International Journal of Scientific and Academic Research (IJSAR) ISSN: 2583-0279
Volume 4, Issue 9, December- 2024
https://ijsar.net
CrossRef DOI: 10.54756/IJSAR.2024.26

# Schizophrenia Diagnosis and Prediction with Machine Learning Models

**Faezeh Norouzi[1], Bruna Lino Modes Santos Machado[2], and Shervin Nematzadeh[3]**

[1] Psychiatry and Behavioral Science, Isfahan University of Medical Science, Isfahan, Iran

[2] Lutheran Institute of Higher Education of Itumbiara, Brazil

[3] North Tehran Branch, Islamic Azad University, Tehran, Iran

## ABSTRACT

*Schizophrenia is a multifaceted mental disorder with varying levels of proneness, making accurate classification essential for early intervention and effective management. This study investigates the application of machine learning techniques to classify schizophrenia proneness levels based on behavioral and demographic features, including age, fatigue, slowing, pain, hygiene, and movement. Using a dataset of 1,000 samples categorized into five levels of proneness—Elevated, High, Moderate, Low, and Very High—we evaluated the performance of Logistic Regression, Support Vector Machine (SVM), Gradient Boosting, and Decision Tree classifiers. Among the models, Logistic Regression achieved the highest accuracy of 94.2%, demonstrating its effectiveness in capturing feature relationships and its suitability for datasets with linear or near-linear patterns. SVM is closely followed with an accuracy of 93%, showcasing its robustness in handling high-dimensional data and non-linear relationships. Gradient Boosting achieved an accuracy of 88.1%, indicating its ability to model complex patterns through iterative corrections, while the Decision Tree, with an accuracy of 75.6%, served as a baseline, reflecting the limitations of single-tree models for complex datasets.*

**Key Words:** Artificial Intelligence, Mental Health, Schizophrenia.

## 1. INTRODUCTION

Mental disorders are a critical global health concern, impacting millions of individuals across diverse age groups, socioeconomic statuses, and cultures. These conditions encompass a wide range of disorders, including depression, anxiety, bipolar disorder, and severe illnesses like schizophrenia, each of which profoundly affects mental, emotional, and social well-being. The growing prevalence of mental disorders has placed immense pressure on healthcare systems worldwide, underscoring the need for robust mental health screening programs. Screening plays a vital role in identifying early signs of mental health issues and facilitating timely and targeted interventions that can significantly improve outcomes. Beyond healthcare settings, integrating mental health screening into schools, workplaces, and community programs has emerged as a pivotal strategy for promoting mental well-being. With advancements in research and technology, mental health assessments have shifted from subjective clinical evaluations to systematic, evidence-based methodologies, making mental health care more accessible (Sacco et al., 2024; Hertenstein et al., 2023).

Schizophrenia, one of the most complex and debilitating mental disorders, is characterized by profound disruptions in thought processes, perception, and emotional regulation. This chronic condition manifests through a combination of positive, negative, and cognitive symptoms, each presenting unique challenges for affected individuals (Tandon et al., 2024; McCutcheon et al., 2023). Positive symptoms include hallucinations, delusions, and disorganized thinking, which can distort reality and impair daily functioning. Negative symptoms, such as reduced motivation, social withdrawal, and diminished emotional expression, often result in profound isolation and diminished quality of life. Cognitive impairments, including memory, attention, and executive functioning difficulties, further complicate the ability to navigate everyday tasks. Schizophrenia often has a significant impact not only on the individuals affected but also on their families and caregivers, who frequently shoulder the burden of long-term care. Despite being a well-recognized disorder, the variability in symptom presentation and severity poses substantial challenges for accurate diagnosis and effective treatment, highlighting the need for innovative approaches to understanding and managing the condition. The analysis of mental health data, particularly in the context of schizophrenia, holds transformative potential for advancing clinical practice and research (Ibrahim et al., 2023). Mental health disorders are often underdiagnosed or misdiagnosed due to the subjective nature of symptom evaluation and the complex interplay of biological, psychological, and social factors. Stigma and limited access to mental health resources further exacerbate these challenges, leaving many individuals without the care they need. By systematically analyzing data on symptoms, behaviors, and patient demographics, researchers can uncover critical patterns that inform clinical practice and guide public health initiatives. In the case of schizophrenia, data-driven insights can help identify early warning signs, stratify risk levels, and personalize interventions. This approach not only improves diagnostic precision but also aids in resource allocation and policymaking. Furthermore, longitudinal analyses of schizophrenia datasets can reveal trends in treatment response and disease progression, contributing to the development of evidence-based guidelines that enhance patient outcomes and reduce the societal burden of the disorder (Van Dee et al., 2023).

The integration of artificial intelligence (AI) into healthcare has ushered in a new era of precision and efficiency, transforming how medical conditions are diagnosed, monitored, and treated (Chen et al., 2022; Abbasi et al., 2024; Kufel et al., 2023; Rauniyar et al., 2023; Kharaji et al., 2024; Haug et al., 2023). AI technologies like machine learning, deep learning, and natural language processing excel at processing vast and complex datasets, identifying patterns, and generating predictive insights beyond human capabilities. In medicine, AI has been applied across various domains, including radiology (Minoo et al., 2024; Radak et al., 2023; Afrazeh et al., 2024), pathology (Mahmoudiandehkordi et al., 2024), and genomics (Guo et al., 2023), to improve diagnostic accuracy and accelerate decision-making. AI-driven tools have been instrumental in advancing personalized medicine by tailoring treatment plans based on individual patient profiles. From disease early detection (Orouskhani et al., 2022; NG et al., 2023) through imaging analysis to optimizing surgical outcomes (Malhotra et al., 2023), AI has demonstrated its potential to bridge gaps in clinical expertise and resource availability, ensuring that more patients receive high-quality care. In the realm of mental health, AI offers promising solutions to long-standing challenges associated with the subjective nature of psychiatric evaluations and the multifaceted presentation of mental disorders (Verma et al., 2023; Abplanal et al., 2024; Sharma et al., 2023). Traditional methods of assessing mental health often rely on clinical interviews and self-reported symptoms, which can be prone to bias and variability. AI can overcome these limitations by analyzing diverse data types, including behavioral patterns, speech characteristics, and physiological signals, to identify subtle indicators of mental health conditions. In the context of schizophrenia, AI-powered models can classify symptoms, predict risk levels, and provide valuable insights into disease progression. For example, machine learning algorithms can analyze speech and behavioral data to detect early signs of psychosis, enabling timely interventions that may alter the course of the disorder. Moreover, AI-driven analytics can optimize treatment strategies by identifying patterns of response to various therapies, paving the way for more personalized and effective care. By leveraging AI in mental health research and clinical practice, we can unlock deeper understanding, reduce diagnostic disparities, and improve the lives of individuals affected by conditions like schizophrenia.

## 2. METHODS

### 2.1 Dataset

The dataset utilized in this study comprises 5,000 samples collected to analyze symptoms associated with schizophrenia, focusing on behavioral and speech patterns. The dataset includes individuals of varying demographic and clinical characteristics, providing a comprehensive overview for analysis. The average age of participants is 74.83 years, with a standard deviation of 9.58 years, highlighting a predominant population in the elderly age group. The gender distribution is nearly balanced, with 2,490 samples from men (49.8%) and 2,510 samples from women (50.2%), ensuring gender representation in the analysis (Kaggle).

Each individual in the dataset is categorized into one of five classes based on their proneness to schizophrenia: Elevated Proneness (3,077 samples, 61.5%), High Proneness (953 samples, 19.1%), Moderate Proneness (912 samples, 18.2%), Low Proneness (45 samples, 0.9%), and Very High Proneness (13 samples, 0.3%). The most prevalent category is Elevated Proneness, representing individuals exhibiting noticeable symptoms but not at the highest risk. This is followed by High Proneness with significant symptoms and a heightened likelihood of developing schizophrenia. The Moderate Proneness level represents individuals with balanced risk, exhibiting symptoms of



DOI: 10.54756/IJSAR.2024.26

intermediate severity. Low Proneness is observed in only 45 samples (0.9%) and corresponds to individuals with minimal symptoms and lower risk. The rarest category is Very High Proneness, encompassing just 13 samples (0.3%) of individuals with severe symptoms and the highest risk of schizophrenia. This categorization not only highlights the diversity within the dataset but also provides a robust framework for developing predictive models to analyze and classify schizophrenia proneness. This distribution reveals that most individuals fall into the Elevated or High Proneness categories, while a smaller proportion exhibit Low or Very High Proneness, providing a varied yet skewed dataset for machine learning tasks. The dataset includes several key features capturing both demographic and clinical aspects. These features encompass numerical scores reflecting fatigue, slowing (reduced responsiveness or movement), pain levels, hygiene (self-care), and movement activity. Demographic attributes such as age, gender, and marital status are also included, allowing for a multifaceted analysis. Notably, some missing values are observed in specific columns, such as the "Pain" score, which will require preprocessing for robust analysis. The diversity and richness of the dataset enable an in-depth exploration of schizophrenia symptoms, making it an ideal resource for applying machine learning algorithms. The balanced gender distribution, combined with a wide range of symptom severity and behavioral characteristics, allows for the development of predictive models and insights into schizophrenia proneness, aiding in early detection and targeted interventions. Figure 1 shows the distributions of six variables: Age, Fatigue, Slowing, Pain, Hygiene, and Movement, providing insights into the dataset's characteristics. The Age variable shows a distribution concentrated in the middle range, particularly between 65 and 85 years, consistent with the dataset's focus on an older population. The other variables—Fatigue, Slowing, Pain, Hygiene, and Movement—display relatively uniform distributions across their respective ranges (0 to 1), indicating a balanced representation of symptom levels among individuals. These uniform distributions suggest that the dataset captures diverse symptom severities without any one range dominating, which is advantageous for machine learning tasks, as it minimizes bias and enhances the model's ability to generalize. Overall, the balanced distributions of these variables underscore the dataset's robustness for analyzing patterns and relationships related to schizophrenia symptoms.

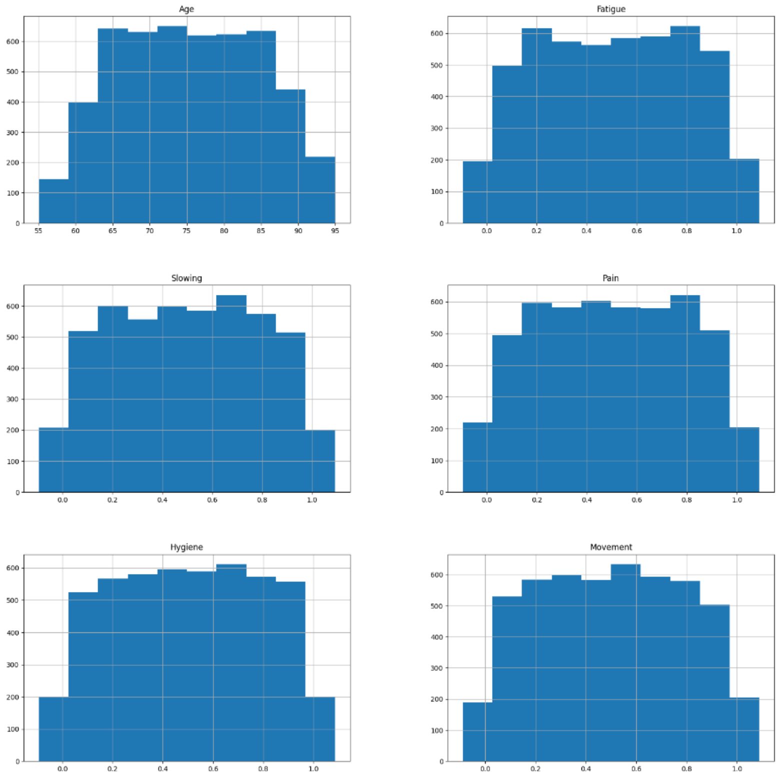


**Figure 1 Histogram illustrations of six variables (Kaggle)**



DOI: 10.54756/IJSAR.2024.26

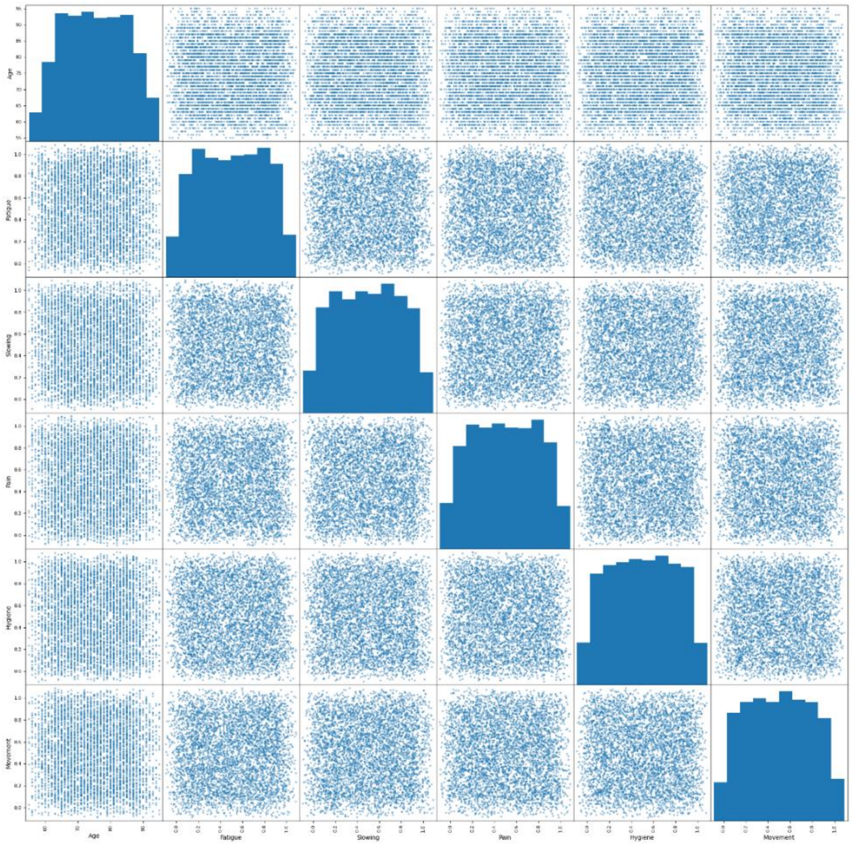


**Figure 2 The scatter plots (Kaggle)**

The scatter matrix in Figure 2 provides a comprehensive visualization of the relationships among the variables **Age**, **Fatigue**, **Slowing**, **Pain**, **Hygiene**, and **Movement**. The diagonal elements of the matrix display the histograms for each variable, reaffirming their individual distributions, with **Age** concentrated between 65 and 85 years and the other variables showing uniform distributions across their ranges. The off-diagonal scatter plots illustrate pairwise relationships between the variables, revealing largely random and unstructured patterns. This indicates weak or no linear correlations between the variables, suggesting that they are largely independent. The lack of strong correlations implies that each variable captures distinct and complementary information about the individuals in the dataset. This independence is advantageous for machine learning applications, as it minimizes feature redundancy and ensures that each feature contributes uniquely to the predictive modeling process. Overall, the scatter matrix highlights the diverse and independent nature of the dataset, making it a robust foundation for analyzing schizophrenia symptoms and developing accurate predictive models.



DOI: 10.54756/IJSAR.2024.26

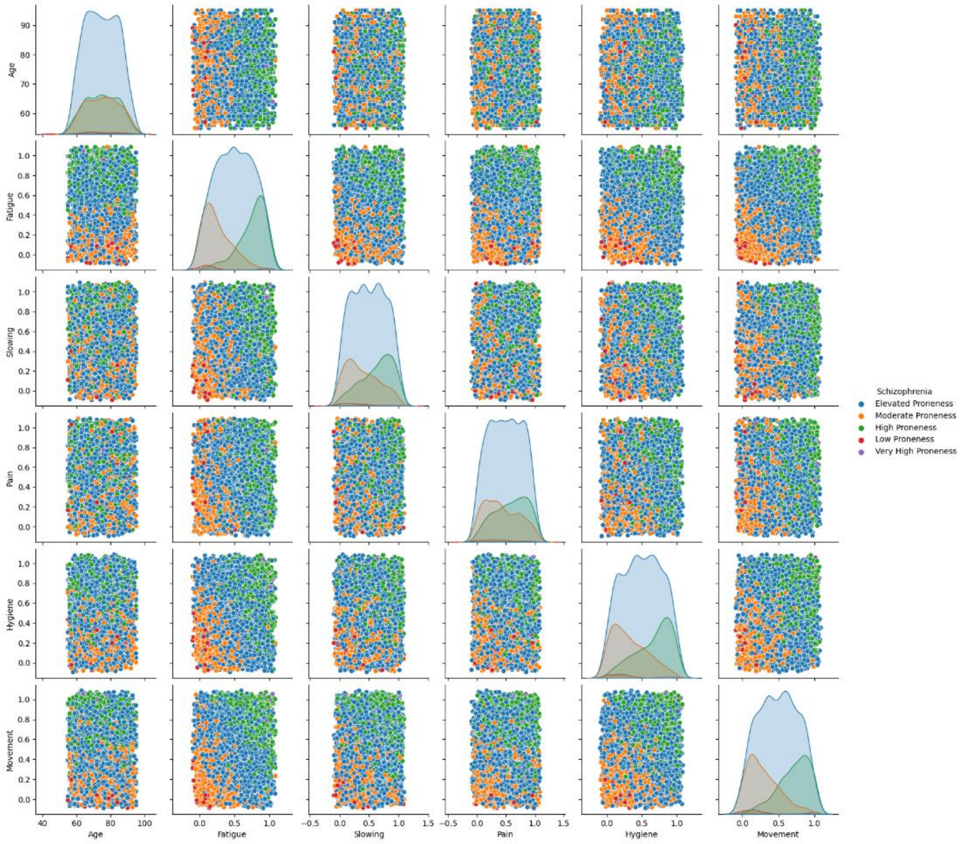


**Figure 3 visualization of the relationships among the variables Age, Fatigue, Slowing, Pain, Hygiene, and Movement, with data points color-coded by the schizophrenia proneness levels (Kaggle)**

The pair plot by Figure 3 provides a comprehensive visualization of the relationships among the variables Age, Fatigue, Slowing, Pain, Hygiene, and Movement, with data points color-coded by schizophrenia proneness levels: Elevated Proneness, Moderate Proneness, High Proneness, Low Proneness, and Very High Proneness. The diagonal elements display kernel density estimates (KDEs) for each variable, illustrating the distribution of values across different proneness levels. While there is considerable overlap between the categories, slight variations can be observed, particularly in features like Fatigue, Pain, and Movement, which show distinct density patterns for certain proneness levels. The scatter plots in the off-diagonal elements highlight pairwise relationships between variables, revealing that most relationships lack strong clustering or separation among the categories. However, some variables show subtle differences in density and spread, suggesting their potential relevance for classification tasks. The significant overlap among the Elevated, Moderate, and High Proneness categories underscores the challenge of distinguishing between these levels, while Low Proneness and Very High Proneness exhibit more distinct patterns. These observations emphasize the need for advanced machine learning techniques capable of capturing subtle, non-linear relationships in the data to accurately classify schizophrenia proneness levels.

### 2.2 Machine Learning Models

Machine Learning (ML) is a branch of artificial intelligence that focuses on developing algorithms capable of learning patterns from data and making predictions or decisions without being explicitly programmed. In recent years, ML has revolutionized numerous fields, including healthcare, by providing powerful tools for analyzing complex datasets. One of its primary benefits is the ability to uncover hidden relationships and insights in data that may not be apparent through traditional statistical methods. Machine learning algorithms excel in tasks such as classification, regression, clustering, and anomaly detection, making them invaluable for identifying trends, predicting outcomes, and optimizing processes. In healthcare, ML has been particularly impactful, enabling early disease detection, patient risk stratification, personalized treatment recommendations, and resource management. Additionally, its capacity to handle diverse data types—ranging from numerical measurements to textual, visual, and audio data—broadens its



DOI: 10.54756/IJSAR.2024.26

applicability across medical domains. By automating repetitive tasks and improving diagnostic precision, ML not only enhances decision-making but also contributes to better patient outcomes and more efficient healthcare systems.

- **Logistic Regression**

Logistic Regression is a foundational machine learning algorithm used for classification tasks, particularly when the target variable is categorical. It works by modeling the relationship between one or more independent variables and a dependent categorical variable. Logistic Regression estimates the probability of a data point belonging to a specific class, making it suitable for binary and multiclass classification problems. It is widely used because of its simplicity, interpretability, and efficiency.

One of the key advantages of Logistic Regression is its ability to provide direct insights into the relationship between features and the target variable. By examining the model's coefficients, practitioners can understand how each feature contributes to the likelihood of an outcome. Logistic Regression performs well when the classes are linearly separable and is often used as a baseline model for classification tasks. However, it may struggle with non-linear patterns in the data, which can limit its effectiveness without additional preprocessing or feature engineering. Logistic Regression is a go-to algorithm in healthcare for disease prediction and risk analysis due to its interpretability and ease of deployment.

- **Support Vector Machine (SVM)**

Support Vector Machine (SVM) is a robust and versatile machine learning algorithm widely used for classification and regression tasks. Its core idea is to find the optimal hyperplane that separates classes in a feature space with the maximum margin, enhancing the model's ability to generalize to unseen data. SVM is particularly effective for datasets with high-dimensional feature spaces, where it identifies patterns that may not be apparent in lower-dimensional representations.

One of the strengths of SVM lies in its flexibility to handle non-linear relationships using kernel functions. These functions transform the original feature space into a higher-dimensional space, enabling the algorithm to draw non-linear decision boundaries. Popular kernels include the radial basis function (RBF), polynomial, and linear kernels, each suitable for different data structures. SVM also incorporates a regularization parameter that balances the trade-off between maximizing the margin and minimizing classification errors, making it resilient to overfitting. While computationally intensive for large datasets, SVM remains a powerful algorithm for tasks like medical diagnostics, where precision and reliability are paramount.

- **Gradient Boosting Classifier**

Gradient Boosting Classifier is an advanced ensemble algorithm that builds a strong predictive model by combining multiple weak learners, typically decision trees. It operates iteratively, where each tree corrects the errors made by the ensemble so far. This sequential approach ensures that the model focuses on difficult-to-classify samples, enhancing overall performance.

The flexibility of Gradient Boosting lies in its ability to handle a wide range of data types and its strong predictive capabilities. By adjusting hyperparameters such as the learning rate, number of estimators, and tree depth, the model can be fine-tuned to balance accuracy and complexity. Gradient Boosting also includes regularization techniques to prevent overfitting, making it suitable for complex datasets with noise. Its adaptability and accuracy have made it a preferred choice for high-stakes applications, such as predicting disease outcomes and identifying risk factors in healthcare.

- **Decision Tree Classifier**

The Decision Tree Classifier is a straightforward yet powerful algorithm that divides data into subsets based on feature values, forming a tree-like structure. Each node in the tree represents a decision based on a feature, while the branches represent the outcomes of those decisions. The algorithm selects the feature that best splits the data at each step, using criteria such as Gini impurity or entropy to evaluate the quality of the splits.



DOI: 10.54756/IJSAR.2024.26

Decision trees are intuitive and easy to interpret, making them a popular choice in applications where explainability is essential. They are capable of handling both categorical and numerical data and can model complex decision boundaries. However, decision trees are prone to overfitting, especially when grown deep, as they can memorize training data. Techniques such as pruning, limiting tree depth, or combining multiple trees in ensemble methods like Random Forest or Gradient Boosting help mitigate this limitation. Despite their simplicity, decision trees remain an essential tool in machine learning, often serving as the foundation for more complex models.

- **XGBClassifier**

XGBClassifier, part of the XGBoost library, is an optimized implementation of the gradient boosting algorithm. It combines the strengths of multiple decision trees to create a powerful predictive model, emphasizing speed, scalability, and accuracy. XGBClassifier introduces enhancements such as parallel processing, efficient handling of missing data, and regularization to improve performance and prevent overfitting.

A standout feature of XGBClassifier is its ability to fine-tune performance through a wide array of hyperparameters, such as the learning rate, maximum tree depth, and subsampling ratios. These options allow the algorithm to be tailored for diverse datasets and objectives. The inclusion of feature importance metrics provides insights into the influence of individual features on predictions, making the model interpretable despite its complexity. XGBClassifier is particularly effective in handling imbalanced datasets and capturing intricate feature interactions, making it a top choice for medical applications where precision and robustness are critical. Its success in competitions and real-world scenarios has cemented its reputation as one of the most powerful algorithms in the machine learning toolkit.

## 3. RESULTS

Four machine learning models—Logistic Regression, Support Vector Machine (SVM), Gradient Boosting, and Decision Tree—were evaluated for classifying schizophrenia proneness levels. The results indicate that all models achieved varying degrees of accuracy, with Logistic Regression and SVM demonstrating superior performance compared to Gradient Boosting and Decision Tree. The **Logistic Regression** model achieved the highest accuracy of **94.2%**, demonstrating its effectiveness in capturing the relationships between features and accurately predicting schizophrenia proneness levels. This result highlights the suitability of Logistic Regression for datasets where linear or near-linear relationships play a significant role in classification. The **Support Vector Machine (SVM)** model followed closely with an accuracy of **93%**, showcasing its robustness and ability to handle complex, high-dimensional data. The high performance of SVM reflects its strength in separating overlapping classes through the use of kernel functions, which are particularly effective in capturing non-linear patterns in the data. The **Gradient Boosting** model achieved an accuracy of **88.1%**, indicating its ability to model complex patterns by iteratively correcting errors from previous iterations. While its performance is slightly lower than that of Logistic Regression and SVM, it still provides valuable insights and may benefit from further hyperparameter tuning to improve accuracy. The **Decision Tree** model attained an accuracy of **75.6%**, making it the least accurate among the four models. This result is expected, as single decision trees are prone to overfitting and may struggle with complex datasets. While interpretable, Decision Trees often perform better when used in ensemble methods, such as Random Forest or Gradient Boosting. Overall, the results highlight the effectiveness of Logistic Regression and SVM as top-performing models for this dataset, offering high accuracy and reliable predictions. Gradient Boosting provides a solid alternative with room for optimization. At the same time, the Decision Tree serves as a baseline, demonstrating the need for more advanced techniques to handle the complexity of schizophrenia symptom data.

A voting classifier is an ensemble machine-learning method that combines the predictions of multiple models to improve overall accuracy and robustness. The voting classifier's performance metrics and confusion matrix provide a comprehensive evaluation of the model's classification ability across the five schizophrenia proneness levels: Elevated Proneness, High Proneness, Low Proneness, Moderate Proneness, and Very High Proneness. For Elevated Proneness, the model achieves a precision of 0.91, a recall of 0.96, and an F1-score of 0.93, along with 592 correctly classified samples in the confusion matrix, demonstrating excellent performance in identifying and accurately predicting the most represented category in the dataset. Similarly, the model performs well for High Proneness, with a precision of 0.88 and a recall of 0.83 while correctly classifying 133 samples. However, some misclassifications are evident, with



DOI: 10.54756/IJSAR.2024.26

28 High Proneness samples being classified as Elevated Proneness, highlighting the challenge of distinguishing between neighboring proneness levels with overlapping feature distributions. For Moderate Proneness, the model achieves a precision of 0.89 and a recall of 0.82, with 157 correctly classified samples, further showcasing its ability to handle well-represented categories. In contrast, the model struggles with the underrepresented categories. For Low Proneness, it achieves a perfect precision of 99.7, indicating that almost all predictions for this category were correct, but the recall is only 0.38, resulting in a low F1-score and just 6 correctly classified samples, with 8 misclassified as Moderate Proneness. The challenge is even greater for Very High Proneness, where the model fails to correctly classify any samples, as indicated by zero values for precision, recall, and F1-score in the metrics and the confusion matrix. Overall, the model achieves a strong overall accuracy of 90% and a weighted average F1-score of 0.90, demonstrating robust performance for well-represented categories like Elevated Proneness, High Proneness, and Moderate Proneness. However, the significant misclassifications for smaller classes such as Low Proneness and the complete lack of correct predictions for Very High Proneness highlight the need for strategies to address class imbalance, such as data augmentation, resampling techniques, or enhanced feature differentiation. These improvements could further refine the model's ability to accurately classify schizophrenia proneness levels across all categories.

## 4. CONCLUSION

This study demonstrated the effectiveness of machine learning models in classifying schizophrenia proneness levels based on behavioral and demographic features. Logistic Regression emerged as the best-performing model, achieving a high overall accuracy of 90% and excelling in predicting well-represented categories such as Elevated and Moderate Proneness. However, the analysis revealed significant challenges in accurately classifying underrepresented categories like Low and Very High Proneness, primarily due to class imbalance and overlapping feature distributions. The results highlighted the potential of machine learning in advancing schizophrenia symptom analysis and enabling early diagnosis. Nonetheless, addressing class imbalance, improving feature selection, and employing techniques such as data augmentation or ensemble methods could have enhanced the model's performance for smaller categories. Future work could focus on integrating additional features or datasets to refine predictions further. This study underscored the promise of machine learning in mental health research, paving the way for more precise and personalized approaches to diagnosing and managing schizophrenia.

## REFERENCES

Abbasi H, Afrazeh F, Ghasemi Y, Ghasemi F. A Shallow Review of Artificial Intelligence Applications in Brain Disease: Stroke, Alzheimer's, and Aneurysm. International Journal of Applied Data Science in Engineering and Health. 2024 Oct 5;1(2):32-43.

Abplanalp SJ, Green MF, Wynn JK, Eisenberger NI, Horan WP, Lee J, McCleery A, Miklowitz DJ, Reddy LF, Reavis EA. Using machine learning to understand social isolation and loneliness in schizophrenia, bipolar disorder, and the community. Schizophrenia. 2024 Oct 5;10(1):88.

Afrazeh F, Shomalzadeh M. Revolutionizing Arthritis Care with Artificial Intelligence: A Comprehensive Review of Diagnostic, Prognostic, and Treatment Innovations. International Journal of Applied Data Science in Engineering and Health. 2024 Sep 10;1(2):7-17.

Chen H, Kim S, Hardie JM, Thirumalaraju P, Gharpure S, Rostamian S, Udayakumar S, Lei Q, Cho G, Kanakasabapathy MK, Shafiee H. Deep learning-assisted sensitive detection of fentanyl using a bubbling-microchip. Lab on a Chip. 2022;22(23):4531-40.

Guo K, Wu M, Soo Z, Yang Y, Zhang Y, Zhang Q, Lin H, Grosser M, Venter D, Zhang G, Lu J. Artificial intelligence-driven biomedical genomics. Knowledge-Based Systems. 2023 Sep 7:110937.

Haug CJ, Drazen JM. Artificial intelligence and machine learning in clinical medicine, 2023. New England Journal of Medicine. 2023 Mar 30;388(13):1201-8.



DOI: 10.54756/IJSAR.2024.26

Hertenstein E, Benz F, Schneider CL, Baglioni C. Insomnia—A risk factor for mental disorders. Journal of sleep research. 2023 Dec;32(6):e13930.

Ibrahim T, Gebril A, Nasr MK, Samad A, Zaki HA. Exploring the Mental Health Challenges of Emergency Medicine and Critical Care Professionals: A Comprehensive Review and Meta-Analysis. Cureus. 2023 Jul;15(7).

Kharaji M, Abbasi H, Orouskhani Y, Shomalzadeh M, Kazemi F, Orouskhani M. Brain Tumor Segmentation with Advanced nnU-Net: Pediatrics and Adults Tumors. Neuroscience Informatics. 2024 Feb 22:100156.

Kufel J, Bargieł-Łączek K, Kocot S, Koźlik M, Bartnikowska W, Janik M, Czogalik Ł, Dudek P, Magiera M, Lis A, Paszkiewicz I. What is machine learning, artificial neural networks and deep learning?—Examples of practical applications in medicine. Diagnostics. 2023 Aug 3;13(15):2582.

Mahmoudiandehkordi S, Yeganegi M, Shomalzadeh M, Ghasemi Y, Kalatehjari M. Enhancing IVF Success: Deep Learning for Accurate Day 3 and Day 5 Embryo Detection from Microscopic Images. International Journal of Applied Data Science in Engineering and Health. 2024 Aug 14;1(1):18-25.

Malhotra K, Wong BN, Lee S, Franco H, Singh C, Silva LA, Iraqi H, Sinha A, Burger S, Breedt DS, Goyal K. Role of artificial intelligence in global surgery: A review of opportunities and challenges. Cureus. 2023 Aug;15(8).

McCutcheon RA, Keefe RS, McGuire PK. Cognitive impairment in schizophrenia: aetiology, pathophysiology, and treatment. Molecular psychiatry. 2023 May;28(5):1902-18.

Minoo S, Ghasemi F. A Narrative Review: Dental Radiology with Deep Learning. International Research in Medical and Health Sciences. 2024 Oct 25;7(5):23-36.

Ng AY, Oberije CJ, Ambrózay É, Szabó E, Serfőző O, Karpati E, Fox G, Glocker B, Morris EA, Forrai G, Kecskemethy PD. Prospective implementation of AI-assisted screen reading to improve early detection of breast cancer. Nature Medicine. 2023 Dec;29(12):3044-9.

Orouskhani M, Zhu C, Rostamian S, Zadeh FS, Shafiei M, Orouskhani Y. Alzheimer's disease detection from structural MRI using conditional deep triplet network. Neuroscience Informatics. 2022 Dec 1;2(4):100066.

Radak M, Lafta HY, Fallahi H. Machine learning and deep learning techniques for breast cancer diagnosis and classification: a comprehensive review of medical imaging studies. Journal of Cancer Research and Clinical Oncology. 2023 Sep;149(12):10473-91.

Rauniyar A, Hagos DH, Jha D, Håkegård JE, Bagci U, Rawat DB, Vlassov V. Federated learning for medical applications: A taxonomy, current trends, challenges, and future research directions. IEEE Internet of Things Journal. 2023 Nov 1.

Sacco R, Camilleri N, Eberhardt J, Umla-Runge K, Newbury-Birch D. A systematic review and meta-analysis on the prevalence of mental disorders among children and adolescents in Europe. European Child & Adolescent Psychiatry. 2024 Sep;33(9):2877-94.

Sharma M, Patel RK, Garg A, SanTan R, Acharya UR. Automated detection of schizophrenia using deep learning: a review for the last decade. Physiological Measurement. 2023 Mar 6;44(3):03TR01.

Tandon R, Nasrallah H, Akbarian S, Carpenter Jr WT, DeLisi LE, Gaebel W, Green MF, Gur RE, Heckers S, Kane JM, Malaspina D. The schizophrenia syndrome, circa 2024: What we know and how that informs its nature. Schizophrenia research. 2024 Feb 1;264:1-28.

van Dee V, Schnack HG, Cahn W. Systematic review and meta-analysis on predictors of prognosis in patients with schizophrenia spectrum disorders: An overview of current evidence and a call for prospective research and open access to datasets. Schizophrenia Research. 2023 Apr 1;254:133-42.

Verma S, Goel T, Tanveer M, Ding W, Sharma R, Murugan R. Machine learning techniques for the Schizophrenia diagnosis: a comprehensive review and future research directions. Journal of Ambient Intelligence and Humanized Computing. 2023 May;14(5):4795-807.



DOI: 10.54756/IJSAR.2024.26

https://www.kaggle.com/code/mdismielhossenabir/schizophrenia-symptoms-analysis

https://www.kaggle.com/datasets/shree23yaa/schizophrenia-symptoms

C. Author: Fanorouzi8@gmail.com



DOI: 10.54756/IJSAR.2024.26