International Conference on Health and Biological Science
ICHBS, Page 141-158, December 2024

Comparison of machine learning algorithms for alzheimer's risk classification

Ria Suci Nurhalizah[1*], *Retno Agus* Setiawan[1], and *Rian* Ardianto[2]

[1]Information System, Faculty of Science and Technology, Harapan Bangsa University, Indonesia
[2]Informatics, Faculty of Science and Technology, Harapan Bangsa University, Indonesia

Abstract. Alzheimer's disease is a progressive form of dementia that affects cognitive function and impacts the quality of life of sufferers. Advancements in artificial intelligence, particularly in machine learning, are creating new possibilities for streamlining the disease classification process using health history data. This study aims to compare the performance of five machine learning algorithms, namely Naive Bayes, Random Forest, Artificial Neural Network (ANN), XGBoost and Support Vector Machine (SVM), in predicting Alzheimer's diagnosis using 2,149 data from Kaggle Open Datasets. The data went through the process of Data Collection, preprocessing, testing using the five algorithms. The evaluation results show that XGBoost has the highest accuracy of 95%, followed by Random Forest with 93% accuracy, ANN 84%, SVM 83%, and Naive Bayes 83%. In conclusion, XGBoost proved to be the most effective model in detecting Alzheimer's, ahead of the other algorithms in this study.

1 Introduction

One of the most prevalent forms of dementia is Alzheimer's, impacting millions of individuals across the globe. This disorder progressively deteriorates, resulting in memory loss and a decline in cognitive abilities (1) (2) (3). Roughly 50 million individuals around the world have alzheimer's. The illness influences approximately 4-9% of individuals matured 60 a long time and over. It is evaluated that the number of individuals with alzheimer's will increment to 152.8 million individuals by 2050 (4). In Indonesia, about 27.9% of the population has Alzheimer's dementia. More than 4.2 million people in Indonesia live with this condition (5). Alzheimer's disease is recognized as a neurological condition that worsens with time and causes cognitive decline and memory loss (6). Planning for long-term care and receiving more effective treatment are made possible by early Alzheimer's disease detection and risk assessment.

New possibilities for streamlining the disease categorization procedure based on medical history data are presented by developments in artificial intelligence technology, especially machine learning (7). Data can be evaluated using machine learning, which can also reveal

* Corresponding author: riasucin@email.com

novel patterns that advance scientific knowledge (8). Several machine learning algorithms have been employed to identify trends in health data in order to predict the risk of Alzheimer's disease (9). Some popular algorithms such as Naive Bayes, Random Forest, Artificial Neural Network (ANN), XGBoost and Support Vector Machine (SVM) have their own strengths and weaknesses in handling various data characteristics.

Previous research (10) aimed to determine which model best predicts Alzheimer's disease. A number of machine learning methods were compared in the study, including Artificial Neural Network (ANN), Logistic Regression (LR), Naive Bayes (NB), Random Forest (RF), and Support Vector Machine (SVM). The Logistic Regression technique had the highest accuracy, 85.71%, according on the results of testing data on 373 Alzheimer's patients from Kaggle.

In the subsequent study, referenced in (11) machine learning models utilizing texture features were employed for the multi-stage classification of Alzheimer's disease, facilitating earlier identification of the condition. Four groups of patients: sMCI (Static Mild Cognitive Impairment), AD, CN (Cognitive Normal), and pMCI (Progressive Mild Cognitive Impairment were created using a dataset of 15,000 MRI scans from the ADNI database. Following preprocessing, 8,856 photos were chosen to be used in the FOS and GLCM procedures for feature extraction. With an accuracy of 66.196%, the Random Forest model outperformed the Decision Tree (56.4%), ANN (58.5%), and SVM (59.2%) models. This approach shows great promise for the multi-stage categorization of Alzheimer's disease.

The application of machine learning techniques for early Alzheimer's disease diagnosis was the focus of another study by (12) Among the algorithms employed was the Light Gradient Boosting Machine (LGBM), which classified five ADNI classes with 99.63% accuracy: the Multilayer Perceptron (MLP), which attained 95.75% accuracy in classifying three OASIS classes; the K-nearest Neighbour (KNN), which reached 87.50% accuracy in classifying two ADNI-OASIS classes; and Gaussian Naive Bayes, which yielded 77.97% accuracy for the OASIS-ADNI approach. The combined (Hybrid) method utilizing LGBM with optimized hyperparameters achieved an accuracy of 99.21%. Additionally, Explainable AI techniques such as LIME were employed to enhance the model's interpretability, proving that machine learning works very well for Alzheimer's disease early diagnosis.

In another study (13), Researchers sought to assess how well artificial neural network (ANN) models could diagnose Alzheimer's disease while accounting for environmental and hereditary variables. The ANN model employed 51 genetic SNPs and 8 environmental factors as features, and the sample included 3,773 healthy people and 184 AD patients. With this model, the Random Forest (RF) and Support Vector Machine (SVM) approaches were compared. Thus, even in the absence of age and genetic data, the ANN model demonstrated remarkable sensitivity (0.95), specificity (0.96), and accuracy (0.98; accuracy: 0.97). For AD prediction, the three primary SNPs that were determined to be crucial were CASS4 rs7274581, PICALM rs3851179, and TOMM40 rs2075650. The ANN model demonstrated the ability to handle the intricacy of Alzheimer's illness by performing similarly to both RF and SVM.

This study analyses the way five machine learning algorithms Support Vector Machine (SVM), Naive Bayes, Random Forest, Artificial Neural Network (ANN), and XGBoost perform in categorizing the risk of Alzheimer's disease. The study was based on 2,149 health history records that were obtained from Kaggle and divided in an 80:20 ratio between training and testing sets. Each algorithm's performance will be assessed using measures including F1-score, recall, accuracy, and precision. Finding the best algorithm to reliably and effectively identify the risk of Alzheimer's disease is the aim.

2 Research methods

The study involves multiple stages, starting with gathering the data required for analysis and dividing it into 80:20 training and test sets. Next, the algorithm models used are Naive Bayes, Random Forest, Artificial Neural Network (ANN), XGBoost and Support Vector Machine (SVM). Training and performance analysis of the machine learning algorithm models examined in this study are carried out in the final stage. The stages of this process can be seen visually below.

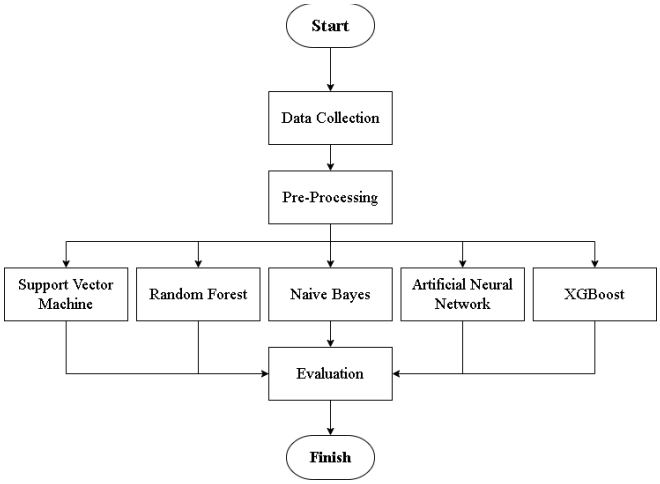

Fig. 1. Flow of Method

2.1 Data collection

Data gathering is one of the most crucial initial steps in the data analysis process. In this stage, there is a collection of data about Alzheimer's disease patients obtained from the Kaggle Open Datasets site which consists of 2,149 data which can be seen in Table 1. (https://www.kaggle.com/datasets/rabieelkharoua/alzheimers-disease-dataset).

Table. 1 Alzheimer's Patient Dataset

	PatientID	Age	Gender	Ethnicity	Education	BMI	Smoking	AlcoholCo	PhysicalA	DietQualit	SleepQual	FamilyHist	Cardiovas	Diabetes	Depressio	HeadInjury	Hypertens	SystolicBP	DiastolicB
0	4751	73	0	0	2	22.928	0	13.297	6.327	1.347	9.026	0	0	1	1	0	0	142	72
1	4752	89	0	0	0	26.828	0	4.543	7.620	0.519	7.151	0	0	0	0	0	0	115	64
2	4753	73	0	3	1	17.796	0	19.555	7.845	1.826	9.674	1	0	0	0	0	0	99	116
3	4754	74	1	0	1	33.801	1	12.209	8.428	7.436	8.393	0	0	0	0	0	0	118	115
4	4755	89	0	0	0	20.717	0	18.454	6.310	0.795	5.597	0	0	0	0	0	0	94	117
...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...
2144	6895	61	0	0	1	39.122	0	1.561	4.050	6.555	7.536	0	0	0	0	0	0	122	101
2145	6896	75	0	0	2	17.858	0	18.767	1.361	2.905	8.555	0	0	0	0	0	0	152	106
2146	6897	77	0	0	1	15.476	0	4.595	9.886	8.120	5.769	0	0	0	0	0	0	115	118
2147	6898	78	1	3	1	15.300	0	8.675	6.354	1.263	8.323	0	1	0	0	0	0	103	96
2148	6899	72	0	0	2	33.290	0	7.891	6.571	7.941	9.879	0	0	0	0	0	0	166	78

	Cholesterol	Cholesterol	Cholesterol	Cholesterol	MMSE	FunctionalA	MemoryC	Behaviora	ADL	Confusion	Disorienta	Personalit	DifficultyC	Forgetfuln	Diagnosis	DoctorInCharge
0	Cholesterol	56.151	33.683	162.189	21.464	6.519	0	0	1.726	0	0	0	1	0	0	XXXConfid
1	Cholesterol	193.408	79.028	294.631	20.613	7.119	0	0	2.592	0	0	0	0	1	0	XXXConfid
2	Cholesterol	153.323	69.772	83.638	7.356	5.895	0	0	7.120	0	1	0	1	0	0	XXXConfid
3	Cholesterol	65.367	68.457	277.577	13.991	8.965	0	1	6.481	0	0	0	0	0	0	XXXConfid
4	Cholesterol	92.870	56.874	291.199	13.518	6.045	0	0	0.015	0	0	1	1	0	0	XXXConfid
...	Cholesterol	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...
2144	Cholesterol	94.870	60.943	234.520	1.201	0.239	0	0	4.493	1	0	0	0	0	1	XXXConfid
2145	Cholesterol	95.411	93.650	367.987	6.458	8.687	0	1	9.205	0	0	0	0	0	1	XXXConfid
2146	Cholesterol	156.267	99.678	294.802	17.011	1.972	0	0	5.036	0	0	0	0	0	1	XXXConfid
2147	Cholesterol	52.483	81.281	145.254	4.030	5.174	0	0	3.785	0	0	0	0	1	1	XXXConfid
2148	Cholesterol	92.200	81.920	217.397	11.115	6.308	0	1	8.328	0	1	0	0	1	0	XXXConfid

2.2 Preprocessing

The stage performed after data collection is preprocessing before it is analyzed or used in machine learning models. This stage includes several important steps, such as addressing missing data, changing the data format, and normalization. In addition, preprocessing also involves removing duplicates, handling outliers, and coding categorical variables (14). This process is very important as clean and well-structured data can improve model performance and produce more accurate results (15). Additionally, the process of breaking a dataset up into several subsets, often divided into two main categories: testing data and training data, is known as data splitting (16). This study uses an 80:20 ratio, which means that half of the data is used to evaluate the model's performance and the other half is used to train the model.

2.3 Support vector machine

The Support Vector Machine (SVM) is a popular machine learning technique for both regression and classification applications (17). By learning the data distribution, it can find an appropriate classification line (or ideal hyperplane) that is not constrained to a linear form (18). The boundaries of the classes are established using the data points and just the support vectors that are closest to the maximum margin (19). The primary objective of SVM is to use a surface that maximizes the margin between classes to differentiate between classes in the training data (20). Equation 1 shows the formula for SVM with a linear kernel:

$$f(x) = wTx + b \quad (1)$$

2.4 Random forest

Using the Bagging technique, A number of decision trees make up Random Forest, an ensemble learning system that is trained on various data subsets (21). This algorithm assists in eliminating outliers and classifying the dataset based on relevant features. Random Forest is especially good at managing complicated and diverse data collected from various sources (22). To determine out how informative a node is on an attribute and to measure the amount of information obtained, the random forest technique starts by computing the gini impurity value and the average gini impurity. The formula for determining gini impurity in Random Forest can be seen in Equation 2:

$$Gini = 1 - \sum_{i=1}^{n} (p_i)^2 \quad (2)$$

2.5 Naive bayes

The foundation of the Naive Bayes classification method is Bayes' Theorem, which operates under the assumption that each predictor functions independently of the others. In essence, one feature's occurrence in a class is assumed to be independent of the existence of any other features by the Naive Bayes algorithm (23). This algorithm employs a straightforward probability approach and does not necessitate complex parameters (24). Naive Bayes is recognized as a highly efficient and uncomplicated probability-based classification algorithm. Bayes' theorem can be used with the following that can be seen in Equation 3:

$$P(H|X) = \frac{P(X|H) \times P(H))}{(P(X))} \quad (3)$$

2.6 Artificial neural network

A computer model that takes inspiration from the way the human brain works is called an artificial neural network (ANN) (25). The calculations in a multilayer ANN are executed by linking numerous artificial neurons, which are composed of multiple output layers, an input layer, and an hidden layer (13). Neurons in each layer are often coupled to neurons in the layers above and below, but connections within the same layer are not allowed. Data flows through the network until the model reaches the desired accuracy level (26).

2.7 XGBoost

A machine learning approach called XGBoost (Extreme Gradient Boosting) enhances the boosting method (27). This algorithm is an ensemble learning method with many advantages, including high flexibility, strong predictive capability, excellent generalization, high scalability, efficient model training, and remarkable robustness (28). XGBoost is specifically designed to enhance model speed and performance by incorporating regularization techniques to mitigate overfitting (29). The minimized objective function can be seen in Equation 4:

$$Obj(t) = \sum_{i=1}^{n} l\,(yi, yi\hat{}(t-1) + ft(xi)) + \Omega(ft) \quad (4)$$

2.8 Evaluation

Metrics assessment, recall, F1-score, accuracy, and precision are used to assess and analyse the performance of the five machine learning algorithms. These standards were chosen because they provide a comprehensive picture of the model's ability to reliably and accurately predict Alzheimer's diagnoses. While precision and recall concentrate on striking a balance between true positive predictions and the model's capacity to recognize relevant cases, accuracy shows the percentage of accurate forecasts. Because the F1-score creates a single statistic that includes precision and recall, it provides a more objective assessment. Additionally, by showing the percentage of accurate and inaccurate forecasts for each class, the confusion matrix offers a more comprehensive understanding of the model's performance. This helps identify any possible problems with the model and areas that require improvement.

3 Result and discussion

3.1 Data collection and visualization

The study's dataset, which was sourced from Kaggle Open Datasets, includes comprehensive medical records for 2,149 individuals, each of whom has a unique ID between 4751 and 6900. This dataset consists of 35 variables, including functional evaluations, cognitive symptoms, clinical measurements, medical history, lifestyle factors, demographic data, and an Alzheimer's disease diagnosis. Below is a detailed explanation of each feature category related to the target variable:

3.1.1 Demographic Details

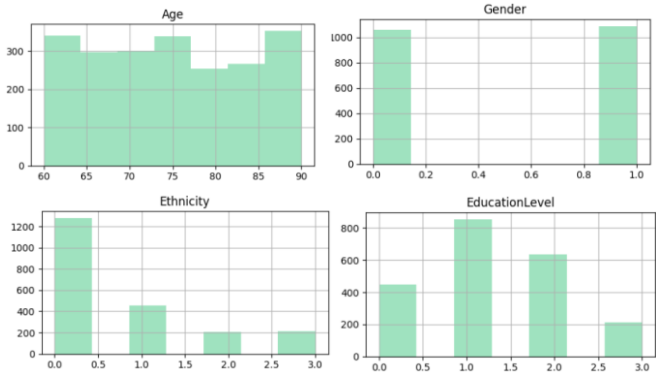

Fig. 2. Histogram of Distribution Demographic Details

Fig. 2. shows the demographic features of the patients, including age, gender, ethnicity, and education level. Patients' ages range from 60 to 90 years old. In terms of gender, 0 denotes male patients, and 1 denotes female patients. Ethnicity is categorized into four groups: 0 for Caucasian, 1 for African American, 2 for Asian, and 3 for other ethnicities. The education levels of the patients are categorized as follows: 0 indicates no formal education, 1 represents a high school graduate, 2 corresponds to a college degree, and 3 signifies higher education.

3.1.2 Lifestyle Factors

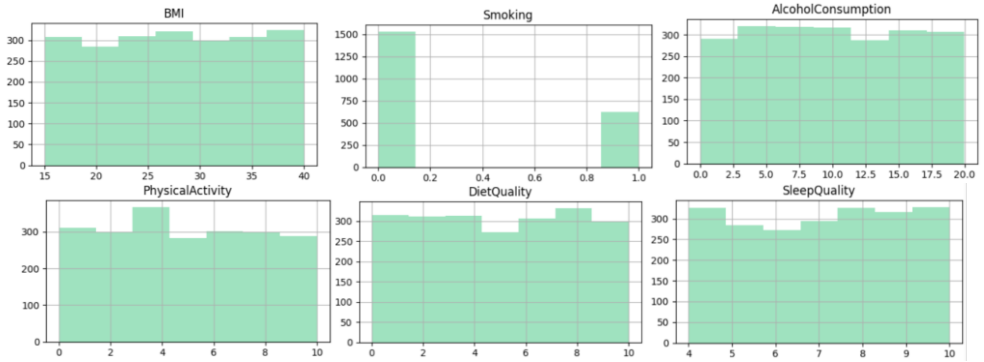

Fig. 3. Histogram of Distribution Lifestyle Factors

Fig. 3. illustrates the lifestyle factors of the patients, including diet quality, physical activity, alcohol consumption, smoking habits, Body Mass Index (BMI), and sleep quality. Patients' BMI varied between 15 and 40. Smokers were assigned a smoking status of 1, while non-smokers were assigned a smoking status of 0. The weekly units used to assess alcohol consumption ranged from 0 to 20. Hours of physical exercise each week were recorded, ranging from 0 to 10. A rating from 0 to 10 was used to evaluate diet quality, while a scale from 4 to 10 was used to evaluate sleep quality. These factors provide an overview of the patients' lifestyles and their impact on health.

3.1.3 Medical History

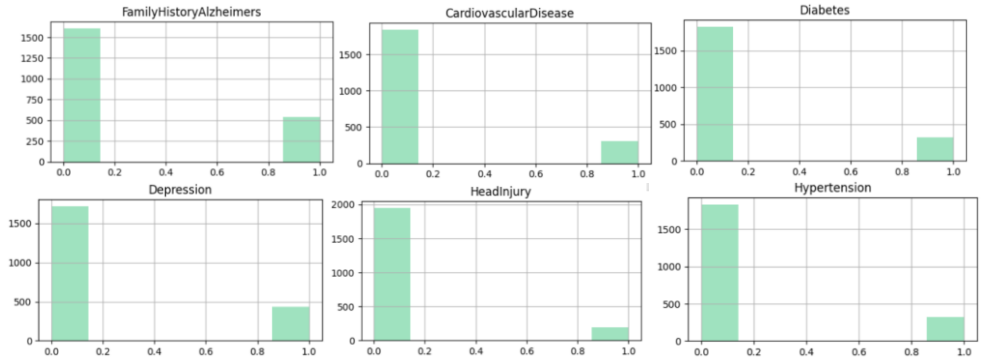

Fig. 4. Histogram of Distribution Medical History

Fig. 4. shows the patients' medical history, which includes several key factors such as family history of Alzheimer's, hypertension, head injury, depression, diabetes, and cardiovascular disease. Family history of Alzheimer's is recorded as 0 for no history and 1 for a positive history. The same binary system is applied to cardiovascular disease, diabetes, depression, head injury, and hypertension, with 0 indicating no disease and 1 indicating the presence of the condition. These factors are crucial for understanding the patients' overall health and risk profile.

3.1.4 Clinical Measurements

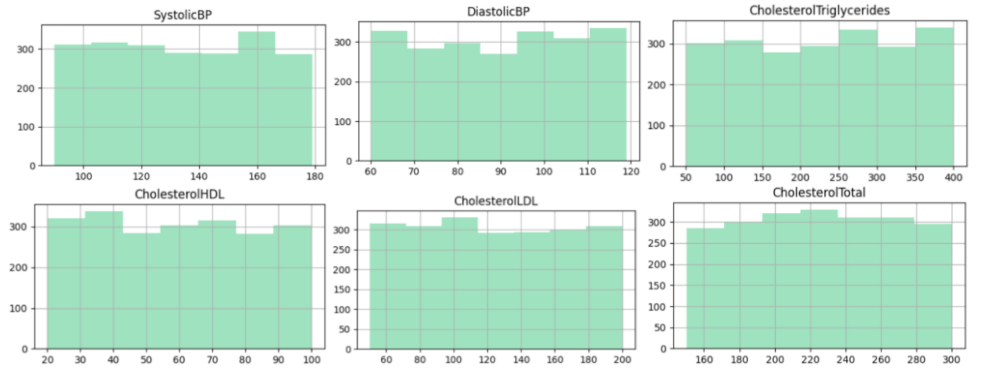

Fig. 5. Histogram of Distribution Clinical Measurements

Fig. 5. presents the patients' clinical measurements, which include systolic and diastolic blood pressure as well as various cholesterol levels. Systolic blood pressure ranges from 90 to 180 mmHg, and diastolic blood pressure ranges from 60 to 120 mmHg. LDL (low-density lipoprotein) cholesterol ranges from 50 to 200 mg/dL, whereas HDL (high-density lipoprotein) cholesterol ranges from 20 to 100 mg/dL. Total cholesterol levels are between 150 and 300 mg/dL. The range of triglyceride values is 50–400 mg/dL. These measurements provide a comprehensive overview of the patients' cardiovascular and lipid health.

3.1.5 Cognitive and Functional Assessments

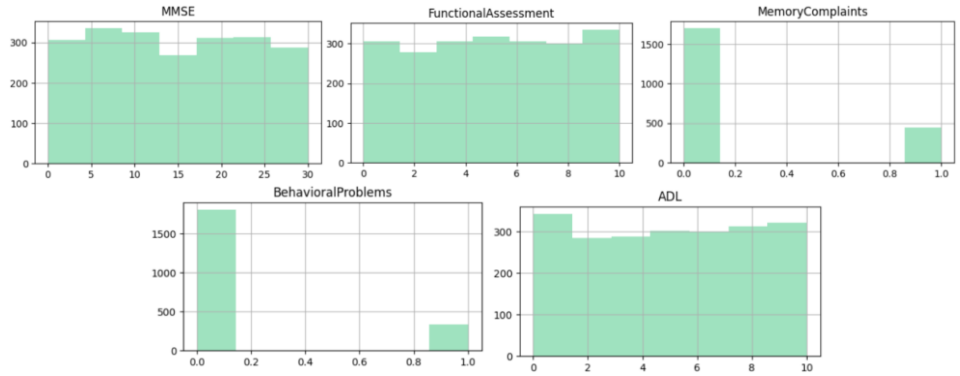

Fig. 6. Histogram of Distribution Cognitive and Functional Assessments

Fig. 6. illustrates the cognitive and functional assessments of the patients, which were conducted through several steps. The Mini-Mental State Examination (MMSE) Pd, which ranges from 0 - 30, indicates cognitive impairment. Functional assessments are rated on a scale of 0 to 10, where a low score signifies greater impairment in daily activities. Memory complaints and behavioural issues are recorded as 0 for no problems and 1 for the presence of problems. Additionally, Activity of Daily Living (ADL) scores also range from 0 to 10, where lower scores indicate more significant difficulty in performing daily tasks.

3.1.6 Symptoms

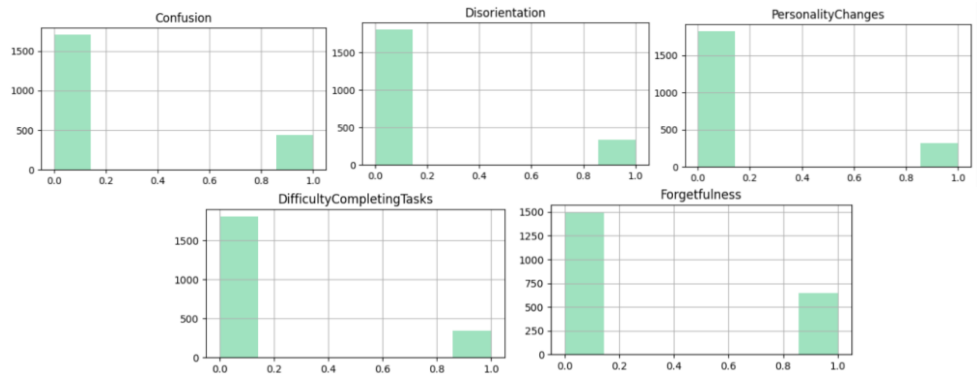

Fig. 7. Histogram of Distribution Symptoms

Fig. 7. shows the various symptoms that can be identified in patients. First, confusion is recorded as 0 for none and 1 for present. Disorientation is assessed in the same manner, with 0 indicating no disorientation and 1 indicating its presence. Additionally, personality changes are noted, where 0 signifies no change and 1 indicates a change. Another symptom, difficulty in completing tasks, is similarly recorded with 0 for no difficulty and 1 for the presence of difficulty. Lastly, forgetfulness is assessed using the same scale, where 0 represents no forgetfulness and 1 indicates forgetfulness.

3.1.7 Diagnosis Information

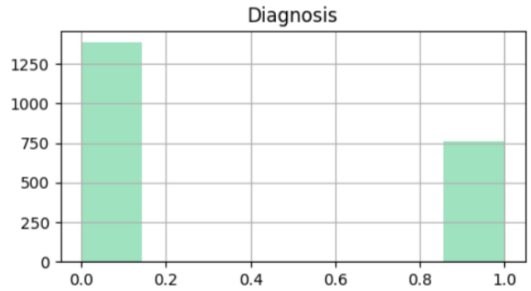

Fig. 8. Histogram of Distribution Diagnosis Information

Fig. 8. displays the diagnosis status for Alzheimer's Disease, utilizing a scale where 0 signifies no diagnosis and 1 indicates a confirmed diagnosis. Besides, Fig. 9. shows the correlation relationship between various health and diagnostic variables related to Alzheimer's. An rise in one variable is linked to an increase in another, as indicated by the color red, It indicates a favorable correlation. Conversely, a negative correlation is shown by blue, indicating that a rise in one measure is associated with a fall in another. For example, Age has a positive correlation with Alzheimer's diagnosis, while MMSE score (which measures cognitive function) is negatively correlated with Alzheimer's, meaning that the lower the cognitive score, the more likely a person is to develop Alzheimer's. Some symptoms such as Confusion, Disorientation, and PersonalityChanges have strong correlations with an Alzheimer's diagnosis, suggesting that these symptoms are often present in patients with the condition. This heatmap helps identify important relationships between health variables and can be used to inform diagnostic measures and disease risk assessment.

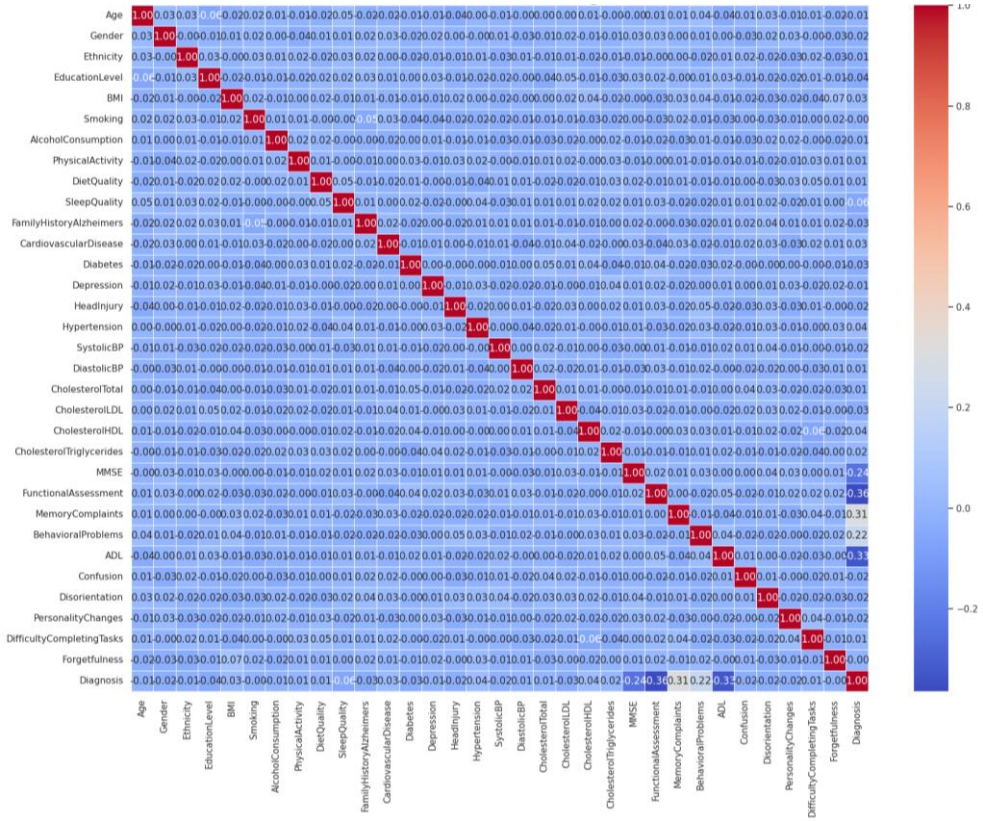

Fig. 9. Heatmap of Correlation Matrix

3.2 Model Result

3.2.1 SVM Model

Table. 2 Classification Report of SVM

	precision	recall	f1-score	support
0	0.85	0.90	0.87	277
1	0.80	0.71	0.75	153
Accuracy			0.83	430
Macro avg	0.82	0.80	0.81	430
Weighted avg	0.83	0.83	0.83	430

Training set score: 0.9401, Testing set score: 0.8326

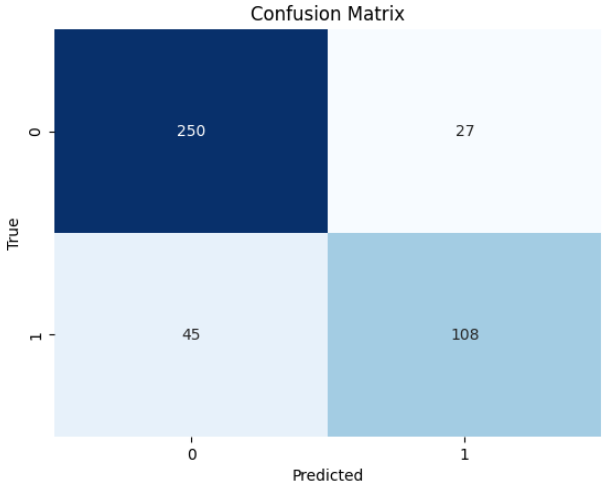

Fig. 10. SVM Model Confusion Matrix

Table. 2 shows the SVM model's performance measures, including precision, recall, and F1-score. On the training data, the model's accuracy was 94%, whereas on the test data, it was 83%. In contrast to class 1, which had 80% precision and 71% recall, class 0 had 85% precision and 90% recall. Furthermore, class 0 and class 1 had F1-scores of 0.87 and 0.75, respectively. The recall, F1-score, and precision, macro averages were 82%, 81%, and 80%, respectively. Given that the model performs better on the training data than the test data, there are signs that it may be overfitting. The confusion matrix in Fig. 10. indicates that there were 358 correct predictions and 72 incorrect predictions.

3.2.2 Random Forest Model

Table. 3 Classification Report of Random Forest Model

	precision	recall	f1-score	support
0	0.91	0.98	0.94	277
1	0.96	0.82	0.89	153
Accuracy			0.93	430
Macro avg	0.94	0.90	0.92	430
Weighted avg	0.93	0.93	0.92	430

Training set score: 1.0000, Testing set score: 0.9256

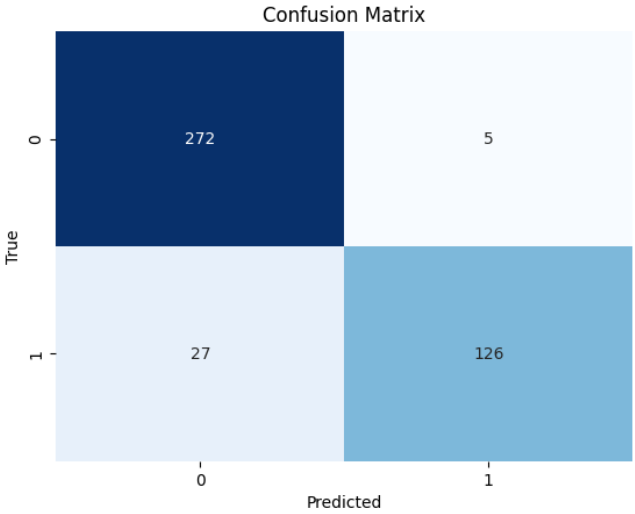

Fig. 11. Random Forest Model Confusion Matrix

The Random Forest model presented in Table 3 achieved 93% accuracy on the test data and 100% on the training data, indicating possible overfitting. The precision for class 0 is 91%, with a recall of 98%, while class 1 exhibits higher precision at 96% but a lower recall of 82%. The model performs better for class 0 as evidenced by the F1-score of 94% for class 0 and 89% for class 1. With a weighted average of almost 93% for each parameter, the overall macro averages for recall, F1-score, and precision, are 90%, 92%, and 94%, respectively. The confusion matrix in Fig. 11. reveals that there were 398 correct predictions and 32 incorrect predictions.

3.2.3 Naive Bayes Model

Table. 4 Classification Report of Naive Bayes Model

	precision	recall	f1-score	support
0	0.85	0.89	0.87	277
1	0.79	0.73	0.76	153
Accuracy			0.83	430
Macro avg	0.82	0.81	0.81	430
Weighted avg	0.83	0.83	0.83	430

Training set score: 0.8202, Testing set score: 0.8326

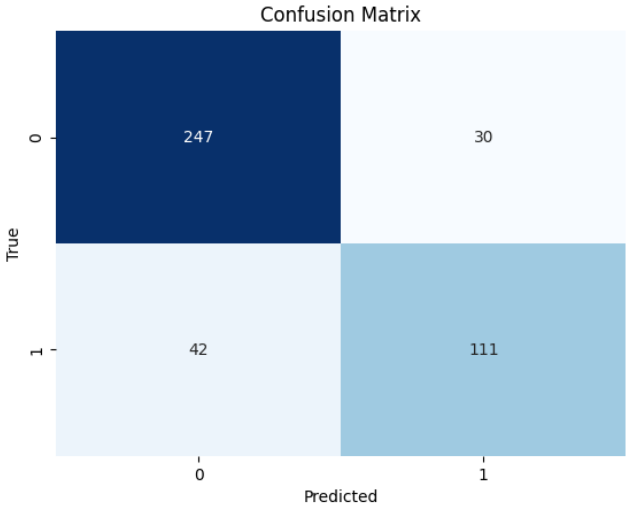

Fig. 12. Naive Bayes Model Confusion Matrix

The Naive Bayes model detailed in Table 4 demonstrated consistent performance on both datasets with an accuracy of 83% on the test data and 82% on the training data. Class 1 has a 79% precision and a 73% recall, but class 0 has an 85% precision and an 89% recall. This implies that class 0 can be identified by the model more successfully than class 1. Class 0's F1-score is 87%, whereas class 1's is 76%. The model's balanced performance across both classes is demonstrated by the overall macro averages for recall, precision, and F1-score, which are 81%, 81%, and 82%, respectively, with a weighted average of over 83% for all measures. According to the confusion matrix in Fig. 12. 358 predictions were right and 72 were wrong.

3.2.4 ANN Model

Table. 5 Classification Report of ANN Model

	precision	recall	f1-score	support
0	0.86	0.90	0.88	277
1	0.81	0.73	0.77	153
Accuracy			0.84	430
Macro avg	0.83	0.82	0.82	430
Weighted avg	0.84	0.84	0.84	430

Training set score: 0.8965, Testing set score: 0.8419

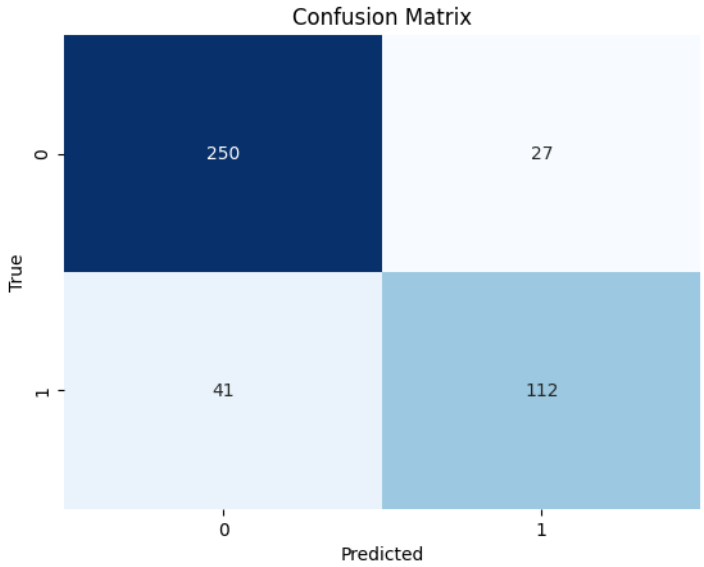

Fig. 13. ANN Model Confusion Matrix

The Artificial Neural Network (ANN) model presented in Table 5 obtained 90% accuracy on training data and 84% accuracy on test data, suggesting good performance with a small chance of overfitting. Class 1 has an 81% precision and a 73% recall, but class 0 has an 86% precision and 90% recall. This suggests that class 0 can be identified by the model more accurately than class 1. Class 0's F1-score is 88%, whereas class 1's is 77%. The model's balanced performance across all classes is demonstrated by the overall macro averages for recall, F1-score, and precision which are 82%, 82%, and 83%, respectively, with a weighted average of almost 84%. The confusion matrix in Fig. 13. indicates that there were 362 correct predictions and 68 incorrect predictions.

3.2.5 XGBoost Model

Table. 6 Classification Report of XGBoost Model

	precision	recall	f1-score	support
0	0.94	0.98	0.96	277
1	0.96	0.90	0.93	153
Accuracy			0.95	430
Macro avg	0.95	0.94	0.94	430
Weighted avg	0.95	0.95	0.95	430

Training set score: 1.0000, Testing set score: 0.9488

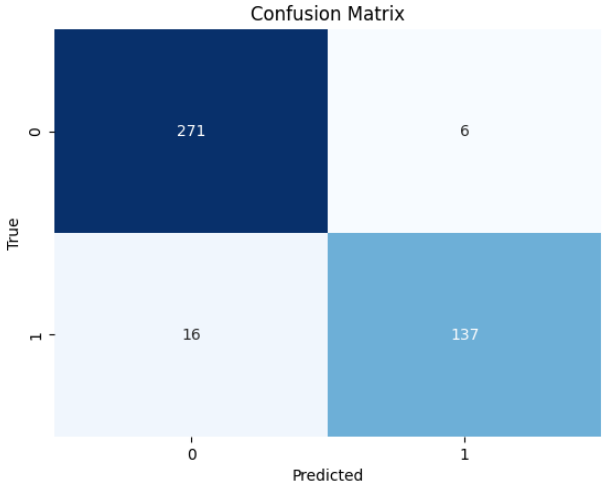

Fig. 14. XGBoost Model Confusion Matrix

The XGBoost model illustrated in Table 6 demonstrates excellent performance, achieving an accuracy of 95% on the test data and 100% on the training data, which suggests potential overfitting. For class 0, the precision is 94% with a recall of 98%, while for class 1, the precision reaches 96% with a recall of 90%, demonstrating how well the model can distinguish between the two classes. The F1-score for class 0 is 96%, while for class 1, it is 93%. Overall, the macro averages for recall, F1-score, and precision are 94%, 94%, and 95%, respectively, with the weighted average also approximately 95% across all metrics, reflecting the model's balanced and reliable performance in both classes. The confusion matrix in Fig. 14. reveals that there were 398 correct predictions and 22 incorrect predictions.

3.3 Model Comparison

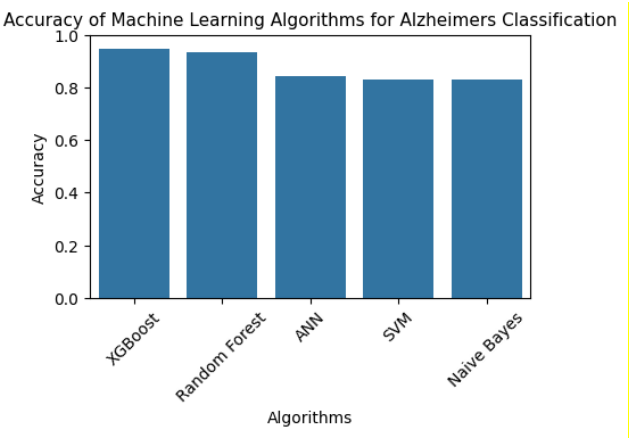

Fig. 15. Accuracy of Machine Learning Algorithms for Alzheimer's Classification

Fig. 15. illustrates the comparison of accuracy scores among several machine learning algorithms, highlighting that the XGBoost model outperforms the others in the system. XGBoost achieves the highest accuracy at 95%, followed by Random Forest with an accuracy

of 93%. The ANN model records an accuracy of 84%, while both the SVM and Naive Bayes models exhibit the same accuracy of 83%. As a result, XGBoost and Random Forest perform significantly more accurate than ANN, SVM, and Naive Bayes. Compared to previous studies, (10) found good accuracy for SVM and Naive Bayes, which aligns with the results of this study. Additionally, (11) using MRI texture features, reported a lower accuracy for Random Forest (66.2%) compared to this study, indicating that structured health data may offer higher predictive power.

4 Conclusion

Based on the analysis and research conducted, this study concludes that it aims to evaluate the effectiveness of five machine learning algorithms: ANN, Random Forest, XGBoost, Naive Bayes, and SVM, utilizing a dataset comprising 2,149 Alzheimer's patient records sourced from Kaggle Open Datasets. The outcomes of the test demonstrate that the XGBoost algorithm exhibits the best performance among the five algorithms, achieving an accuracy rate of 95%. This finding confirms that XGBoost is the most effective algorithm in this study for detecting and predicting Alzheimer's disease, highlighting its significant potential to aid in the early diagnosis process.

References

1. Shah S, Shah M. The effects of machine learning algorithms in magnetic resonance imaging (MRI), and biomarkers on early detection of Alzheimer’s disease. Adv Biomark Sci Technol [Internet]. 2024;6:191–208. Available from: https://linkinghub.elsevier.com/retrieve/pii/S2543106424000152
2. Zhang S, Yuan J, Sun Y, Wu F, Liu Z, Zhai F, et al. Machine learning on longitudinal multi-modal data enables the understanding and prognosis of Alzheimer’s disease progression. iScience. 2024 Jul 19;27(7).
3. Sharma A, Kaur S, Memon N, Jainul Fathima A, Ray S, Bhatt MW. Alzheimer’s patients detection using support vector machine (SVM) with quantitative analysis. Neuroscience Informatics. 2021 Nov;1(3):100012.
4. Sánchez-Reyna AG, Celaya-Padilla JM, Galván-Tejada CE, Luna-García H, Gamboa-Rosales H, Ramirez-Morales A, et al. Multimodal early alzheimer’s detection, a genetic algorithm approach with support vector machines. Healthcare (Switzerland). 2021 Aug 1;9(8).
5. Bestari AP. KEMENKES. 2023 [cited 2024 Sep 19]. Mengenal Demensia Alzheimer. Available from: https://yankes.kemkes.go.id/view_artikel/2819/mengenal-demensia-alzheimer
6. Taie SA, Ghonaim W. A new model for early diagnosis of alzheimer’s disease based on bat-svm classifier. Bulletin of Electrical Engineering and Informatics. 2021;10(2):759–66.
7. Tanveer M, Richhariya B, Khan RU, Rashid AH, Khanna P, Prasad M, et al. Machine learning techniques for the diagnosis of alzheimer’s disease: A review. ACM Transactions on Multimedia Computing, Communications and Applications. 2020 Apr 1;16(1s).
8. Purwono P, Wirasto A, Setiawan RA, Triwibowo DN, Zuhrufillah I, Sumantri RBB, et al. Analysis of the relationship between fetal health prediction features with machine learning feyn qlattice regression model. In: 3rd Borobudur International Symposium On Science And Technology 2021 [Internet]. AIP Conference Proceedings; 2023

[cited 2024 Nov 14]. p. 020183. Available from: https://pubs.aip.org/aip/acp/article-abstract/2706/1/020183/2889424/Analysis-of-the-relationship-between-fetal-health?redirectedFrom=fulltext
9. Khan A, Zubair S. Development of a three tiered cognitive hybrid machine learning algorithm for effective diagnosis of Alzheimer's disease. Journal of King Saud University - Computer and Information Sciences. 2022 Nov 1;34(10):8000–18.
10. Firman Akbar, Rahmaddeni. Komparasi Algoritma Machine Learning Untuk Memprediksi Penyakit Alzheimer. Jurnal Komputer Terapan. 2022 Nov 30;8(2):236–45.
11. Gill LS, Kaur J, Goel N. Machine learning and texture features based approach for classifying Alzheimer's disease. In: Procedia Computer Science. Elsevier B.V.; 2024. p. 2741–8.
12. Chakraborty M, Naoal N, Momen S, Mohammed N. Analyze-Ad: A comparative analysis of novel AI approaches for early Alzheimer's detection. Array. 2024 Jul 1;22.
13. Huang YH, Chen YC, Ho WM, Lee RG, Chung RH, Liu YL, et al. Classifying Alzheimer's disease and normal subjects using machine learning techniques and genetic-environmental features. Journal of the Formosan Medical Association. 2024 Jun 1;123(6):701–9.
14. Fan C, Chen M, Wang X, Wang J, Huang B. A Review on Data Preprocessing Techniques Toward Efficient and Reliable Knowledge Discovery From Building Operational Data. Vol. 9, Frontiers in Energy Research. Frontiers Media S.A.; 2021.
15. Maharana K, Mondal S, Nemade B. A review: Data pre-processing and data augmentation techniques. Global Transitions Proceedings. 2022 Jun;3(1):91–9.
16. Joseph VR. Optimal ratio for data splitting. Stat Anal Data Min. 2022 Aug 1;15(4):531–8.
17. Afzal S, Maqsood M, Khan U, Mehmood I, Nawaz H, Aadil F, et al. Alzheimer disease detection techniques and methods: A review. Vol. 6, International Journal of Interactive Multimedia and Artificial Intelligence. Universidad Internacional de la Rioja; 2021. p. 26–38.
18. Vichianin Y, Khummongkol A, Chiewvit P, Raksthaput A, Chaichanettee S, Aoonkaew N, et al. Accuracy of Support-Vector Machines for Diagnosis of Alzheimer's Disease, Using Volume of Brain Obtained by Structural MRI at Siriraj Hospital. Front Neurol. 2021 May 10;12.
19. Farhadian M, Shokouhi P, Torkzaban P. A decision support system based on support vector machine for diagnosis of periodontal disease. BMC Res Notes. 2020 Jul 13;13(1).
20. Mohamed ES, Naqishbandi TA, Bukhari SAC, Rauf I, Sawrikar V, Hussain A. A hybrid mental health prediction model using Support Vector Machine, Multilayer Perceptron, and Random Forest algorithms. Healthcare Analytics. 2023 Nov 1;3.
21. Franciotti R, Nardini D, Russo M, Onofrj M, Sensi SL. Comparison of Machine Learning-based Approaches to Predict the Conversion to Alzheimer's Disease from Mild Cognitive Impairment. Neuroscience. 2023 Mar 15;514:143–52.
22. Patel L, Shukla T, Huang X, Ussery DW, Wang S. Machine Learning Methods in Drug Discovery. Molecules. 2020 Nov 2;25(22).
23. Mahesh B. Machine Learning Algorithms-A Review. International Journal of Science and Research [Internet]. 2018; Available from: www.ijsr.net
24. Wildah SK, Agustiani S, S MRR, Gata W, Nawawi HM. Deteksi Penyakit Alzheimer menggunakan Algoritma Naïve Bayes dan Correlation Based Feature Selection. Jurnal Informatika. 2020 Sep 7;7(2):166–73.

25. Abdolrasol MGM, Suhail Hussain SM, Ustun TS, Sarker MR, Hannan MA, Mohamed R, et al. Artificial neural networks based optimization techniques: A review. Vol. 10, Electronics (Switzerland). MDPI; 2021.
26. Niazkar HR, Niazkar M. Application of artificial neural networks to predict the COVID-19 outbreak. Glob Health Res Policy. 2020 Dec 1;5(1).
27. Ibrahem Ahmed Osman A, Najah Ahmed A, Chow MF, Feng Huang Y, El-Shafie A. Extreme gradient boosting (Xgboost) model to predict the groundwater levels in Selangor Malaysia. Ain Shams Engineering Journal. 2021 Jun 1;12(2):1545–56.
28. Zhang P, Jia Y, Shang Y. Research and application of XGBoost in imbalanced data. Int J Distrib Sens Netw. 2022 Jun 1;18(6).
29. Liew XY, Hameed N, Clos J. An investigation of XGBoost-based algorithm for breast cancer classification. Machine Learning with Applications. 2021 Dec;6:100154.