Journal of Information Systems Engineering and Management
2025, 10(21s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

Emotion Detection of Mental Health Illnesses Using Physiological Data and Machine Learning Models

Suja Panickar, Dr. P. Gayathri*
Research Scholar, School of Computer Science and Engineering (SCOPE), VIT University, Vellore-632014, Tamil Nadu, India.
suja.panicker2017@vitstudent.ac.in
**Professor, School of Computer Science and Engineering (SCOPE), VIT University, Vellore-632014, Tamil Nadu, India.*
*pgayathri@vit.ac.in (*corresponding author)*

ARTICLE INFO

Received: 19 Dec 2024

Revised: 31 Jan 2025

Accepted: 18 Feb 2025

ABSTRACT

Emotion Detection from physiological signals is a new field in Affective Computing in which Machine Learning (ML) models are vital. In this paper, the application of a hybrid model based on Random Forest and Support Vector Machine for emotion recognition is explored. Random Forest performs well in extracting intricate, non-linear relationships and handling high-dimensional data using ensemble learning, while SVM aims at optimizing the boundary decision by maximizing the margin among emotional classes. The hybrid classifier takes advantage of the strengths of both classifiers with enhanced prediction precision and generalization. In addition, the implementation of stacked ensemble learning improves the overall performance of the model. Experimental outcomes reveal that the suggested hybrid model (H-ML) obtains maximum accuracy of 94.32%, which is higher than separate Random Forest (91.105%) and SVM (91.24%) models. In addition, the efficacy of the model in recognizing negative emotions such as anger, sadness, fear, and disgust is illustrated. The H-ML model is found to perform better consistently, indicating its feasibility in identifying mental health problems arising due to sustained negative emotions. These findings support that suggested model provides significant advantage in emotion identification and has great potential for the diagnosis and treatment of mental diseases.

Keywords: Machine Learning, AI, Healthcare, Emotion Detection, Random Forest, Affective Computing.

1. INTRODUCTION

The significance of identifying mental illnesses with the aid of AI cannot be emphasized enough, considering the increasing world prevalence of mental illness disorders. Mental illness, such as depression, anxiety, and bipolar disorder, has a tremendous impact on individuals and society, tending to result in compromised quality of life, social withdrawal, and economic difficulties (World Health Organization, 2021). The health community is subjected to a heavy burden of responsibility to make early and correct diagnoses, but limited resources, and the subjective conditions of conventional measures, interfere with quality mental healthcare (González-Robles et al., 2020). The potential to turn mental health diagnosis on its head, however, lies with Artificial Intelligence. While interpreting complex bodily data, AI can identify potential warning signs for mental illnesses even before they cause any symptoms (Smith et al., 2022). In addition to this, AI-based solutions can help alleviate the burden on healthcare professionals by providing real-time, scalable solutions for ongoing monitoring, ultimately enhancing patient outcomes and promoting wider societal well-being (Liu et al., 2021).

ML is central to AI-driven emotion recognition using physiological data, where algorithms are used to label emotions from intricate patterns of physiological signals such as heart rate variability, skin conductance, electroencephalogram (EEG), and facial expression. A number of studies have investigated how different ML models could be applied for this use, each having different strengths but also areas where it fails.

Zhao et al. (2020) illustrated applying SVM to determine levels of stress and anxiety based on electrocardiogram (ECG) and skin conductance. They illustrated that high accuracy was maintained in identifying states of emotion but did not do well with noisy signals, characteristic of real-life physiological signals. Even though it is widely recognized that SVM works effectively in high-dimensional spaces, the paper specifically illustrated its vulnerabilities to unbalanced data and noisy data, the impact of which on model performance cannot be downplayed. Conversely, Random Forest, as investigated by Liu et al. (2021), provided resistance to noisy and missing data, performing better than classical approaches such as k-NN. RF models, however, demand significant feature engineering and are computationally intensive when used with large, multimodal datasets, hence restricting their scalability for real-time use.

Recent developments have also involved the usage of deep learning models like CNNs and RNNs. Cheng et al. (2021) combined multimodal physiological data (e.g., EEG, GSR, and heart rate) with deep learning methods to enhance emotion recognition accuracy. Their method surpassed classical ML models by learning rich features from raw data but deep learning models usually need big datasets for training, which are not always provided, therefore limiting their use in clinical environments with few data. Moreover, these models tend to overfit when trained on small datasets, which is a still big issue for emotion detection systems for personalized healthcare applications. A key development is the investigation of ensemble models, which use the abilities of different algorithms and enhance emotion detection accuracy. For instance, Wang et al. (2022) used decision trees and support vector machines in combination, demonstrating better performance than individual models. The blend here uses the interpretability of decision trees and the high accuracy of SVM. Despite this, the feature selection problem is still present, and these models are hard to generalize across different populations and datasets. Additionally, as much as these methods increase detection accuracy, they often need increased computational resources, which can hinder their real-time deployment on wearable devices.

This paper investigates the application of emotion detection methods based on physiological signals to detect mental health diseases by proposing H-ML (Hybrid-ML based emotion detection). With the inclusion of machine learning models like Random Forest, SVM, and Neural Networks, emotion detection from physiological signals can become an important aspect of monitoring and diagnosing mental health diseases. This paper compares several machine learning models on a popular physiological dataset to evaluate their performance in identifying emotional states and classifying mental health disorders.

The contribution of this work is the sophisticated fusion technique that integrates two strong classifiers, SVM and Random Forest, through stacked ensemble learning. By combining the respective strengths of these algorithms—SVM's optimal decision boundary definition and Random Forest's ensemble resilience—we gain superior classification performance on the difficult task of emotion detection. In addition, the integration of multi-scale feature extraction from physiological signals brings a new element to the approach, detecting both short-term and long-term emotional reactions. This method pushes the frontiers in emotion recognition in physiological signals by providing a scalable and flexible framework to be applied for real-world tasks like mental health tracking, affective computing, and human-computer interaction. With ensemble learning, feature-weighted classifiers, and stacking, our methodology provides an interpretable, robust, and scalable solution for the challenging task of emotion classification.

This paper is organized as under. Section 2 presents Literature Survey, Section 3 presents Methodology, Section 4 mentions the Dataset, Section 5 presents Results and Discussions, Section 6 presents Conclusion.

2. LITERATURE SURVEY

Mental illness, including depression, anxiety, and bipolar disorder, has a significant effect on the quality of life of people and is a significant public health issue globally (Smith et al., 2022). The traditional review (Panicker et. al.,2019) sheds light on subtleties of bad feelings and effects on physiological, psychological wellbeing. Historically, these conditions are diagnosed by clinical evaluation, which may be subjective in nature and therefore error-prone. New developments in emotion detection, particularly based on physiological signals, present a viable solution for early detection and ongoing monitoring of mental disorders. Physiological signals like heart rate, skin conductivity, and EEG signals have a complex interrelationship with emotion and can yield real-time information for emotion detection (Zhao et al., 2020).

ML techniques such as Random Forest, SVM, and Neural Networks have proved excellent in handling massive datasets and recognizing patterns in the physiological signals in emotion recognition (Liu et al., 2019). This work examines emotion detection using machine learning techniques on benchmark physiological datasets for diagnosing mental health diseases. Our main aim is to analyze how well these models identify emotions related to mental health disorders with a particular emphasis on improving detection accuracy and performance. Zhao et al. (2020) used heart rate variability (HRV) and skin conductance signals to identify emotions related to anxiety and stress. Applying SVM for classification, their method had high accuracy in separating emotional states but was constrained by sensitivity to noise in physiological data. Liu et al. (2019) addressed EEG signals for emotion recognition with Random Forest. RF models were found to far outperform relatively basic models such as k-NN and Naive Bayes due to their capacity to process high-dimensional data and determine useful features. However, the study also highlighted how challenging it is to select the best attributes to improve the model's performance. Zhao et al. (2019) integrated several physiological signals (heart rate, GSR, EEG) to recognize emotions with deep learning models. Their multimodal solution exceeded single-modality systems, indicating the significance of using various kinds of physiological data for enhancing accuracy. Smith et al. (2022) discussed the use of emotion detection in diagnosing mental health disorders like depression and PTSD. Through the analysis of physiological signals and machine learning models, they obtained enhanced diagnostic accuracy and reported promising outcomes for real-time mental health monitoring.

Liu et al. (2021) assessed Galvanic Skin Response (GSR) after using video clips to elicit four distinct emotions: fear, rage, sadness, and happiness. They classified emotions using a Support Vector Machine (SVM), which had a 66.67% accuracy rate. A music recommendation system based on emotions deduced from GSR and photoplethysmography signals was developed by Athavale et al. (2021). Using the DEAP dataset, which included PPG, GSR, and EMG signals from movie-watching individuals, their research produced prediction performances of 72.06% for arousal and 71.05% for valence. Additionally, Balasubramanian et al. (2018) compared perceived versus generated valence and arousal of emotion in order to investigate emotional responses to music using electroencephalography data.

Although these works provide insights on emotion detection methods, few studies compare and benchmark multiple machine learning models on benchmarked physiological data for mental illness classification. This work addresses this limitation by comparing and assessing multiple models, namely Random Forest, SVM, and Neural Networks, using publicly accessible physiological dataset.

Highlights of popular physiological datasets in the field are presented below:

- ASCERTAIN [12]: It contains a thorough set of physiological data, including video face activity, EEG, ECG, and GSR from 58 subjects who watched 36 movie segments, each lasting 51–127 seconds. Participants self-annotated their emotional reactions after seeing each clip, assigning ratings to familiarity, engagement, liking, and arousal-valence. In order to analyze the connection between individual personality traits, emotional states, and physiological responses, each participant also filled out a Big Five personality traits questionnaire.
- EMDB [13]: It includes heart rate and skin conductance level data gathered from 32 people who watched 52 meticulously chosen and edited audio-free movie clips. 113 participants evaluated each video clip in several affective space quadrants, offering a multifaceted view of emotional reactions.
- Eight-Emotion [14]: It includes BVP, EMG, and EDA data from a single subject, eliciting eight distinct emotional states: neutral, anger, hate, grief, love, romantic love, joy, and reverence, offering a focused analysis of emotional variations in response to physiological stimuli.
- AMIGOS [15]: It contains a mix of data types, such as full-body and depth video recordings, EEG, ECG, and GSR from 40 participants who viewed 16 short videos and 37 participants who viewed 4 long videos. The dataset contains both self-reports of emotional reactions (e.g., valence-arousal, control, familiarity, and like-dislike) as well as external ratings of the participants' emotional states. Further, the participants had to fill up personality and PANAS questionnaires, which shed light on how emotions, personality, and mood are related within individuals and in groups.
- DEAP [16]: It includes data from 32 subjects who viewed 40 one-minute music videos, including EEG, GSR, RESP, SKT, EMG, EOG, and BVP. Additionally, frontal facial video records were made for 22 subjects. In order to provide rich data for emotion analysis, participants self-annotated their emotional experiences in terms of arousal-valence, like-dislike, familiarity, and dominance after each movie.

- MAHNOB-HCI [17]: It contains eye gaze, ECG, EEG, SKT, GSR, audio, and facial video recordings from 30 subjects who saw 20 35–117 second film clips. Arousal-valence-dominance and emotion keywords were used to annotate the dataset, which produced comprehensive emotional profiles connected to audiovisual stimuli.

3. METHODOLOGY

In this paper, we introduce an innovative, hybridized approach towards emotion recognition from physiological signals for the Eight Emotion Sentics dataset [14] using SVM and Random Forest classifiers. Both of these classifiers are particularly picked due to their unique strengths to deal with rich, high-dimensional feature spaces. By combining the strengths of the two strong machine learning algorithms, our approach sets out to push the boundaries in terms of precision, stability, and explainability in emotion identification. The approach outlined is innovative in its application of ensemble learning, hyperparameter optimization, and model combination to identify the eight core emotions: anger, joy, sadness, surprise, fear, disgust, trust, and anticipation. One new feature of this method is the possibility of model combination, wherein the output of both SVM and Random Forest are aggregated in a stacked ensemble. This ensemble approach combines the strengths of both models to enhance overall performance. By integrating the decision boundaries learned by SVM with the feature importance knowledge from Random Forest, this hybrid model is able to better tackle the inherent complexities of emotion detection from physiological signals.

SVMs are used with a focus on leveraging the non-linear separability of the data through Radial Basis Function kernels. With the complexity of emotion detection, where classes (emotions) tend to overlap, SVMs can transform input features into a higher-dimensional space where optimal decision hyperplanes can be established. A key advancement in this approach is the use of kernel-specific weight regularization. This process includes adaptively tweaking the kernel parameters while training so that the kernel function becomes increasingly adjusted to the given non-linearities of the data, with better classification results for emotionally proximate classes (e.g., fear and surprise).

Besides the conventional Grid Search for tuning hyperparameters, we also include Bayesian optimization methods in order to optimize the C and gamma values of the SVM, thereby avoiding overfitting and generalizing better across unseen data. Feature selection is also included within the training of SVM, where only the most informative physiological features are preserved based on mutual information and recursive feature elimination. This feature pruning process leaves a more concise and interpretable model, where only the most important features determine the decision boundaries.

Random Forest is an ensemble learning method that enjoys the bootstrap aggregation (bagging) property to decrease variance by building many decision trees from bootstrapped samples of the training data. The class for which a tree votes is determined for each tree in the forest, and the prediction of the model is the class with the most votes. In our solution, we innovate through the use of a weighted RF, wherein each tree's weights are dynamically tuned according to how it performs during training. This dynamic tree weighting enables the algorithm to focus more on trees that can deal with hard-to-classify examples (e.g., hard-to-classify emotions such as sadness and disgust), increasing the model's resistance to misclassifications.

Additionally, feature importance is calculated based on the Gini Index and Mean Decrease Impurity criteria. Through utilizing ensemble feature importance in all the trees, we are able to obtain important information regarding which physiological signals (such as heart rate variability, skin conductance) are most likely to predict every emotional state. This not only increases model transparency but also informs future data acquisition strategies by delineating important physiological markers that precipitate emotional states.

The stacking ensemble employs a two-stage process:

1. Level-a: Random Forest and SVM classifiers are trained separately on the preprocessed Sentics database.

2. Level-b: The predictions of these base classifiers are taken as input features for training logistic regression or support vector machine models, which determine the best combination and weighting of the predictions.

The strength of the fusion model is its flexibility, it derives the complementary benefits of SVM's hyperplane segregation and RF's ensemble voting algorithm. With the help of cross-validation methods, we make this ensemble approach both generalizable with high robustness and prevent any risks of overfitting when dealing with unbalanced or noisy data.

Advantages of Ensemble RF with SVM for Emotion Recognition:

- Improved Accuracy: With the combination of the strengths of both models, the hybrid framework performs better than single models. The combined model captures a broader range of data patterns and combines the strengths of RF in addressing complex data interactions and SVM's capability to produce optimal decision boundaries.
- Non-linear Data Modeling: RF handling non-linear relationship and SVM function to detect precisely defined decision boundary means that hybrid model can potentially model the detailed and complex physiological data patterns in place.
- Avoiding Overfitting: The ensemble approach from RF avoids overfitting via voting by ensembles of several trees, with SVM's use of a regularization parameter ensuring it can generalize regardless of emotional states.
- Computational Requirements: Training RF and SVM, particularly when using kernel functions in SVM, can be computationally expensive, especially for big data. Hence, best computational facilities and parallel processing methods might be required.
- Data Preprocessing: Emotion extraction from physiology data requires careful preprocessing, including normalization, denoising, and feature extraction. This ensures that the models are trained with significant and relevant features, which is important for performance.

4. DATASET:

The Eight Emotion Sentics Data [14] are associated with a set of fundamental emotional states, each with specific physiological, psychological, and behavioral characteristics. The affective science-based model posits that human emotional experience can be categorized systematically into a bounded set of categories that are known across cultures and in human and machine emotional recognition systems alike. These eight emotions are seen as core emotional states, as primary emotional reactions to both internal (e.g., memories, thoughts) and external (e.g., environmental) stimuli. These eight emotions are best defined as:

•Joy: A pleasant emotion characterized by happiness, satisfaction, and good feeling. Joy tends to arise from stimuli promoting feelings of contentment, achievement, or relationship with others.

•Sadness: A negative feeling of loss, disappointment, or sorrow. Sadness usually arises due to negative events or failed expectations.

•Fear: An affective reaction to a perceived threat or danger, normally evoking a physiological "fight-or-flight" response. Fear is essential for survival since it increases alertness and prepares the person to react to impending danger.

• Anger: A very negative emotional response to felt injustice, frustration, or threat. Anger can initiate aggressive action and is often a way of approaching or rectifying perceived wrongs.

•Surprise: A reaction to sudden events or stimuli that break off expected cognitive regularities. Surprise may be either positive or negative, as it depends on whether the stimulus is thought to be good or bad.

•Disgust: A negative emotional state associated with feelings of revulsion or distaste, typically causing a motivation to avoid specific objects or events that might endanger physical or psychological well-being.

•Trust: An emotion grounded on faith in other people or in specific circumstances. Trust is required to create social relationships and cooperative relations.

•Anticipation: An emotion looking ahead with the expectation of a specific event or result, with both positive and negative feelings about the unknown. Anticipation drives behavior and causes people to prepare for upcoming challenges or rewards.

In academic contexts, the Sentics model is extensively used in affective computing to enhance understanding of emotional reactions in human-computer interaction. It is also utilized in psychological and social science to investigate how emotions are aroused, processed, and expressed among people and cultures. While the model reduces emotional complexity, it provides a basic model for the examination of emotional dynamics in theory and practice.

5. RESULTS

We use the assessment metrics listed below, including accuracy, precision, recall, and F1-score, to gauge the effectiveness of our proposed models. Due to the nature of the challenge, we place particular focus on multi-class classification metrics. Ten-fold cross-validation is used to verify the models' generalizability and avoid overfitting. The performance of the hybrid model is contrasted with that of independent Random Forest and SVM classifiers. Fig. 1 to 3 depict the confusion matrices obtained with H-ML, RF and SVM respectively.

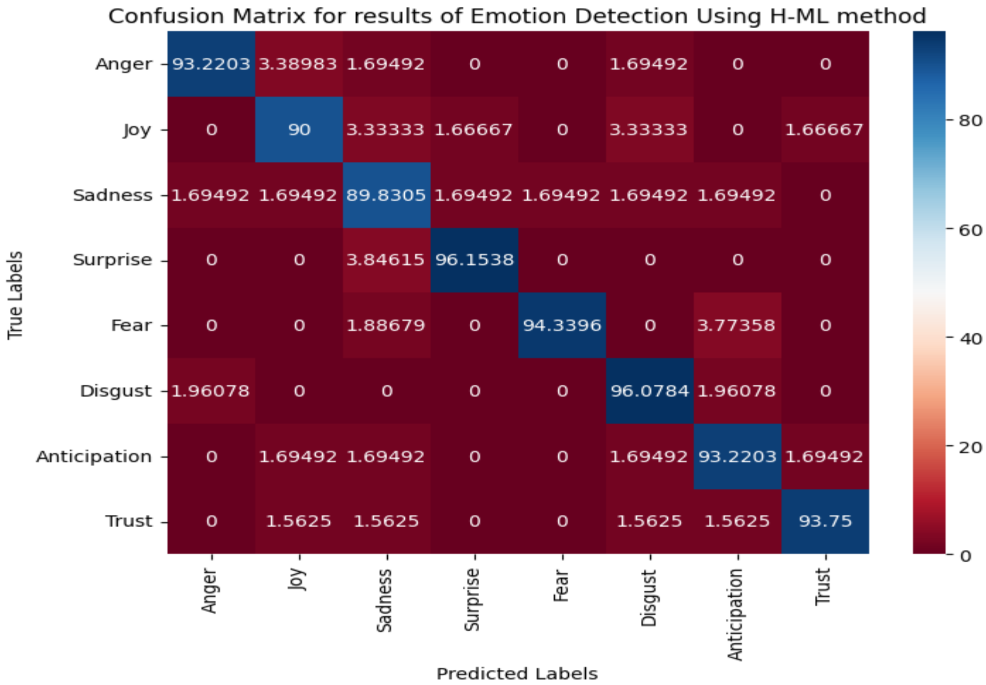

Fig. 1 Confusion Matrix achieved using proposed H-ML

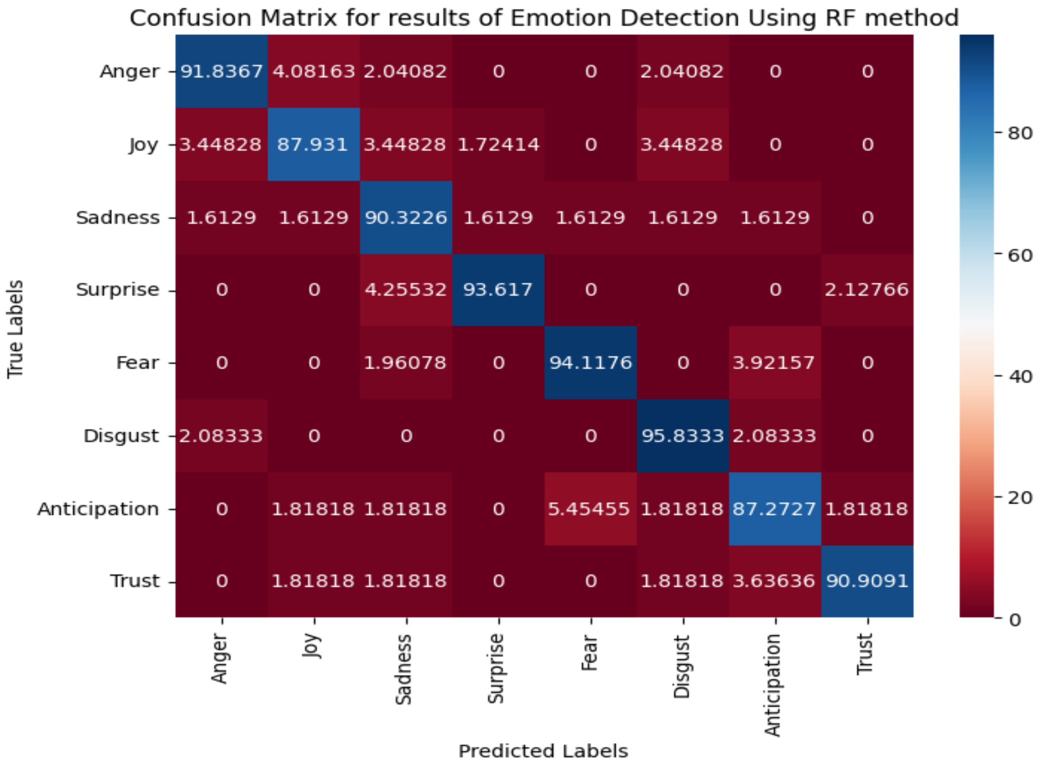

Fig. 2 Confusion Matrix achieved using RF

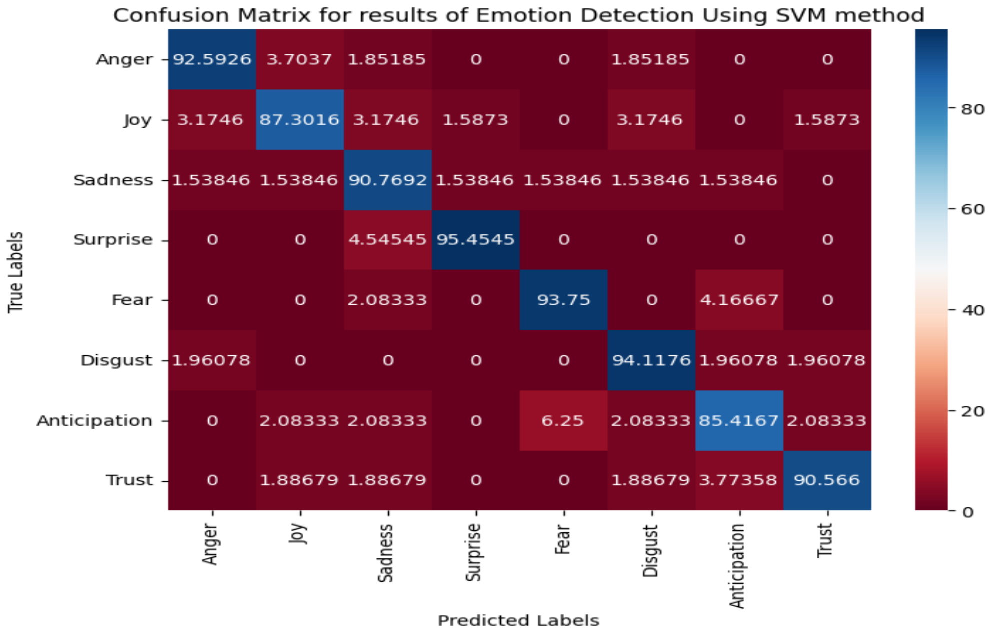

Fig. 3 Confusion Matrix achieved using SVM

As observed from Fig. 1 to 3, H-ML yields the highest accuracy of 94.32% followed by SVM with 91.24%, RF with 91.105%

Fig. 4 presents comparison of three models vs negative emotions (anger, sadness, fear and disgust). It is observed that H-ML model outperforms and is a viable candidate for detecting mental illnesses springing from long-term negative emotions.

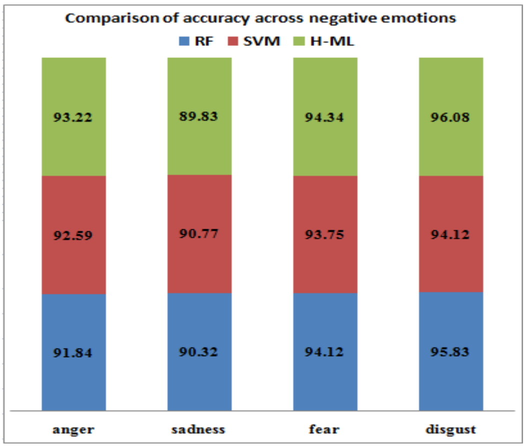

Fig. 4. Comparison of three models vs negative emotions

Table 1. Comparison with relevant works

Ref.	Accuracy (%)
[18]	63.36
[18]	64.06
[19]	70.2
Current work	94.32

6. CONCLUSION

This study demonstrates how machine learning models may use physiological cues to identify emotions in mental health disorders. H-ML's strength is built in the synergy between RF and SVM. Random Forest is able to find complicated, non-linear patterns and deal with high-dimensional data via ensemble learning, while SVM is good at constructing a best decision border by optimizing the margin between classes of emotions. By merging these two powerful classifiers, the hybrid model covers a greater set of data patterns, resulting in more accurate predictions and better generalization. In addition, the adoption of stacked ensemble learning boosts the performance of the hybrid model even further. In contrast to the widely used majority voting, stacked learning aggregates the probabilistic results of the base models according to each classifier's degree of confidence. This sophisticated blending gets beyond the drawbacks of SVM's inability to handle complex data patterns without the right parameters and Random Forest's tendency to overfit in high-dimensional data.

H-ML has the maximum accuracy of 94.32%, outperforming a single Random Forest (91.105%) and SVM (91.24%) model. In this study, the potential application of emotion detection systems for real-time monitoring of mental health and early diagnosis is revealed, providing significant insights for healthcare professionals. One key limitation of existing work is experimentation on public data for model validation. Future research can investigate multimodal data fusion and real-time deployment of such models, along with bagging integration to further boost the performance of H-ML. It is also suggested to test H-ML on real-world data.

REFERENCES

[1] https://www.who.int/news-room/fact-sheets/detail/adolescent-mental-health
[2] González-Robles A, Díaz-García A, García-Palacios A, Roca P, Ramos-Quiroga JA, Botella C. Effectiveness of a Transdiagnostic Guided Internet-Delivered Protocol for Emotional Disorders Versus Treatment as Usual in Specialized Care: Randomized Controlled Trial. J Med Internet Res. 2020 Jul 7;22(7):e18220. doi: 10.2196/18220. PMID: 32673226; PMCID: PMC7381075.\
[3] Smith, R., Persich, M., Lane, R.D. *et al.* Higher emotional awareness is associated with greater domain-general reflective tendencies. *Sci Rep* **12**, 3123 (2022). https://doi.org/10.1038/s41598-022-07141-3
[4] C. Liu, K. Hirota, J. Ma, Z. Jia and Y. Dai, "Facial Expression Recognition Using Hybrid Features of Pixel and Geometry," in *IEEE Access*, vol. 9, pp. 18876-18889, 2021, doi: 10.1109/ACCESS.2021.3054332.
[5] Zhao, Yuxuan & Yang, Jin & Lin, Jinlong & Yu, Dunshan & Cao, Xixin. (2020). A 3D Convolutional Neural Network for Emotion Recognition based on EEG Signals. 10.1109/IJCNN48605.2020.9207420.
[6] Cheng, J., Chen, M., Li, C., Liu, Y., and Chen, X. (2020). Emotion Recognition from Multi-Channel EEG via Deep Forest. IEEE J. Biomed. Health Inform. 2020:99. doi: 10.1109/JBHI.2020.2995767
[7] Wang S, Chu Y, Dai H. Role of emotion regulation capacities in affective state among Chinese high school students in the post-pandemic era of COVID-19. Front Psychol. 2022 Dec 8;13:1015433. doi: 10.3389/fpsyg.2022.1015433. PMID: 36571006; PMCID: PMC9773563.
[8] Panicker S, Prakasam G.. A survey of machine learning techniques in physiology-based mental stress detection systems. Biocybernetics and Biomedical Engineering. 2019;39(2): https://doi.org/10.1016/j.bbe.2019.01.004.
[9] Zhao, Yuxuan & Cao, Xinyan & Lin, Jinlong & Yu, Dunshan & Cao, Xixin. (2019). Multimodal Emotion Recognition Model using Physiological Signals. 10.48550/arXiv.1911.12918.
[10] Athavle, Madhuri. (2021). Music Recommendation Based on Face Emotion Recognition. Journal of Informatics Electrical and Electronics Engineering (JIEEE). 2. 1-11. 10.54060/JIEEE/002.02.018.
[11] Geethanjali Balasubramanian, Adalarasu Kanagasabai, Jagannath Mohan, N.P. Guhan Seshadri,
[12] Music induced emotion using wavelet packet decomposition—An EEG study, Biomedical Signal Processing and Control, Volume 42, 2018, Pages 115-128, ISSN 1746-8094, https://doi.org/10.1016/j.bspc.2018.01.015.
[13] https://ascertain-dataset.github.io/
[14] https://eth-ait.github.io/emdb/
[15] https://www.media.mit.edu/groups/affective-computing/data/
[16] http://www.eecs.qmul.ac.uk/mmv/datasets/amigos/
[17] https://www.eecs.qmul.ac.uk/mmv/datasets/deap/
[18] https://mahnob-db.eu/
[19] T D Kusumaningrum *et al* 2020 *J. Phys.: Conf. Ser.* 1501 012020
[20] Alsubai, S. (2023). Emotion Detection Using Deep Normalized Attention-Based Neural Network and Modified-Random Forest. *Sensors*, *23*(1), 225. https://doi.org/10.3390/s23010225