



# Machine Learning Framework for Early Detection of Mental Health Conditions from Textual Data

Basheer Riskhan<sup>1</sup>, Abdullah Al Hadi<sup>1\*</sup>, S M Asiful Islam Saky<sup>1</sup>, Md Saiful Arefin<sup>1</sup>  
and Khalid Hussain<sup>1</sup>

<sup>1</sup>School of Computing and Informatics, Albukhary International University, Alor Setar, Kedah, Malaysia

\*Corresponding author's email: [abdullah.hadi@student.aiu.edu.my](mailto:abdullah.hadi@student.aiu.edu.my)

**Abstract.** Mental health disorders significantly affect global populations, placing heavy burdens on healthcare systems worldwide. Traditional diagnostic methods, mainly clinical assessments and self-reports, lack real-time monitoring, are prone to biases, and often result in delayed interventions. Recent advancements in machine learning (ML) offer promising opportunities to enhance mental health detection through behavioural and physiological data analysis. This study evaluates four widely used machine learning algorithms—Support Vector Machines (SVM), Logistic Regression, Naïve Bayes, and Random Forests—in identifying early indicators of mental health conditions from textual data. A dataset of 27,978 textual records from the “Analysis and Modelling on Mental Health Corpus” was analysed. Data preprocessing involved normalization, stop word removal, lemmatization, and TF-IDF vectorization to prepare robust features for model training. Model performance was assessed using accuracy, precision, recall, and F1-score metrics. Results showed that SVM and Logistic Regression outperformed other models, achieving accuracy rates of 92% and 91%. These findings demonstrate the potential of ML-based frameworks to support earlier and more accurate mental health interventions. Integrating such techniques into clinical practice can improve diagnostic accuracy, reduce healthcare workload, and enhance patient outcomes.

**Keyword:** Mental Health Detection; Textual Data Analysis; Machine Learning Algorithms; Early Intervention.

## 1. Introduction

The field of mental health presents a significant global challenge, impacting millions of individuals worldwide and constituting a substantial portion of the global disease burden [1]. Mental health issues such as depression, anxiety, schizophrenia, and bipolar disorder not only affect an individual's well-being but also strain healthcare systems due to their chronic and recurring nature. Timely detection and effective management of these conditions are critical for improving treatment outcomes and reducing societal costs associated with long-term care.

The gold standard for evaluating mental health is a clinical assessment conducted at regular intervals. However, self-reports are known to be biased and cannot reflect changes in symptoms over short time intervals, making them intrusive. To overcome these hurdles, there has been increased engagement with



real-time digital monitoring in mental health (RTDMH), particularly through machine learning (ML). ML algorithms can process large volumes of data—from social media posts to smartphone and wearable sensor streams—to detect subtle behavioral and physiological changes characteristic of mental health conditions with high accuracy and within brief periods [2,3].

Emerging ML applications offer improved diagnostic methods, early intervention programs, customized treatments, and better prognoses for patients [3]. These technologies can scale far more economically than conventional diagnostic strategies, potentially transforming mental health care delivery. Despite these advances, mental health disorders remain stigmatized, and conventional diagnostic procedures are often time-consuming and sometimes inaccurate, delaying proper treatment. This challenge is compounded by social rejection experienced by those diagnosed, which may deter them from seeking help [1].

The key objectives of this study are threefold. First, to evaluate the effectiveness of widely used machine learning algorithms, including Support Vector Machines (SVM), Naïve Bayes classifiers, Random Forest, and Logistic Regression, in identifying mental health conditions from textual data. Second, to examine and identify the most influential features contributing to classification performance, with a particular focus on data from high school leavers, college entrants, and early-career professionals, to understand feature impact across these demographic groups. Third, to design a machine learning-based diagnostic framework for the early detection of mental health conditions, supporting timely clinical intervention and enhancing the overall quality of life for affected individuals through continuous monitoring and predictive analysis [4]. The paper is organized as follows- **section 2** provides a comprehensive review of the relevant literature on the application of machine learning techniques in mental health detection, highlighting key developments and research gaps. **Section 3** describes the research methodology, including details on the dataset, preprocessing techniques, feature extraction methods, model development, and evaluation procedures. **Section 4** presents the experimental results and discusses their implications in comparison with existing approaches. Finally, **section 5** summarizes the main findings of the study, outlines its limitations, and proposes directions for future research.

## 2. Literature Review

Various machine learning (ML) techniques have been applied to mental health detection, each offering unique advantages. Support Vector Machines (SVMs) are effective for high-dimensional, binary classification tasks such as depression and anxiety detection. Liu et al. used SVMs to achieve high accuracy on physiological and text data [5]. Decision Trees and their ensemble variant, Random Forests, provide interpretability while reducing overfitting [6]. Deep learning architectures (CNNs, LSTMs) excel at capturing hierarchical patterns in EEG signals and social media text, with Shatte et al. demonstrating strong performance in mental health classification [3,7]. These indicate that classical and deep learning models play complementary roles: classical methods offer explainable and computationally efficient solutions, while deep architectures capture complex latent relationships within multimodal data. Together, they contribute to the ongoing transition from rule-based diagnostics to data-driven predictive modelling in mental health research.

Ensemble methods such as AdaBoost and Gradient Boosting combine multiple weak learners into a robust classifier; Katarya and Maan reported high accuracy using these approaches [8]. Structured survey instruments (e.g., DASS-21, PHQ-9) supply tabular inputs for ML pipelines [9], whereas unstructured social media posts can be mined via NLP for sentiment signals indicative of depression and anxiety [10]. Real-time physiological monitoring through smart clothing and wearable sensors (EEG, ECG, HRV) has also been integrated with ML frameworks [11]. By combining self-reported data, behavioural indicators, and physiological measures, researchers can develop systems capable of continuous, passive, and objective monitoring—marking a major shift toward proactive mental health management.



Despite these advances, challenges remain. Data representativeness and quality issues can introduce bias, exacerbating health disparities [12]. Highly accurate deep models often lack interpretability, hindering clinical adoption [13]. The use of sensitive personal data raises ethical and privacy concerns, necessitating robust protections and informed consent [14]. Recent research highlights a growing emphasis on explainable AI (XAI) and multimodal fusion approaches for mental health diagnostics [15]. By combining text, voice, physiological, and behavioural signals, these models aim to capture a more holistic view of an individual's mental state. Furthermore, federated learning has emerged as a promising solution to preserve privacy while enabling collaborative training across multiple institutions without centralizing sensitive data [16]. Some recent studies demonstrate that hybrid frameworks integrating deep learning with XAI not only improve classification accuracy but also enhance interpretability, fostering clinical trust and adoption [17, 18].

In parallel with advances in explainability and federated approaches, recent developments in **transformer-based architectures** have significantly reshaped the landscape of mental health detection from textual data. Chen and Guestrin's **XGBoost** algorithm [19] have been widely adopted for integrating structured and unstructured inputs, demonstrating strong predictive stability and efficiency. Orabi et al. [20] further extended deep learning applications to social media data, showing that depression detection from Twitter posts benefits from richer contextual modelling. Subsequently, the introduction of **BERT** [21] and related transformer models has revolutionized natural language understanding by enabling context-dependent feature extraction, while Matero et al. [22] applied BERT-based embeddings to identify suicidal ideation in online mental health forums with superior precision. Together, these studies illustrate a clear evolution toward **context-aware, semantically interpretable, and scalable** AI systems that advance both the accuracy and inclusivity of modern mental health analytics.

### 3. Research Method

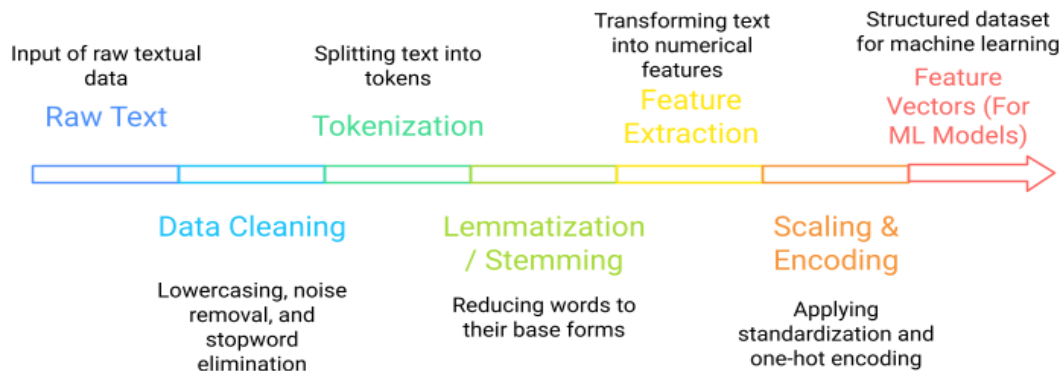
#### 3.1. Data Collection

The **Analysis and Modelling on Mental Health Corpus (AMMHC)** dataset used in this study was obtained from **Kaggle**, where it was originally curated by contributors within the mental health research community. The corpus comprises **27,978 anonymized textual entries** collected from online mental health forums and self-reported user posts discussing topics such as emotional well-being, stress, anxiety, and related psychological conditions. Each record in the dataset contains two primary attributes: a **text** field and a **label** field. The text field includes raw, user-generated content of varying lengths, while the label field provides a binary annotation—**0** indicating the absence and **1** indicating the presence of mental health concerns.

The dataset was selected based on key criteria, including **data quality, source credibility, annotation consistency, and representativeness** of diverse mental health expressions. Its suitability for this research lies in its rich linguistic and emotional cues, which support the development of text-based models for early detection of mental health issues. Additionally, the AMMHC dataset has been referenced in previous studies involving **psychological text mining** and **sentiment analysis**, reinforcing its reliability and validity for machine learning applications in mental health prediction.

#### 3.2. Data Pre-processing

Data pre-processing is a crucial stage for cleaning and standardizing data before feeding it into machine learning models. This stage includes several steps as outlined below in the figure 1.



**Figure 1.** Workflow of Text Cleaning, Transformation, and Feature Engineering

### 3.2.1. Data Cleaning

Data preprocessing ensures the quality and consistency of textual data for mental health detection. In this study, preprocessing was implemented using Python on Jupyter Notebook, leveraging libraries such as Pandas, NumPy, NLTK, Regular Expressions (re), and Scikit-learn. The text data was normalized by converting all text to lowercase. Special characters, numbers, punctuation marks, and HTML tags were removed using Regular Expressions to eliminate noise. Stopwords were filtered out using NLTK's stopword corpus, reducing redundancy and improving feature extraction efficiency.

To standardize word representations, lemmatization was applied using NLTK's WordNet Lemmatizer, ensuring that words were reduced to their base form. Missing values were handled using Pandas' *dropna* function, while Scikit-learn's *SimpleImputer* was used for imputation where necessary. For feature extraction, Term Frequency–Inverse Document Frequency (TF–IDF) vectorization was applied. TF–IDF assigns importance scores to words based on their occurrence in a document relative to their presence in the entire dataset. The selection of these steps was guided by the goal of improving data quality and model interpretability. Lemmatization and stopword removal reduce vocabulary sparsity and promote semantic clarity, while handling missing values prevents inconsistencies that could bias model predictions. These operations ensure that the dataset remains both representative and reliable for downstream analysis.

### 3.2.2. Data Encoding

Encoding was carried out to transform text into formats suitable for ML models. The process began with tokenization, splitting text into words or tokens. Lemmatization and stemming were applied to reduce redundancy, treating words such as *running* and *ran* as the same base form.

Encoding techniques included one-hot encoding and word embeddings such as Word2Vec and GloVe. One-hot encoding creates sparse binary vectors, whereas word embeddings produce dense vectors that capture semantic relationships, thereby enhancing model generalization and interpretability. This combination of encoding techniques allows the model to capture both surface-level linguistic structure and deeper contextual meaning. As a result, the encoded data provides a balanced foundation for machine learning algorithms to recognize emotional, psychological, and linguistic cues within the text.



### 3.2.3. *Scaling and Fitting*

Scaling ensures that all features contribute equally to the learning process. For algorithms such as SVM and neural networks, scaling input data is crucial. Standardization (zero mean, unit variance) was applied to bring features to a comparable scale, improving gradient-based optimization and consistency across training and testing phases. Scaling also improves the efficiency of gradient-based optimization processes in neural networks and prevents features with larger numerical ranges from dominating the learning phase. Proper scaling contributes to faster convergence, better model stability, and improved generalization during evaluation.

### 3.2.4. *Tuning*

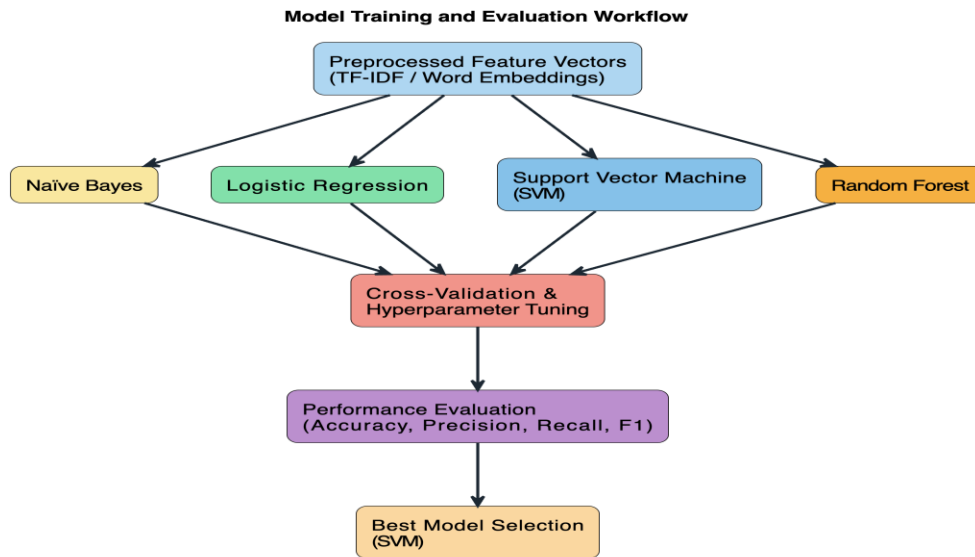
Hyperparameter tuning optimizes model performance. Parameters were tuned using Grid Search and Random Search. Grid Search systematically explores combinations of hyperparameters, while Random Search samples from the parameter space more efficiently. Both methods aimed to select optimal hyperparameters, improving model generalization and reducing risks of overfitting or underfitting. This tuning process ensures that each model operates under optimal conditions, minimizing the risks of overfitting or underfitting. The approach improves the model's robustness, enhances prediction accuracy, and ensures consistent performance across various data splits. By applying a structured tuning procedure, the study achieves a balance between computational efficiency and model reliability.

### 3.2.5. *Feature Extraction*

Feature extraction converts pre-processed text into numerical feature vectors. Common methods include TF-IDF and word embeddings (Word2Vec, GloVe). TF-IDF captures word importance across documents, while embeddings encode semantic meaning and contextual relationships. These methods ensured that only informative features were retained, thereby improving classification accuracy. The combination of statistical and semantic feature extraction techniques provided a balanced representation of both frequency-based and contextual information. This hybrid approach enabled the models to better understand subtle linguistic variations and emotional cues present in mental health-related text.

## 3.3. *Modelling*

Extracted features were used to train multiple machine learning algorithms for detecting mental health concerns. The figure presents the workflow for model training and evaluation in this study. Pre-processed features were used to train four classifiers—Naïve Bayes, Logistic Regression, SVM, and Random Forest—followed by cross-validation and hyperparameter tuning. Model performance was then evaluated using accuracy, precision, recall, and F1-score, as supported by the confusion matrix results. Among the tested models, SVM emerged as the best choice, achieving the most balanced accuracy and recall values. The whole workflow is demonstrated in the figure 2, which provides an overall overview of the system model training and evaluation steps.



**Figure 2.** Model Training and Evaluation Workflow

### 3.3.1. Random Forest

Random Forest builds multiple decision trees using bootstrap samples, aggregating results through majority voting. This reduces variance and improves generalization. Feature importance scores further enhance interpretability, identifying key predictors of mental health indicators.

### 3.3.2. Support Vector Machine (SVM)

SVMs are widely used for binary classification tasks such as mental health detection. The algorithm identifies the optimal hyperplane that maximizes the margin between classes. Kernel functions (RBF, polynomial, sigmoid) extend SVMs to nonlinear problems. Their robustness in high-dimensional feature spaces makes them suitable for text-based classification tasks.

### 3.3.3. Logistic Regression

Logistic Regression models the probability of mental health concerns using a logistic function. Its strengths include simplicity, interpretability, and efficiency on large datasets. Model coefficients reveal relationships between text features and mental health indicators, aiding clinical insights and timely interventions.

### 3.3.4. Naïve Bayes

Naïve Bayes is a probabilistic classifier well-suited for text classification. It assumes independence among features, enabling fast computation and scalability to large datasets. Its efficiency makes it valuable for screening tasks, though accuracy depends on feature representation.

## 3.4 Evaluation

The evaluation of the proposed machine learning framework was carried out using a comprehensive set of performance metrics, namely accuracy, precision, recall, and F1-score. These measures were selected to provide a holistic view of model performance beyond simple accuracy, ensuring that both the

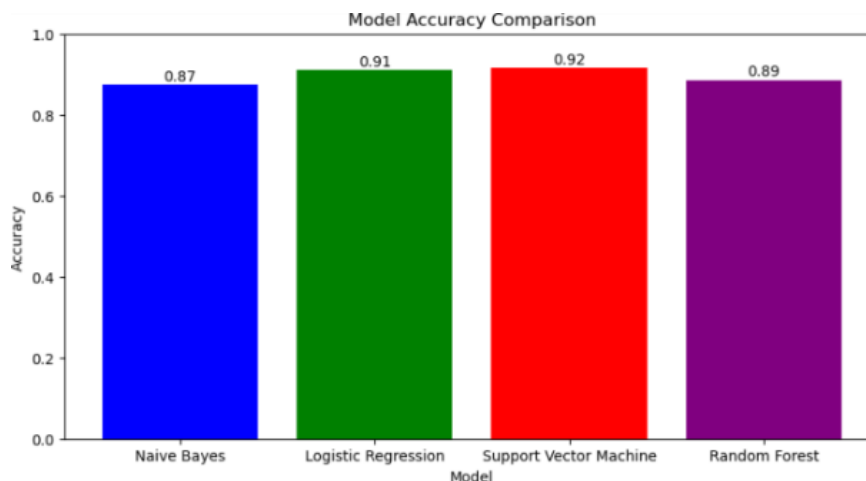


detection of true positives and the minimization of false predictions were adequately assessed. Accuracy served as a general indicator of overall correctness, precision quantified the reliability of positive classifications, and recall reflected the model's ability to identify all relevant mental health cases. The F1-score, as the harmonic mean of precision and recall, provided a balanced perspective, particularly in addressing potential class imbalance within the dataset. To strengthen the reliability of the evaluation, cross-validation techniques were employed, thereby reducing the risk of overfitting and ensuring that the reported outcomes remained stable across different data partitions.

In addition to the evaluation metrics, hyperparameter tuning played a critical role in optimizing model performance. Grid search and random search strategies were applied to systematically explore parameter combinations and identify the most effective configurations for each classifier. This approach ensured that the models were not only compared on equal footing but also optimized for their respective strengths. Furthermore, the evaluation framework emphasized transparency and reproducibility by adopting a standardized process for preprocessing, training, validation, and testing. By integrating multiple evaluation metrics, cross-validation, and systematic parameter optimization, the assessment approach ensured a robust and unbiased comparison of the selected machine learning models for mental health detection.

#### 4. Result and Discussion

The study on mental health detection using ML techniques provided promising and informative outcomes. The experimental evaluation of four machine learning models, namely Naïve Bayes, Logistic Regression, SVM, and Random Forest, yielded strong results for mental health classification. Each model was assessed based on accuracy, precision, recall, and F1-score, providing valuable insights into their classification performance.



**Figure 3.** Model Accuracy Across Naive Bayes, Logistic Regression, SVM, and Random Forest

As stated in the figure 3, the Naïve Bayes model achieved an accuracy of 87%, demonstrating strong recall for mental health cases (0.95) but lower precision (0.83), indicating a tendency to over-classify. It correctly identified 2,794 mental health cases, misclassifying 553 non-mental health cases. The model's probabilistic nature allowed it to be highly sensitive in detecting mental health cases, making it a useful choice where minimizing false negatives is crucial. However, its assumption of feature independence limited its precision, leading to false positives.

The Logistic Regression model demonstrated an improved accuracy of 91.12%, achieving a precision of 0.91 for label 0 (non-mental health cases) and 0.92 for label 1 (mental health cases). The recall values were 0.92 and 0.90, respectively, indicating well-balanced classification performance. The

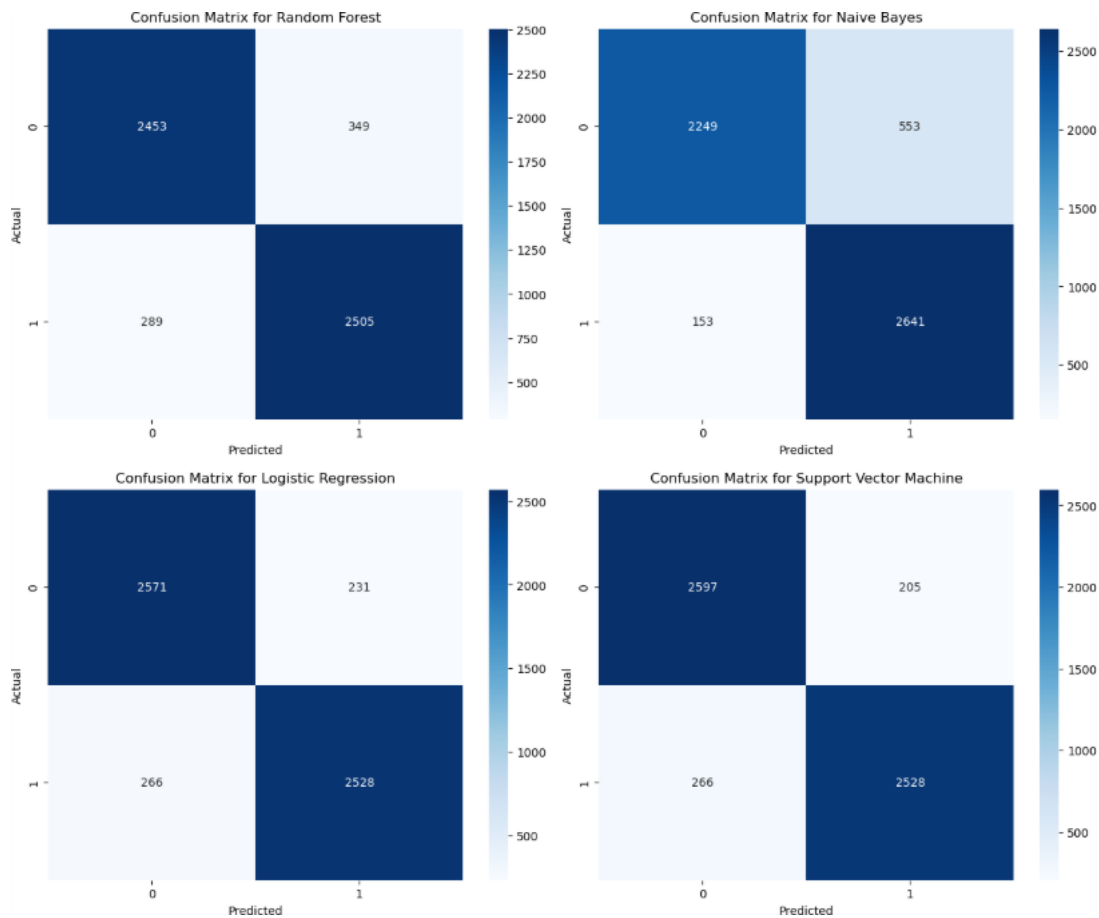


model correctly classified 2,571 non-mental health cases and 2,528 mental health cases, with misclassifications of 231 and 266 cases, respectively. Logistic Regression effectively distinguished between the two classes, making it a robust choice for general-purpose mental health classification.

The SVM model emerged as the best-performing classifier, achieving the highest accuracy of 91.58%. This strong performance can be attributed to a combination of effective hyperparameter tuning, particularly with the RBF kernel, which captures complex decision boundaries, and the use of TF-IDF vectorization, which transformed the textual data into high-dimensional numerical features, enabling SVM to extract meaningful linguistic patterns. Compared to other models, SVM leveraged high-dimensional text data more effectively, making it particularly suitable for natural language classification. SVM's performance was further enhanced through hyperparameter tuning, particularly adjusting the regularization parameter (C) to balance margin maximization and classification error, ensuring better generalization. The Radial Basis Function (RBF) kernel was used to capture complex decision boundaries, allowing the model to differentiate between subtle variations in mental health-related text. Experimental results show that SVM achieved a precision of 0.91 for label 0 and 0.92 for label 1, with recall values of 0.93 and 0.90, respectively. The higher recall for label 0 indicates that SVM was slightly better at classifying non-mental health cases while maintaining strong performance in detecting mental health cases. It correctly classified 2,597 non-mental health cases and 2,528 mental health cases, with 205 and 266 misclassifications, respectively. The combination of optimized feature representation, effective high-dimensional data handling, and fine-tuned hyperparameters made SVM the most suitable model for this study.

The Random Forest model achieved an accuracy of 88.59%, performing reliably across both classes. It attained a precision of 0.89 for label 0 and 0.88 for label 1, with recall values of 0.88 and 0.90, respectively. While Random Forest performed well in classifying non-mental health cases, its lower precision for label 1 suggests a tendency to misclassify some mental health-related text, making it slightly less effective than SVM and Logistic Regression. The model misclassified 349 non-mental health cases and 289 mental health cases, indicating that while its ensemble learning approach balanced classification, it was not as effective as SVM in generalizing across both classes. Random Forest was used as the final algorithm, and it leveled the results well, but its performance was slightly lower than SVM and Logistic Regression.

Analyzing figure 4, although all four models demonstrated strong classification performance, SVM and Logistic Regression outperformed the others, achieving the highest accuracy and well-balanced precision–recall values. Naïve Bayes exhibited high recall for mental health cases but was prone to over-classification, leading to more false positives. Random Forest provided a reliable alternative, but its lower precision for mental health cases made it slightly less suitable for real-world classification tasks. The results confirm that SVM, with its ability to optimize high-dimensional text features and its robust hyperparameter tuning, is the most effective model for mental health classification.



**Figure 4.** Confusion Matrix Comparison of Random Forest, Naive Bayes, Logistic Regression, and SVM

## 5. Conclusion

This study evaluated the effectiveness of machine learning models in detecting mental health conditions, focusing on Support Vector Machine (SVM), Logistic Regression, Random Forest, and Naïve Bayes. The experimental results demonstrated that SVM (92% accuracy) and Logistic Regression (91% accuracy) were the most effective models, exhibiting well-balanced precision and recall values, making them suitable for mental health classification tasks. While Naïve Bayes (87%) showed high sensitivity (recall of 0.95 for label 1), it produced a higher false positive rate. Random Forest (88%) provided competitive performance, but its slightly lower precision made it less suitable compared to SVM and Logistic Regression.

Although machine learning models present significant potential in mental health diagnostics, several constraints must be considered. The trade-off between accuracy and computational cost is a key concern, as models like SVM require substantial processing power and memory, making them less suitable for real-time applications in resource-limited environments. Additionally, bias in training data can impact model fairness, leading to misclassifications, particularly for underrepresented groups. Ethical concerns related to data privacy, consent, and regulatory compliance must also be addressed to ensure responsible AI deployment in healthcare.

To further improve mental health classification, future research should explore hybrid machine learning models, leveraging the strengths of multiple classifiers to enhance predictive performance. A combination of SVM with probabilistic models like Naïve Bayes or ensemble methods such as boosting



and stacking could yield a more robust system, optimizing both accuracy and sensitivity. Additionally, integrating electronic healthcare records (EHRs) with machine learning could provide a more comprehensive assessment by combining structured clinical data with unstructured social media text features.

Future work should also focus on explainable AI (XAI) techniques, ensuring model transparency and interpretability to gain trust from healthcare professionals and patients. The continued advancement of AI-driven mental health screening tools holds great potential for enhancing early diagnosis, reducing misclassification rates, and improving patient outcomes. However, to fully realize these benefits, further research is needed to address computational challenges, data bias, ethical considerations, and model interpretability, ensuring that AI-powered mental health assessment tools are accurate, ethical, and widely accessible. Efforts should also emphasize the development of computationally efficient algorithms for resource-constrained environments, improved dataset collection practices to reduce bias, and robust data privacy protocols.

## References

- [1] World Health Organization, *Depression and Other Common Mental Disorders: Global Health Estimates*. Geneva, Switzerland: World Health Organization, 2017.
- [2] E. Garcia-Ceja, M. Riegler, T. Nordgreen, P. Jakobsen, K. J. Oedegaard, and J. Tørresen, "Mental health monitoring with multimodal sensing and machine learning: A survey," *Pervasive and Mobile Computing*, vol. 51, 2018, doi: 10.1016/j.pmcj.2018.09.003.
- [3] A. B. R. Shatte, D. M. Hutchinson, and S. J. Teague, "Machine learning in mental health: A scoping review of methods and applications," *Psychological Medicine*, vol. 49, no. 9, pp. 1426–1448, 2019, doi: 10.1017/S0033291719000151.
- [4] Allam, H., Davison, C., Kalota, F., Lazaros, E., & Hua, D. (2025). AI-Driven Mental Health Surveillance: Identifying Suicidal Ideation Through Machine Learning Techniques. *Big Data and Cognitive Computing*, 9(1), 16
- [5] Y. Liu, J. Hankey, B. Cao, and P. Chokka, "Screening for major depressive disorder in a tertiary mental health center using EarlyDetect: A machine learning-based pilot study," *Journal of Affective Disorders Reports*, vol. 3, 2021, doi: 10.1016/j.jadr.2020.100062.
- [6] K. Vaishnavi, U. N. Kamath, B. A. Rao, and N. V. S. Reddy, "Predicting mental health illness using machine learning algorithms," *Journal of Physics: Conference Series*, 2022, doi: 10.1088/1742-6596/2161/1/012021.
- [7] K. Hussain, S. A. Khan, M. Muzaffar, Z. R. Khan, and S. U. Farooq, "Performance evaluation of CNN architectures for COVID-19 detection from X-ray images," *Computer Methods in Biomechanics and Biomedical Engineering: Imaging & Visualization*, vol. 11, no. 1, pp. 33–41, 2023, doi: 10.1080/21681163.2022.2052750.
- [8] R. Katarya and S. Maan, "Predicting mental health disorders using machine learning for employees in technical and non-technical companies," in *Proc. IEEE Int. Conf. Advances and Developments in Electrical and Electronics Engineering (ICADEE)*, 2020, doi: 10.1109/ICADEE51157.2020.9368923.
- [9] D. J. Joshi, M. Makhija, Y. Nabar, N. Nehete, and M. S. Patwardhan, "Mental health analysis using deep learning for feature extraction," in *Proc. ACM Int. Conf. Proceeding Series*, 2018, doi: 10.1145/3152494.3167990.
- [10] R. A. Rahman, K. Omar, S. A. M. Noah, M. S. N. M. Danuri, and M. A. Al-Garadi, "Application of machine learning methods in mental health detection: A systematic review," *IEEE Access*, vol. 8, pp. 183 952–183 970, 2020, doi: 10.1109/ACCESS.2020.3029154.
- [11] E. Sajno, S. Bartolotta, C. Tuena, P. Cipresso, E. Pedroli, and G. Riva, "Machine learning in biosignals processing for mental health: A narrative review," *Frontiers in Psychology*, vol. 13, 2023, doi: 10.3389/fpsyg.2022.1066317.
- [12] N. K. Iyortsuun, S. H. Kim, M. Jhon, H. J. Yang, and S. Pant, "A review of machine learning and deep learning approaches on mental health diagnosis," *Healthcare (Switzerland)*, vol. 11, no. 3, 2023, doi: 10.3390/healthcare11030285.
- [13] S. M. et al., "Mental health prediction models using machine learning in higher education institution," *Turkish Journal of Computer and Mathematics Education (TURCOMAT)*, vol. 12, no. 5, pp. 1442–1452, 2021, doi: 10.17762/turcomat.v12i5.2181.
- [14] S. A. Alanazi et al., "Public's mental health monitoring via sentimental analysis of financial text using machine learning techniques," *International Journal of Environmental Research and Public Health*, vol. 19, no. 15, 2022, doi: 10.3390/ijerph19159695.
- [15] P. Dubey, P. Dubey, and P. N. Bokoro, "Federated learning for privacy-enhanced mental health prediction with multimodal data integration," *Computer Methods in Biomechanics and Biomedical Engineering: Imaging & Visualization*, vol. 13, no. 1, p. 2509672, 2025, doi:10.1080/21681163.2024.2509672.



- [16] A. Kumar, J. Kumari, and J. Pradhan, “Explainable deep learning for mental health detection from English and Arabic social media posts,” *ACM Transactions on Asian and Low-Resource Language Information Processing*, 2023.
- [17] Y. Yousufi, R. Damaševičius, and R. Maskeliūnas, “Multimodal fusion of EEG and audio spectrogram for major depressive disorder recognition using modified DenseNet-121,” *Brain Sciences*, vol. 14, no. 10, Art. 1018, Oct. 2024.
- [18] X. Du, C. Jiang, X. Yuan, et al., “GAME: Generalized deep learning model towards multimodal data integration for early screening of adolescent mental disorders,” *arXiv preprint arXiv:2309.10077*, Sep. 2023.
- [19] T. Chen and C. Guestrin, “XGBoost: A scalable tree boosting system,” in *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD)*, San Francisco, CA, USA, 2016, pp. 785–794.
- [20] A. H. Orabi, P. Buddhitha, B. Hachey, and M. Rouhizadeh, “Deep learning for depression detection of Twitter users,” in *Proceedings of the Fifth Workshop on Computational Linguistics and Clinical Psychology: From Keyboard to Clinic (CLPsych)*, New Orleans, LA, USA, 2018, pp. 88–97.
- [21] J. Devlin, M.-W. Chang, K. Lee, and K. Toutanova, “BERT: Pre-training of deep bidirectional transformers for language understanding,” in *Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies (NAACL-HLT)*, Minneapolis, MN, USA, 2019, pp. 4171–4186.
- [22] M. Matero, A. Idnani, Y. Son, S. Giorgi, H. Vu, M. Zamani, and H. A. Schwartz, “Suicidal ideation detection via contextual embeddings of posts from online mental health forums,” in *Proceedings of the Sixth Workshop on Computational Linguistics and Clinical Psychology (CLPsych)*, Minneapolis, MN, USA, 2019, pp. 167–171.