CatForest: Deep Contextual Sentiment Modeling for Mental Health Detection from Social Media

Afsana Akter Tusa[1], Md Jakaria Zobair[1] *, Toufiq Mehraz[1], Mahir Ashef[2], Refat Tasfia Orpa[3], and Md Arifuzzaman[4]

[1] MARS Lab, Department of Computer Science and Engineering, Daffodil International University, Dhaka-1216, Bangladesh
[2] Department of Computer Science and Engineering, Bangladesh University of business and technology, Dhaka-1216, Bangladesh
[3] Department of Computer Science and Engineering, Hajee Mohammad Danesh Science and Technology University, Dinajpur-5200, Bangladesh
[4] Department of Mathematics, East Texas A & M University, Commerce, Texas, USA
afsanatusa38@gmail.com, jakariacse16@gmail.com, toufiqmehraz121@gmail.com, mahirashef30@gmail.com, refattasfiacse16@gmail.com,arif.ringku@gmail.com

Abstract. Mental health is one of the most important part of our overall well-being. Anxiety, depression, and emotional distress are becoming more common issues day-by-day for all age groups. People share their daily thoughts, emotions, and experiences on social media platforms such as Facebook, Twitter, and Instagram, which represent their mental states. In this study, we explored how machine learning can detect these signals by applying models like Logistic Regression (52.4% accuracy), Random Forest (97.08%), XGBoost (96.1%), and Support Vector Machines (79.6%). Among these, a newly developed hybrid model, CatForest, which combines the strengths of Random Forest with additional optimizations, performed the best, achieving 97.09% accuracy. CatForest shows strong potential for identifying emotional states through social media, which can be a promising tool for early and accessible mental health support. Dataset characteristics, ethical considerations, and evaluation processes are discussed to ensure reproducibility and reliability . . .

Keywords: Mental Health Detection, Social Media, Machine Learning, Ensemble Models, Random Forest, SVM

1 Introduction

"Mental health" is a crucial aspect of overall well-being, through which we can understand someone's feelings, thoughts, and behaviors. People, day by day, feel more depressed leading emotional distress. According to the World Health Organization (WHO), one in every seven people lives with a mental health problem [1], with adolescents and young adults being particularly vulnerable [2].

* Corresponding author: jakariacse16@gmail.com

M. S. Arefin et al. (eds.), *Proceedings of the International Conference on Intelligent Data Analysis and Applications (IDAA 2025)*, Advances in Intelligent Systems Research 206,
https://doi.org/10.2991/978-94-6239-664-7_26

Anxiety, depression, and related disorders have become pressing global challenges. So urgent action is needed for early detection and intervention of mental health. With the rise of social media platforms such as Facebook, Twitter, and Instagram, people increasingly share their emotions and experiences online, and it carries valuable signals of their mental state. Detecting such hidden emotional patterns through natural language processing (NLP) and machine learning (ML) has become a promising direction. In this respect some powerful algorithms like Logistic Regression, SVM, and BERT have shown potential for identifying stress, depression, or anxiety. Recent research also highlights the effectiveness of hybrid and ensemble approaches in healthcare-related prediction tasks: Zobair et al. [3] proposed a hybrid ViT-GRU model for breast cancer detection to address class imbalance, Zobair et al. [4] developed a lightweight and generalizable deep learning model for COVID-19 prediction from X-ray images, and Hasan et al. [5] applied ensemble-based machine learning to predict mothers' delivery modes. These works demonstrate the strength of combining models to improve robustness and accuracy.

Motivated by this, we propose CatForest, a novel hybrid ensemble model that integrates Random Forest and CatBoost to detect mental health signals from social media more effectively, uncovering subtle emotional cues that simpler models often miss.

2 Literature review

In a paper, Saeed et al. [6] developed a multi-modal deep learning model combining BiLSTM text analysis with temporal posting patterns and attention mechanisms to detect depression and anxiety from over 700k Reddit posts. The model achieved 74.5% accuracy, which indicated the importance of combining content and timing for early mental health detection. Ding et al. [7] compared machine learning and deep learning models for mental health detection on social media. Deep learning models such as ALBERT and GRU outperformed traditional ML models, with ALBERT achieving an F1-score of 0.96 and an AUC of 0.99. Garg [8] reviewed numerous studies and highlighted how AI techniques, including ML and DL, are being used to detect stress, depression, and suicidal tendencies from online posts. Bhuiyan et al. [9] proposed a hybrid CNN–BiLSTM model and used SHAP (SHapley Additive exPlanations) for interpretability, improving accuracy from 92.81% to 94.29%. Shah et al. [10] explored the effectiveness of fine-tuned large language models (LLMs), such as GPT-3.5 Turbo and LLaMA2-7B, achieving 96% accuracy in classifying depressive vs. non-depressive posts. Hasan and Saquer [11] analyzed multiple transformer and LSTM-based models using Reddit data, where RoBERTa achieved 93.22% accuracy, 93.14% F1-score, and 94.41% recall. Zhao and Kerz [12] explored hybrid and ensemble models for multiclass mental health prediction, while Renjith et al. [13] proposed an ensemble technique for suicidal ideation detection, achieving strong performance. Ezerceli and Dehkharghani [14] used deep neural networks for mental disorder and suicidal ideation detection, achieving F1-scores up to 0.97. Tavchioski et al. [15] used transformers such as BERT, RoBERTa, BERTweet, and Mental-BERT for depression detection on Reddit and Twitter, finding ensemble models most effective. Pokrywka et al. [16] evaluated transformer models—including DeBERTa and GPT-4o—for suicide risk detection, where fine-tuned GPT-4o achieved the highest weighted F1 of 75.5%. Rahman et al. [17] surveyed mental health detection techniques on online social networks, and they observed issues such as language barriers, privacy concerns, and limited features. Yazdavar et al. [18] proposed a multimodal framework combining text, images, and user interactions, achieving 90% accuracy on data collected from Twitter. Ansar et al. [19] reviewed studies from 2015–2025 and found that social media can raise awareness and reduce stigma but also spread misinformation. Febrian and Awangga [20] conducted a systematic literature review, finding that XGBoost achieved the highest average accuracy (80.1% ± 4.2%) for predicting mental health indicators from digital physical activity data. Espino Carrasco et al. [21] reviewed 62 studies on AI-assisted mental health interventions, emphasizing ethical considerations and personalization. Li et al. [22] reviewed 64 studies on digital data sources in public health, highlighting social media's role in disease prevention and health behavior analysis. Singh et al. [23] proposed an embedded LSTM framework achieving 94.88% accuracy in detecting depression from Reddit text, while Ghosal and Jain [24] used fastText embeddings with XGBoost to detect depression, achieving 71% accuracy. Finally, De Choudhury et al. [25] analyzed Twitter data from users with self-reported depression, using behavioral, emotional, linguistic, and network features to achieve 72.38% accuracy.

3 Methodology

In this section we will discuss the overall system flow where at first we collect the dataset, then we apply some preprocessing techniques and then extract features. After that, training testing splitting is performed before applying different models. Overall system flow is given in figure 1

3.1 Dataset Description

We use the "Social Media Usage and Emotional Well-Being" dataset from Kaggle [26] , which includes both quantitative social media activity metrics and qualitative emotional indicators. The dataset contains 1,000 instances and 10 attributes, covering demographic information (age, gender), platform usage (daily usage time in minutes, posts per day, likes and comments received, messages sent), and an emotional label representing the dominant emotion expressed by each user.A class distribution analysis reveals a slight imbalance: Neutral and Happiness appear more frequently than minority classes such as Fear and Anger. This imbalance indicates that classification performance may vary across classes, and proper evaluation metrics (e.g., F1-score) are necessary.Because the dataset is relatively small, there is a risk of model overfitting. To mitigate this, we apply systematic preprocessing techniques, regularization, and cross-validation.

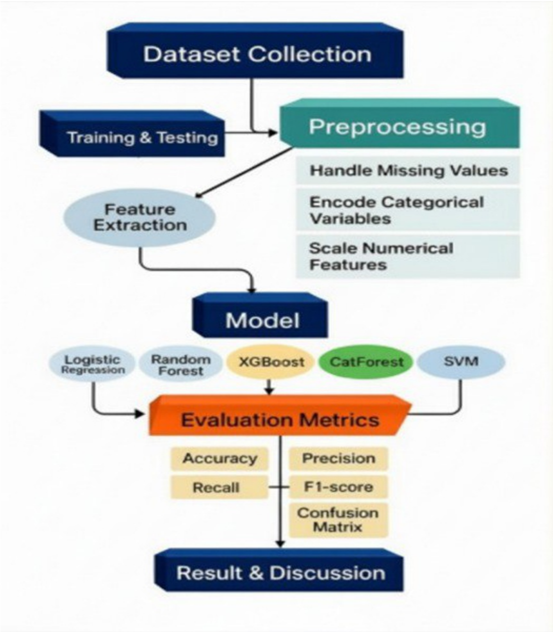

Fig. 1. Overall System Flow.

3.2 Data Preprocessing

Atfrist, we identify the missing rows and remove the inconsistent information from the dataset. We also check the duplicate values to maintain the integrity of the dataset. We convert the Gender and Platform to categorical form as well as the target coloum Emotion is converted into label-encoding. We also standarized the numerical features like the Daily Usage Time and Posts Per Day so that no signal feature would dominate the learning process. All preprocessing steps were applied consistently across models to ensure fair comparisons.

3.3 Feature Extraction

we applied Recursive Feature Elimination (RFE) using a Random Forest classifier to extract the most relevent features form the dataset. This method evaluates the importance of each feature for predicting emotions and we remove the less important columns which has smaller contribution to the target class. From the eight original features, RFE identified four as particularly valuable for predicting mental health: gender, hours spent online, likes, and comments. The remaining features—age, platform, posts, and shares—were excluded, as they offered minimal predictive value within our dataset.

3.4 Model Training and Testing

We split the dataset into training and testing sets with an 80:20 stratified split to preserve the original class distribution. Although some emotions appeared less frequently, we kept the sampling natural to avoid introducing artificial bias. To make sure our models generalized well and to reduce overfitting on this relatively small dataset, we applied 5-fold cross-validation during training.

3.5 Description of The Model

In this study, we explored several machine learning models to predict emotions from social media behavior.

Logistic Regression It is a statistical model which is mostly used to predict binary class by observing one or more variables. It use a S-shape curve instead of straight line, and gives a probability as output.

XGBoost It is a powerfull model which uses gradient descent. Also the regularization is used here, making it powerful for large, non-linear datasets. But it requires careful tuning and may confuse overlapping emotions.

Support Vector Machine (SVM) It works by finding the best hyperplane that maximizes class separation. It is also a powerful model which is used for classification, regresion as well as outlier detection.

Advanced SVM It is the upgradation of the SVM model where some preprocessing steps like SMOTE for balancing, RFE for feature selection, PCA for dimensionality reduction are used. Also hyperparameter tuning is performed in this model. These improvements make it more robust and generalizable.

CatForest CatForest is a hybrid model that combines CatBoost and Random Forest, leveraging the strengths of both. CatBoost efficiently captures complex feature interactions and handles categorical variables. One the otherhand, Random Forest provides robustness by adding multiple decision trees. Together, they model the relationships between social media usage and emotional well-being effectively and achieve higher accuracy than a single model. The approach is also interpretable through feature importance analysis, and 5-fold cross-validation ensures the model generalizes well without overfitting. CatForest combines RF & CatBoost using optimized weighted fusion ($0.4 \times$ RF + $0.6 \times$ CatBoost)

Algorithm 1: CatForest Hybrid Model

Input: Dataset $D = \{X, y\}$; RF and CatBoost hyperparameters
Output: Predicted labels and evaluation metrics

1. Split the dataset into 80% training and 20% testing.
2. Identify categorical and numerical features.
3. Encode categorical features using `LabelEncoder`.
4. Scale numerical features using `StandardScaler`.
5. Train a Random Forest model on the scaled training features.
6. Train a CatBoost model on the preprocessed (raw + encoded) training data.
7. Compute prediction probabilities from both models on the test set.
8. Combine probabilities using weighted fusion:
   HybridProb = 0.4 × RF + 0.6 × CatBoost.
9. Select the final predicted class using the highest value in HybridProb.
10. Evaluate performance using Accuracy, Precision, Recall, and F1-Score.

3.6 Novelty and Contribution of CatForest

CatForest is a simple but effective hybrid model that blends the strengths of Random Forest and CatBoost. Instead of using a complex stacking method, it combines the prediction probabilities of the two models using a lightweight weighted-fusion approach (40CatBoost).This lets Random Forest reduce variance while CatBoost captures deeper feature interactions, especially in categorical data. To our knowledge, this specific fusion strategy has not been applied before for mental-health emotion detection. The improved accuracy and stability across all classes show that this hybrid design adds meaningful value beyond using the individual models alone.

3.7 Evaluation Metrics

To evaluate the performance of the machine learning models, several standard metrics were employed. These metrics provide insights into the predictive accuracy and the areas where models may misclassify samples. The following metrics were used: Accuracy: Accuracy represents the proportion of correctly predicted instances among all predictions:

$$Accuracy = \frac{TrueP + TrueN}{TrueP + TrueN + FalseN + FalseP} \tag{1}$$

Here TrueP means true positive, TrueN means true negative,FalseP means false positive and FalseN means false negative. Precision (Positive Predictive Value): Precision measures the proportion of positive predictions that are actually correct:

$$Precision = \frac{TrueP}{TrueP + FalseP} \tag{2}$$

This metric is particularly useful when minimizing false positives is critical, such as avoiding misdiagnosis of healthy individuals as mentally ill.

Recall (Sensitivity / True Positive Rate): Recall quantifies the proportion of actual positives that the model correctly identifies:

$$Recall = \frac{TrueP}{TrueP + FalseN} \tag{3}$$

It is important in contexts where reducing false negatives is essential, such as ensuring that at-risk individuals are not missed. F1-Score: The F1-score is the harmonic mean of Precision and Recall:

$$F1 - score = 2 * \frac{Precision * Recall}{Precision + Recall} \tag{4}$$

It provides a balanced measure, particularly useful for imbalanced datasets. v. Confusion Matrix A confusion matrix is a tabular representation of True Positives, True Negatives, False Positives, and False Negatives. It is useful for visualizing where the model is making errors across different classes. vi. ROC-AUC (Receiver Operating Characteristic – Area Under Curve) ROC-AUC measures how effectively a model separates different classes. AUC = 1 → Perfect classifier AUC = 0.5 → Random guessing

4 Result and discussion

We experimented with several machine learning models to evaluate their ability to detect mental health states from social media posts. Here is a table of our used models.

Table 1. Testing Accuracy of Different Machine Learning Models.

Model	**Testing Accuracy(4)**	**Testing Accuracy(6)**	**Testing Accuracy(All)**
Logistic Regression	0.4563	0.5145	0.5243
XGBoost	0.9320	0.9417	0.9612
SVM	0.4466	0.5048	0.7961
Advanced SVM	0.5436	0.7378	0.9320
Hybrid Model (CatForest)	**0.9515**	**0.9612**	**0.9709**

Table 1 shows the testing accuracy of different models for 4-class, 6-class, and all-category setups. Logistic Regression struggled (0.45–0.52), while SVM did slightly better, especially in the all-class case (0.7961). Advanced SVM showed clear improvement, reaching 0.9320 with kernel optimization.Among our models

XGBoost performed very high which is (0.9320-9612), but we proposed another model CatForest,a hybrid model, and got more better result which is (0.9515-0.9709). We use another three models like Logistic Regression, SVM and Advanced SVM. But get best result after combining two models (Random Forest & CatBoost). We also use RFE(Recursive Elimination Factor) and as our dataset is small we see that there is a slight drop in our accuracy. Finally we can say that our hybrid model performs best in our small dataset for detecting mental health signals.

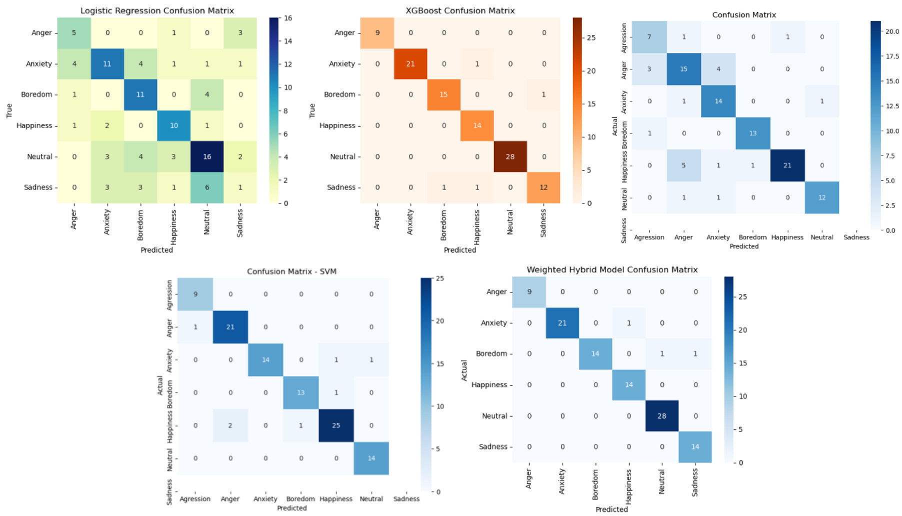

Fig. 2. Confusion matrix comparison across models — (a) Logistic Regression, (b) XGBoost, (c) SVM, (d) Advanced SVM, (e) CatForest (Random Forest + CatBoost)..

Figure 2 shows the difference of confusion matrix for different models. Here five models consist of six emotions are shown. If we compare, here XGBoost performs better specially for Neutral and Happiness but Logistic regression has frequent mix-ups. For SVM we get steady results but when we applied advanced SVM then it reduced confusion between Anxiety and Sadness. Our proposed model CatForest which is a hybrid model stood out as the most balanced which offers strong and consistent predictions across all emotions. It highlights the strength of ensemble approaches.

Figure 3 shows a clear performance on five machine learning models in which one is a hybrid model named CatForest (CatBoost+Random Forest) which delivered the best result. From this chart we can easily see the highlights about different algorithm and say that different algorithm can create a stronger and more accurate predictor. XGBoost and Advanced SVM also performs very good reflecting their strength in tackling complex patterns. On the other hand Logistic regression and SVM perform a bit lag behind, most probably they were

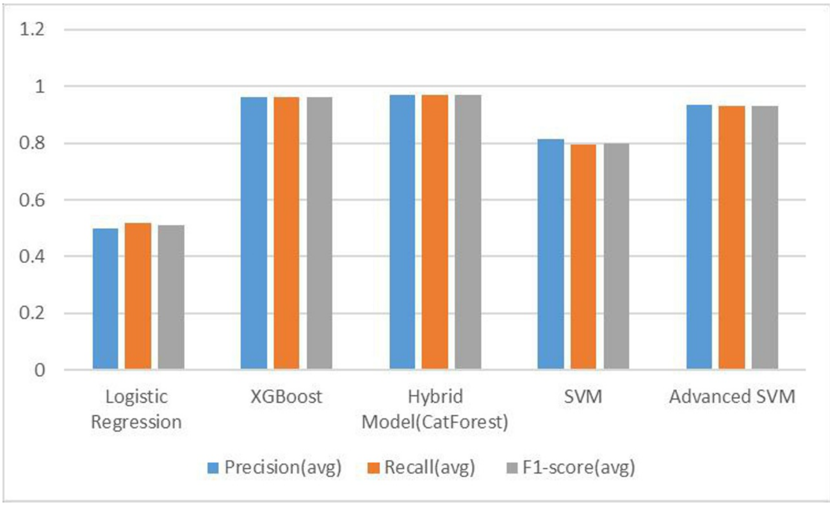

Fig. 3. Precision, Recall and F1-score Comparison of Different Model.

struggling with capturing more complicated relationships in the data. Overall, from this chart we can say a story that using hybrid model boost a model's accuracy and reliability.

Table 2. Comparison of Performance with Existing Studies.

Reference	Dataset	Method	Performance
[6] Qasim Bin Saeed et al. (2025)	Reddit (700k+ posts)	BiLSTM + Temporal features + Attention	Accuracy = 74.5%
[11] Hasan, Saque	Reddit	RoBERTa, LSTM, Hybrid	F1 = 93.14%; Accuracy = 93.22%
[24] Ghosal & Jain (2023)	Reddit (SuicideWatch, Depression)	fastText + TF-IDF + XGBoost	Accuracy = 71%, F1 = 0.71
[25] De Choudhury et al. (2013)	Twitter (1-year activity logs)	Behavioral + Emotional + Linguistic + Network features	Accuracy ≈ 72.38%, Precision = 0.74
Ours	Kaggle: Social Media Usage & Emotional Well-being	**Hybrid CatForest** (CatBoost + Random Forest) vs LR, SVM, XGB	**Accuracy = 97.09%, Precision = 0.97**

Table 2 shows a comparison of past research about mental health detection using social media data. This table highlights the dataset which they were used, method and performance where we can see accuracy, F1 score and others. Previous works mostly relied on Reddit or Twitter and they used BiLSTM techniques with attention, transformer models such as BERT, RoBERTa, fastText with TF-IDF, and feature-based behavioral or emotional models. Their performance is between 71-76.61%. In our research we used the Kaggle "Social media usages & emotional well-being" dataset and applied a hybrid CatForest model (combining CatBoost and Random forest). Our this hybrid model performs very much good comparing previous models, it achieved a very high accuracy 97.09% and precision was 0.97. The Evidences demonstrates that ensemble models can actively capture complex patterns in social media data.

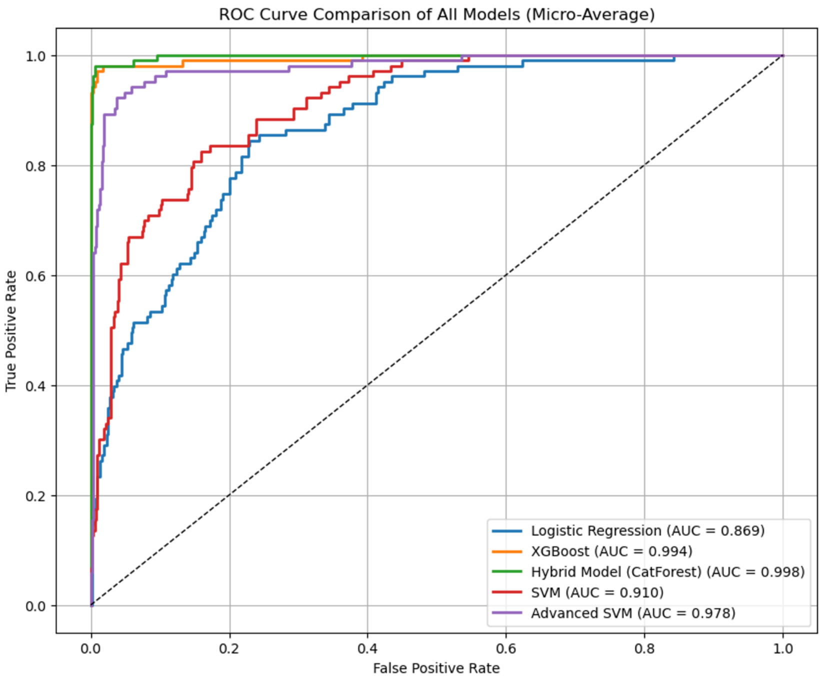

Fig. 4. Comparison of ROC curves for Logistic Regression, XGBoost, CatForest, and SVM models.

Figure 4 shows the mental health classes which are reveled to the readers through all the models by ROC curves. The hybrid model (Catforest) is the most effective among all other models. From the ROC curve we can see its AUC is 0.998. Others models like Advanced SVM and XGBoost also do a better job with AUC 0.978 and 0.994. They also has a very good separation power. Logistic regression and SVM has some lacking but still they offers meaningful insights. For clarity the curves are color coded and the micro average approach ensures a balanced view of all classes. In general, the figure depicts that ensemble and hybrid techniques exemplify pushing the boundaries of prediction. Logistic Regression predicted results with relaticely low accuracy, due to its linearity and inability to trace complex nonlinear social media pathways associated with emotional data. The standard SVM outcomes were moderate but got better with advanced tuning, although the algorithm is limited in handling high-dimensional data and complex feature interactions.Tree-based models, includ- Thus, the best performing models, includ- Besides, the dataset is relatively small, with 1000 samples and limited features, limiting model generalization. And finally, despite the hybrid model achieving the highest level of accuracy (97.09), it presents limitations for implementation in real-time or limited resource applications. Our findings indicate that correctly selecting the model and preprocessing method is vital for detecting social health signals from posts. Logistic Regression, a simple model, was inadequate for dealing with the complexity of the data. Ensemble methods such as random forest and catboost, and their hybrid combination catboost called catforest performed the best (97), and captured subtle emotional cues with reliable accuracy. XGBoost also shows strong results without the highest average accuracy, which is explained by several wrongly classified instances in emotions such as Sadness, meaning that some emotions are easier to confuse than others because of similar words or phrases within the training dataset and the number of instances used to train the model. Standard SVM showed underwhelming results at first but, after performing proper preprocessing (oversampling, PCA, and hyperparameter tuning), its accuracy rose to 93.2, highlighting the significance of preparation work when dealing with high-dimensional textual data. Although these accuracies are high, there was a risk of overfitting because the dataset is small. To mitigate this, 5-fold cross-validation was applied such that the performance metrics are not a result of a particular train-test split but rather generalizable results. The slight class imbalance was maintained by stratified sampling so that there was no artificial bias introduced. The feature importance analysis of CatForest gave interpretable insights based on the final most impacting features on emotional outcome from the social media usage behaviors. Last, all data were anonymized, and no personal identifiers were used, thus tackling privacy and ethical issues related to the use of sensitive social media data. Besides, we tried RFE with selecting 4 and 6 features and with training without RFE. Surprisingly, the accuracy of the model went down with RFE. This is probably due to the fact that the dataset provided is relatively small (about 1,000 rows and 10 columns, including serial number and the target), and further discarding features in a relatively small dataset can discard much important information, on which the effective working of features dependon, thereby worsening the performance. To summarize, advanced models and intelligent preprocessing make a huge difference. CatForest hybrid, in particular, shows high performance in all categories of emotions. Although the detection of certain closely related emotions remains problematic, the obtained results allow one to speak about the potential of ML methods for the task of monitoring the state of mental health with the help of social media activities.

5 Interpretability, Bias, and Ethical Considerations

At first we were testing that which features influenced the model decisions most. After testing we found that the daily usage of time, likes and comments were the strongest signals. This helps us to understand the model behavior more than previous. We also checked for basic demographic bias and did not observe major issues through the small dataset limits deeper fairness testing. As the dataset is publicly available and anyone can access and also it is fully anonymize so no personal information was involved here. Our model is only for research not for any medical or sensitive decision making.

6 Conclusion

Our study demonstrated the applicability of machine learning for the detection of psychosomatic state from social media messages. We discovered that simple models such as Logistic Regression failed to cope with the complexity of online text. On the contrary, ensemble models such as Random Forest, CatBoost, and a carefully tuned Support Vector Machine (SVM) performed well. On the other hand, CatBoost and Random Forest provide enhanced accuracy, indicating their ability to identify subtle emotional cues within posts. Preprocessing and fine-tuning also showed a difference, particularly for SVM; However, some of the emotions such as Sadness and Anxiety were difficult for the model to comprehend and require further experiments. In short, the result obtained proves, with the appropriate models and strategies, machine learning techniques can be useful for social media mental health monitoring. In the future, work will be done on combining different models or adding more contexts to understanding people's online sentiment better.

References

1. World Health Organization. (2022). Mental disorders. Geneva: WHO.Available: https://www.who.int/news-room/fact-sheets/detail/mental-disorders (access on 10 August,2025)
2. World Health Organization. (2023). World Mental Health Day 2023. Geneva: WHO.Available: https://www.who.int/news-room/events/detail/2023/10/10/default-calendar/world-mental-health-day-2023---mental-health-is-a-universal-human-right (access on: 15 August,2025)
3. Zobair, M.J. et al. A hybrid ViT-GRU model for breast cancer detection. *2025 ECCE Conference*, pp.1–7. doi:10.1109/ECCE64574.2025.11013095 (2025)
4. M. J. Zobair, R. T. Orpa, M. Asef, S. M. T. Siddiquee, N. R. Chakraborty and A. Habib, "A lightweight and generalizable deep learning model for the prediction of COVID-19 from X-ray images", *International Journal of Electrical and Computer Engineering (IJECE)*, Vol-14, No-4, doi: http://doi.org/10.11591/ijece.v14i4.pp4068-4077 (2024)

5. Hasan, M., Zobair, M.J., Akter, S., Ashef, M., Akter, N. and Sadia, N.B. Ensemble based machine learning model for early detection of mother's delivery mode. *2023 International Conference on Electrical, Computer and Communication Engineering (ECCE)*, Chittagong, Bangladesh, pp. 1–6. doi:10.1109/ECCE57851.2023.10101558 (2023)
6. Saeed, Q.B. and Ahmed, I. Early detection of mental health issues using social media posts. *arXiv preprint arXiv:2503.07653* (2025)
7. Ding, Z., Wang, Z., Zhang, Y., Cao, Y., Liu, Y., Shen, X., Tian, Y. and Dai, J. Trade-offs between machine learning and deep learning for mental illness detection on social media. *Scientific Reports*, 15(1), p.14497 (2025)
8. Garg, M. Mental health analysis in social media posts: A survey. *Archives of Computational Methods in Engineering*, 30(3), pp.1819–1842. doi:10.1007/s11831-022-09863-z (2023)
9. Bhuiyan, M.I., Kamarudin, N.S. and Ismail, N.H. Enhanced suicidal ideation detection from social media using a CNN-BiLSTM hybrid model. *arXiv preprint arXiv:2501.11094* (2025)
10. Shah, S.M., Gillani, S.A., Baig, M.S.A., Saleem, M.A. and Siddiqui, M.H. Advancing depression detection on social media platforms through fine-tuned large language models. *Online Social Networks and Media*, 46, p.100311 (2025)
11. Hasan, K. and Saquer, J. A comparative analysis of transformer and LSTM models for detecting suicidal ideation on Reddit. In *2024 International Conference on Machine Learning and Applications (ICMLA)*, pp.1343–1349 (2024)
12. Zhao, Y. and Kerz, E. Exploring hybrid and ensemble models for multiclass prediction of mental health status on social media. *arXiv preprint arXiv:2212.09839* (2022)
13. Renjith, S., Abraham, A., Jyothi, S.B., Chandran, L. and Thomson, J. An ensemble deep learning technique for detecting suicidal ideation from posts in social media platforms. *Journal of King Saud University – Computer and Information Sciences*, 34(10), pp.9564–9575 (2022)
14. Ezerceli, Ö. and Dehkharghani, R. Mental disorder and suicidal ideation detection from social media using deep neural networks. *Journal of Computational Social Science*, 7, pp.2277–2307 (2024)
15. Tavchioski, I., Robnik-Šikonja, M. and Pollak, S. Detection of depression on social networks using transformers and ensembles. *arXiv preprint arXiv:2305.05325* (2023)
16. Pokrywka, J., Kaczmarek, J.I. and Gorzelańczyk, E.J. Evaluating transformer models for suicide risk detection on social media. *2024 IEEE International Conference on Big Data (BigData)*, pp.8566–8573 (2024)
17. Abd Rahman, R., Omar, K., Noah, S.A.M. and Danuri, M.M. A survey on mental health detection in online social networks. *International Journal of Advanced Science, Engineering and Information Technology*, 8(4-2), p.1431(2018)
18. Yazdavar, A.H. et al. Multimodal mental health analysis in social media. *PLoS One*, 15(4), e0226248. (2020)
19. Ansar, N., Qamar, A. and Ansar, Q. Social media's role in mental health awareness: A review study. *Annual Methodological Archive Research Review*, 3(5), pp.346–363. (2025)
20. Febrian, D. and Awangga, R.M. Machine learning models for predicting mental health indicators using digital physical activity data: A systematic literature review. *Journal Informatics, Education and Management (JIEM)*, 7(2), pp.258–275. (2025)

21. Espino Carrasco, D.K. et al. Sustainability of AI-assisted mental health intervention: A review of literature (2020–2025). *International Journal of Environmental Research and Public Health*, 22(9), p.1382. (2025)
22. Li, L., Novillo-Ortiz, D., Azzopardi-Muscat, N. and Kostkova, P. Digital data sources and their impact on people's health. *Frontiers in Public Health*, 9, 645260. (2021)
23. Singh, J., Wazid, M., Singh, D.P. and Pundir, S. An embedded LSTM-based scheme for depression detection and analysis. *Procedia Computer Science*, 215, pp.166–175. (2022)
24. Ghosal, S. and Jain, A. Depression and suicide risk detection on social media using fastText and XGBoost. *Procedia Computer Science*, 218, pp.1631–1639. (2023)
25. De Choudhury, M., Gamon, M., Counts, S. and Horvitz, E. Predicting depression via social media. *Proceedings of AAAI ICWSM*, pp.128–137. (2013)
26. Kaggle Dataset: Social Media Usage and Emotional Well-being. Available: https://www.kaggle.com/datasets/emirhanai/social-media-usage-and-emotional-well-being