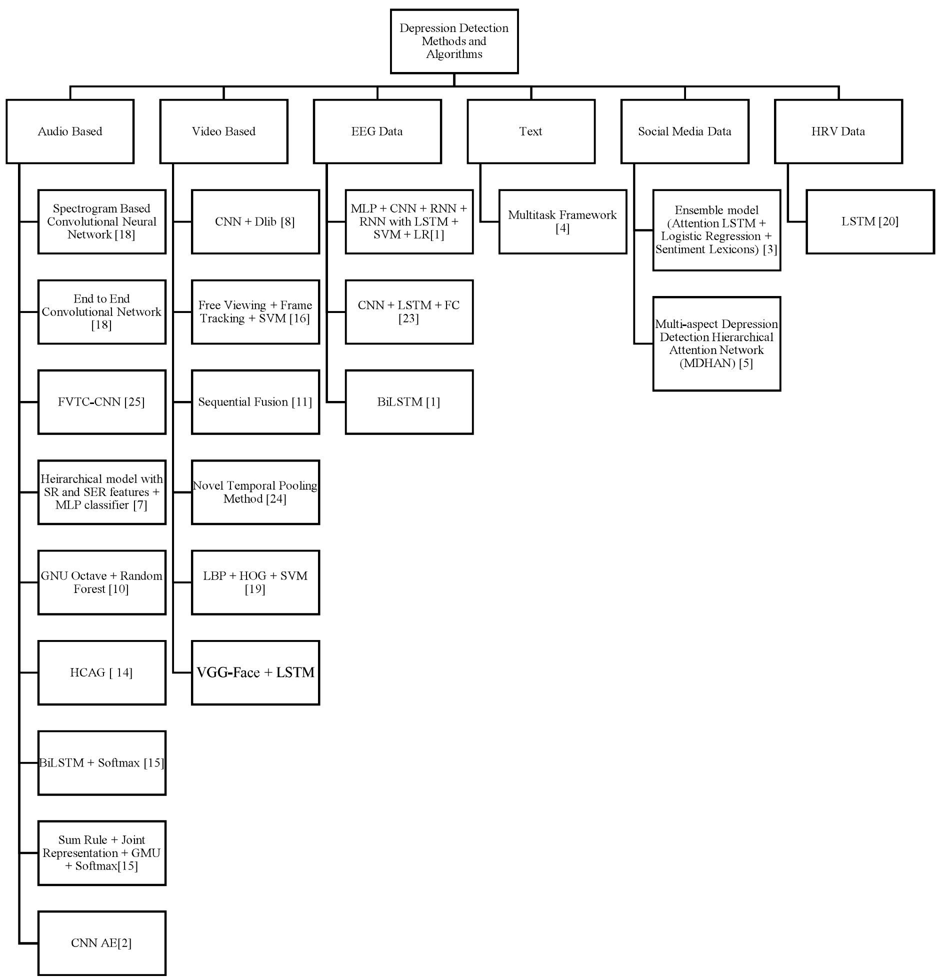

Fig 4. Depression Detection Categories and Methods

3. Conclusion:

Depression, which is increasingly turning into a huge illness, often has an impact on people from all social levels, ethnicities, and countries. Due to the intrinsic nature of solitude, it might be difficult to identify those who need care due to a mental condition but are unable to communicate it. This problem commonly goes unreported even by those who are sad. Since textual sentiment analysis is a non-invasive method that can be continuously observed and managed, it helps identify diseases. This is a huge help in the fight against depression since it enables us to tell the difference between happy and sad periods without seeing a psychologist, giving us the ability to respond promptly when necessary. Using chatbots, emotional AI, mixed inputs, and facial expression analysis (video and image processing), it is possible to recognize, understand, and prevent depression. Several AL and ML approaches, such as Linear Support Vector, Naive-Bayes, Logistic Regression, PCA, LSTM-RNN, and KNN Classification, are used to discern emotions and subsequently identify depression. The effectiveness and performance of several algorithms, including SVM and Multinomial Naive-Bayes, are examined to see which is more effective at identifying emotions and, consequently, depression, in tweets. AI-based solutions driven by interactive technology are also highlighted. For instance, when a chatbot notices that a user is depressed, it will reply with a joke or music to lift their spirits. The identification, analysis, and treatment of depression as well as its prevention may be made possible by such emotional AI and ML-based solutions. In the future, these methods can be included in a whole system to clinically categorize sad people based on their emotional profiles. In conclusion, using different AI and ML algorithms, depression, mood, and emotion may all be detected usinggestures, text, videos, speech, photographs, etc.

References:

[1] Avik Sarkar, Ankita Singh, Rakhi Chakraborty, "A deep learning-based comparative study to track mental depression from EEG data", Neuroscience Informatics,Volume 2, Issue 4, 2022, 100039, ISSN 2772-5286

[2] Sara Sardari, BaharehNakisa, Mohammed NaimRastgoo, Peter Eklund, "Audio based depression detection using Convolutional Autoencoder", Expert Systems with Applications, Volume 189, 2022, 116076, ISSN 0957-4174

[3] L. Ansari, S. Ji, Q. Chen and E. Cambria, "Ensemble Hybrid Learning Methods for Automated Depression Detection", in IEEE Transactions on Computational Social Systems, vol. 10, no. 1, pp. 211-219, Feb. 2023

[4] Ghosh, S., Ekbal, A. & Bhattacharyya, P."A Multitask Framework to Detect Depression, Sentiment and Multi-label Emotion from Suicide Notes", CognComput 14, 110–129 (2022).

[5] Zogan, H., Razzak, I., Wang, X. et al. "Explainable depression detection with multi-aspect features using a hybrid deep learning model on social media", World Wide Web 25, 281–304 (2022).

[6] Babu, N.V., Kanaga, E.G.M. "Sentiment Analysis in Social Media Data for Depression Detection Using Artificial Intelligence: A Review." SN COMPUT. SCI. 3, 74 (2022).

[7] Dong, Yizhuo, and Xinyu Yang. "A hierarchical depression detection model based on vocal and emotional cues", Neurocomputing 441 (2021): 279-290

[8] Singh, Jaiteg, and GauravGoyal. "Decoding depressive disorder using computer vision", Multimedia Tools and Applications 80, no. 6 (2021): 8189-8212

[9] Tadalagi, Manjunath, and Amit M. Joshi. "AutoDep: automatic depression detection using facial expressions based on linear binary pattern descriptor." Medical & biological engineering & computing 59, no. 6 (2021): 1339-1354,

[10] Espinola, Caroline Wanderley, Juliana Carneiro Gomes, JessianeMônica Silva Pereira, and Wellington Pinheiro dos Santos. "Detection of major depressive disorder using vocal acoustic analysis and machine learning—an exploratory study." Research on Biomedical Engineering 37, no. 1 (2021): 53-64

[11] Chen, Qian, ItiChaturvedi, ShaoxiongJi, and Erik Cambria. "Sequential fusion of facial appearance and dynamics for depression recognition." Pattern Recognition Letters 150 (2021): 115-121

[12] de Melo, WheidimaCarneiro, Eric Granger, and Miguel Bordallo Lopez. “MDN: A deep maximization-differentiation network for spatio- temporal depression detection.” EEE Transactions on Affective Computing (2021),
[13] Hamid, DannielShazmeer Bin Abdul, S. B. Goyal, and Pradeep Bedi. “Integration of Deep Learning for Improved Diagnosis of Depression using EEG and Facial Features.” Materials Today: Proceedings (2021),
[14] Niu, Meng, Kai Chen, Qingcai Chen, and Lufeng Yang. “Hcag: A hierarchical context-aware graph attention model for depression detection.” In ICASSP 2021-2021 IEEE international conference on acoustics, speech and signal processing (ICASSP), pp. 4235-4239. IEEE, 2021
[15] Aloshban, Nujud, Anna Esposito, and Alessandro Vinciarelli. “What you say or how you say it? depression detection through joint modeling of linguistic and acoustic aspects of speech.” Cognitive Computation (2021): 1-14,
[16] Shen, Ruizhe, Qi Zhan, Yu Wang, and Huimin Ma. “Depression detection by analysing 16 movements on emotional images.” In ICASSP 2021-2021 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), pp. 7973-7977. IEEE, 2021,
[17] A. Ashraf, T. S. Gunawan, F. D. A. Rahman, M. Kartiwi, N. Ismail and Ulfiah, “A Summarization of the Visual Depression Databases for Depression Detection,” 2020 6th International Conference on Wireless and Telematics (ICWT), 2020, pp. 1-6
[18] N. Srimadhur and S. Lalitha. “An end-to-end model for detection and assessment of depression levels using speech”, Procedia Computer Science, 171:12–21, 2020,
[19] Pampouchidou, A., Pediaditis, M., Kazantzaki, E. et al. “Automated facial video-based recognition of depression and anxiety symptom severity: cross-corpus validation” Machine Vision and Applications 31, 30 (2020),
[20] Louise V. Coutts, David Plans, Alan W. Brown, John Collomosse, “Deep learning with wearable based heart rate variability for prediction of mental and general health”, Journal of Biomedical Informatics, Volume 112, 2020, 103610, ISSN 1532-0464,
[21] Liu, JQ. et al. (2020). “Dynamic Facial Features in Positive-Emotional Speech for Identification of Depressive Tendencies” In: Chen, YW., Tanaka, S., Howlett, R., Jain, L. (eds) Innovation in Medicine and Healthcare. Smart Innovation, Systems and Technologies, vol 192. Springer,
[22] H. Sanyal, S. Shukla and R. Agrawal, “Study of Depression Detection using Deep Learning,” 2021 IEEE International Conference on Consumer Electronics (ICCE), 2021, pp. 1-5,
[23] Thoduparambil, P.P., Dominic, A. & Varghese, S.M, “EEG-based deep learning model for the automatic detection of clinical depression” Phys Eng Sci Med 43, 1349–1360 (2020).
[24] W. Carneiro de Melo, E. Granger and M. B. Lopez, “Encoding Temporal Information For Automatic Depression Recognition From Facial Analysis,” ICASSP 2020 - 2020 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), 2020, pp. 1080-1084,
[25] Z. Huang, J. Epps and D. Joachim, “Exploiting Vocal Tract Coordination Using Dilated CNNS For Depression Detection In Naturalistic Environments”, ICASSP 2020 - 2020 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), 2020, pp. 6549- 6553,
[26] A. Ashraf, T. S. Gunawan, B. S. Riza, E. V. Haryanto, and Z. Janin, “On the review of image and video-based depression detection using machine learning”, Indonesian Journal of Electrical Engineering and Computer Science (IJEECS), vol. 19, no. 3, pp. 1677-1684, 2020,
[27] W. Zheng, L. Yan, C. Gou and F. -Y. Wang, “Graph Attention Model Embedded With Multi-Modal Knowledge For Depression Detection”, 2020 IEEE International Conference on Multimedia and Expo (ICME), 2020, pp