SciNexuses

Journal Homepage: sciencesforce.com/scin

SciNexuses Vol. 1 (2024) 174–183

Paper Type: Original Article

Alzheimer's Disease Prediction using Hybrid Machine Learning Techniques

Raghad Ahmed Gad [1,*] and Ahmed Abdelhafeez [2,3]

[1] Faculty of Information Systems and Computer Science, October 6th University, 12585, Egypt; Raghadahmed213@gmail.com.
[2] Computer Science Department, Faculty of Information System and Computer Science, October 6 University, Giza, 12585, Egypt; aahafeez.scis@o6u.edu.eg.
[3] Applied Science Research Center. Applied Science Private University, Amman, Jordan; a_abdelhafeez@asrc.asu.edu.jo.

Received: 25 May 2024 **Revised:** 26 Nov 2024 **Accepted:** 27 Dec 2024 **Published:** 30 Dec 2024

Abstract

Artificial intelligence (AI) and machine learning (ML) have shown benefits in many domains during their growth, particularly considering the enormous amount of data generated recently. For quicker and more precise decision-making regarding illness projections, it might be more dependable. Models can be used to analyze and visualize diseases. The article compares several machine learning algorithms and hybrid machine learning models. A range of machine learning techniques is also available. The following techniques were tested: Random Forest, AdaBoost Classifier, Gaussian NB, Decision Tree, and Logistic Regression (LR). Following the process, the RF classifier performs better than previous algorithms with an accuracy of 92.5%. We created two sets of hybrid models: two-classifier and three-classifier combinations. The best-performing models yielded impressive results; RF and AdaBoost achieved 92.55% accuracy in the two-classifier combinations. Of the three classifier combinations, the accuracy of DT, AdaBoost, and LR was the greatest at 95.46%.

Keywords: Artificial Intelligence; Machine Learning; Logistic Regression; Alzheimer's Disease.

1 | Introduction

The most common cause of dementia in older people globally is Alzheimer's disease (AD), a neurological condition. It gradually impairs cognitive and everyday functioning, resulting in a loss of independence and the need for ongoing care. Early AD discovery is crucial because it allows for prompt treatment, which may enhance the lives of those afflicted and slow the disease's progression. There is presently no known cure for AD, and early diagnosis is still difficult because clinical symptoms can differ from person to person [1].

Machine learning, or ML, has emerged as a viable method for detecting AD in recent years due to its ability to analyze complex patient data and identify trends that may indicate the existence and progression of the illness. Using machine learning (ML) for AD diagnosis could be helpful to doctors by providing additional insights based on patterns that are hard to find using traditional techniques [2]. Several supervised machine learning techniques were employed in this work, including LR, RF, DT, Gaussian NB, AdaBoost Classifier,

Corresponding Author: aahafeez.scis@o6u.edu.eg
https://doi.org/10.61356/j.scin.2024.1517

and Gradient Boosting Classifier. Additionally, we propose and evaluate hybrid machine learning models that integrate two or three classifiers for the diagnosis of AD. Measures such as F1-score, recall, accuracy, and precision were employed to assess how effectively these models predicted the diagnosis of AD based on a dataset that included a variety of patient data. This dataset contains lifestyle characteristics, demographic information, medical history, clinical measurements, cognitive and functional assessments, and a range of AD-related symptoms.

It includes information on patients aged 60 to 90 years, as well as demographics such as age, sex, ethnicity, and educational attainment. Body mass index, smoking status, and level of physical activity are examples of lifestyle factors. Medical history includes familial illnesses such as Alzheimer's disease, diabetes, depression, high blood pressure, and cardiovascular disease. The MMSE score, systolic and diastolic blood pressure, and cholesterol level, when combined with clinical data, provide cognitive diagnostics that provide insight into each patient's physical and mental health [2].

2 | Methodology

This section outlines the methodology and analysis conducted in the current research as shown in Figure 1.

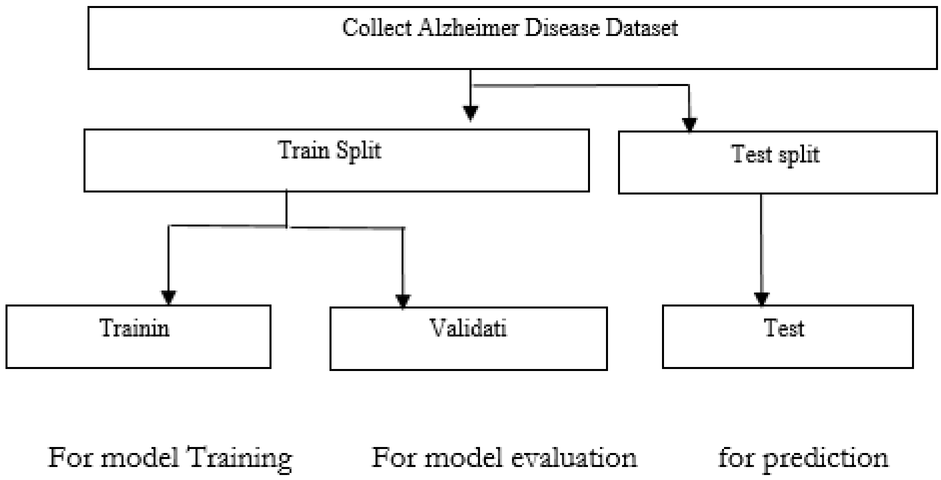

Figure 1. Process Alzheimer disease prediction [4].

This section describes the steps involved in making an Alzheimer's disease prediction, from gathering the data needed for the model to identifying the key characteristics that influence the prediction process. Data separation and preprocessing are then covered. The database is managed and separated into training and testing sets in this step so that different methods may be used for the trained dataset and the results can be seen by assessing the accuracy using the testing data. Five distinct machine learning methods were applied to this dataset: AdaBoost Classifier, Random Forest, Gaussian NB, Decision Tree, Gradient Boosting Classifier, and Logistic Regression.

2.1 | Dataset and Attribute

This section outlines the methodology and analysis conducted in the current research.

AD Dataset from Kaggle

Dataset link: https://www.kaggle.com/datasets/rabieelkharoua/alzheimers-disease-dataset

This dataset is specifically designed to cater to comprehensive coverage in terms of wide spectrums of patient characteristics, which are crucially required in analysis and the forecast of outcomes relative to Alzheimer's disease. Among other important dimensions, the features involved in this dataset reflect demographic data concerning patients, lifestyle, medical history, clinical measurements, cognitive and functional assessments, and observed symptoms. The dataset used for this analysis is split into an 80:20 training and testing split, respectively. Each attribute is explained in detail below.

2.1.1 | Demographic Details

This section outlines the methodology and analysis conducted in the current research.

Table 1 shows the patients' demographic and identification data with unique IDs, age, and gender. It further encodes ethnicity into four categories and education levels from None to Higher Education. The features mentioned above are crucial when analyzing the characteristics of the patients on a systematic basis.

Table 1. Demographic details of Dataset.

Attribute	Description
Patient ID	A unique identifier was assigned to each patient, ranging from 4751 to 6900.
Age	The age of the patients ranges from 60 to 90 years
Gender	Binary indicator: 0 = Male, 1 = Female.
Ethnicity	Encodes patient ethnicity: 0 = Caucasian, 1 = African American, 2 = Asian, 3 = Other
Education Level	Encoded as: 0 = None, 1 = High School, 2 = Bachelor's, 3 = Higher Education.

2.1.2 | Lifestyle Factors

Table 2 summarizes several important lifestyle factors affecting patient health: BMI, smoking history, and weekly alcohol intake. It further considers physical activity, diet quality, and sleep quality to provide insight into daily habits that influence overall health.

Table 2. Lifestyle Factors details of the dataset.

Attribute	Description
BMI	Body Mass Index, ranging from 15 to 40.
Smoking	History of smoking: 0 = No, 1 = Yes.
Alcohol Consumption	Weekly alcohol intake (0 to 20 units).
Physical Activity	Weekly hours of physical activity (0 to 10 hours).
Diet Quality	Scores from 0 to 10, reflect diet quality.
Sleep Quality	Score from 4 to 10, assess sleep quality.

2.1.3 | Medical History

Table 3 highlights some fundamental medical history that could impact the current health condition of a patient. This data includes a family medical history of Alzheimer's, cardiovascular disease, diabetes, depression, head injury, and hypertension, all of which are binary. Thus, this kind of feature provides valuable information on the risk factors associated with different conditions.

Table 3. Medical History details of the dataset.

Attribute	Description
Family History of Alzheimer's	0 = No, 1 = Yes.
Cardiovascular Disease	0 = No, 1 = Yes.
Diabetes	0 = No, 1 = Yes.
Depression	0 = No, 1 = Yes.
Head Injury	0 = No, 1 = Yes.
Hypertension	0 = No, 1 = Yes.

2.1.4 | Clinical Measurements

Table 4 presents selected clinical measures relevant to cardiovascular and metabolic health. Measures include blood pressure and cholesterol levels: total cholesterol, LDL, HDL, and triglycerides. These indicators will provide a general view of the state of the heart and its risks.

Table 4. Clinical Measurements details of the dataset.

Attribute	Description
Systolic BP	Systolic blood pressure in mmHg, ranging from 90 to 180.
Diastolic BP	Diastolic blood pressure in mmHg, ranging from 60 to 120
Cholesterol Total	Total cholesterol levels in mg/dL, ranging from 150 to 300.
Cholesterol LDL	Low-density lipoprotein cholesterol levels in mg/dL, from 50 to 200.
Cholesterol HDL	High-density lipoprotein cholesterol levels in mg/dL, from 20 to 100.
Cholesterol Triglycerides	Triglyceride levels in mg/dL, ranging from 50 to 400.

2.1.5 | Cognitive and Functional Assessments

Table 5 Cognitive and functional health measures, including the MMSE score, with lower scores indicating cognitive impairment, and a functional assessment score measuring independence in daily functioning. This is further supported by other factors such as memory complaints, behavioral problems, and ADL scores, to present a fuller picture of the patient's state of mind and functionality.

Table 5. Cognitive and Functional Assessments Details of the dataset.

Attribute	Description
MMSE (Mini-Mental State Examination)	Scores from 0 to 30, with lower scores indicating cognitive impairment.
Functional Assessment	Scores ranged from 0 to 10, with lower scores indicating greater functional impairment.
Memory Complaints	0 = No, 1 = Yes, if the patient has reported memory issues.
Behavioral Problems	0 = No, 1 = Yes, indicating the presence of behavioral issues.
ADL (Activities of Daily Living)	Scores from 0 to 10, assessing functional independence.

2.1.6 | Symptoms

Table 6 shows the symptoms of cognitive and behavioral health. It consists of binary representations of symptoms such as confusion, disorientation, personality changes, inability to complete tasks, and forgetfulness. These features can show the presence and intensity of the cognitive decline or change in the patient's behavior.

Table 6. Symptoms details of the dataset.

Attribute	Description
Confusion	0 = No, 1 = Yes, reports the presence of confusion.
Disorientation	0 = No, 1 = Yes, indicates episodes of disorientation.
Personality Changes	0 = No, 1 = Yes, presence of personality shifts or mood changes.
Difficulty Completing Tasks	0 = No, 1 = Yes, difficulty with task completion.
Forgetfulness	0 = No, 1 = Yes, presence of forgetfulness or memory loss.

2.1.7 | Diagnosis Information

Table 7 presents the final diagnosis status of the patients who have been categorized based on the presence and development of Alzheimer's AD. It will also be useful for deciding on the severity and stage of the disease for appropriate medical management.

Table 7. Diagnosis Information of Dataset.

Attribute	Description
Diagnosis	Final diagnosis status, classifying patients based on the presence and progression of Alzheimer's disease.

2.2 | Pre-processing of Data

Data preprocessing is one of the most crucial features in the preparation of the dataset for machine learning algorithms, as it can ensure that all the variables will be on the same scale and optimally suited for the model. In this paper, the definition of preprocessing was narrowed down to the structural treatment of the two numerical features using two techniques, namely, Standard Scaling and Min-Max Scaling. This was important to enable the variables to be within the same range and thus allow higher efficiency and steadiness of the algorithms.

2.2.1 | Standard Scaling

Standard Scaling normalizes the data so that each variable distribution takes a mean equal to zero and a standard deviation equal to one. This is a common method of normalization, making sure all variables become standardized and therefore appropriate for algorithms that are sensitive to the magnitude of data values, such as Logistic Regression and Support Vector Machines.

The formula for Standard Scaling is:

$$z = \frac{x-\mu}{\sigma} \tag{1}$$

Where x Is the original data point, μ Is the mean, and μ Is the standard deviation of the feature.

The following features were standardized using the Standard Scaler from sklearn.

2.2.2 | Min-Max Scaling

Min-max scaling was used to rescale the features into a common range. It rescales the data into a fixed range, usually between 0 and 1, thereby making it suitable for algorithms whose input needs to be bound, such as Neural Networks.

The formula for Min-Max Scaling is:

$$x' = \frac{x-\min(x)}{\max(x)-\min(x)} \tag{2}$$

Where x Is the original value, and $\min(x)$and $\max(x)$ EndThese are the minimum and maximum values of the feature.

The same set of features used for Standard Scaling was also rescaled using the Min-Max Scaler from sklearn. Preprocessing. The rescaled features had values confined to the range [0, 1].

3 | Machine Learning Algorithms

3.1 | Logistic Regression (LR)

Logistic regression is a guided machine learning technique by labels, which allows us to do binary classification tasks by predicting the probability of an outcome or an event. The model gives two outcomes: Yes→one or no →zero. In our case, it tells us whether the patient has the disease or not [6].

3.2 | Gaussian NB

Gaussian NB is one of the techniques that classify new data based on the relation between inputs and outputs. It is used in classification tasks where the goal is to predict the category or class of given data based on the training data. In our case, the model starts to learn from the training data to see how often unique features like age or cholesterol level will affect the outcome of whether a person has heart disease or not. Based on this information, when the model gets a new record from the testing set, the Naive Bayes algorithm checks the new patient's age, cholesterol level, and blood pressure and sees how similar these features are to those of patients who had heart disease or did not [7].

3.3 | Decision Tree (DT)

It is like the previous algorithm, but instead of combining branches, it combines multiple decision trees for the result. Each tree in the forest makes its prediction, and then the Random Forest algorithm combines these predictions to get the result [8].

3.4 | Logistic Regression (LR)

It is a supervised learning algorithm. The main process is splitting the features into subsets that look like tree structures; every node shows a feature, and every branch shows a decision rule. The decision tree learns how to make decisions or predictions based on this training data, and it then uses this information to predict the labels of new, unseen data [9].

3.5 | AdaBoost Classifier

The AdaBoost Classifier (Adaptive Boosting) is an ensemble learning algorithm that combines multiple weak classifiers to form a strong classifier. By iteratively adjusting the weights of misclassified samples, AdaBoost emphasizes harder-to-classify instances, thereby enhancing overall model accuracy. It is known for its simplicity and effectiveness, particularly in binary classification tasks. AdaBoost is often applied in medical and healthcare applications due to its interpretability and capacity to improve prediction accuracy without overfitting [10].

4 | Experimental Results

4.1 | Model Performance

Our model link: https://github.com/raghadahmed195/Alzhymer-Model

Table 8. Result of models.

Model	Accuracy	Precision	Recall	F1-Score
Gaussian NB	83.25%	83.06%	83.25%	83.09%
Decision Tree	90.69%	90.64%	90.69%	92.41%
Random Forest	92.55%	92.82%	92.55%	92.41%
Logistic Regression	83.02%	82.81%	83.02%	82.80%
AdaBoost	91.76%	91.11%	91.16%	91.12%

From Figure 2 and Table 8, we can see that Random Forest has the highest accuracy and classification report, and logistic regression has the lowest accuracy and classification report among all five models that perform.

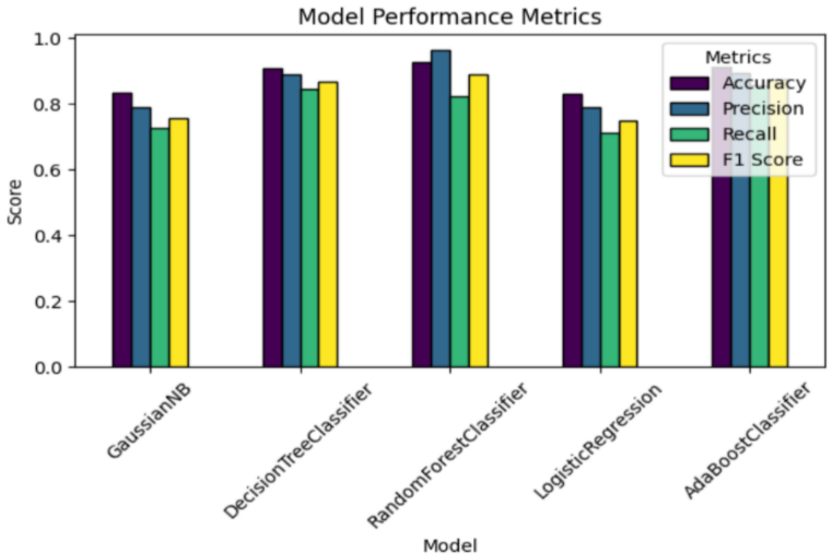

Figure 2. Learning Models v/s Accuracy, Recall, and F1 Score [11].

Table 9. Result of Hybrid models (2 models).

Model	Accuracy	Precision	Recall	F1-score
Gaussian NB +Random Forest	89.3%	89.25%	89.3%	89.26%
Decision Tree +Random Forest	90.69%	90.64%	90.69%	90.63%
Random Forest +Adaboost	92.55%	92.82%	92.55%	92.41%
AdaBoost +Decision tree	91.16%	91.7%	91.16%	90.9%

From Table 9, we can see that when Random Forest and Adaboost perform together, they achieve the highest accuracy and classification report.

Table 10. Result of Hybrid models (3 models).

Model	Accuracy	Precision	Recall	F1-score
Gaussian NB +Random Forest +logistic Regression	88.37%	88.28%	88.37%	88.27%
Decision Tree +AdaBoost +logistic Regression	95.46%	91.59%	91.62%	91.57%
Random Forest +Adaboost +Gaussian NB	89.76%	89.72%	89.76%	89.73%

From Table 10, we can see that when Decision Tree, AdaBoost, and logistic Regression perform together, they achieve the highest accuracy and classification report [13].

Table 11. Result of first reference paper [3].

Model	Accuracy	Recall
Logistic Regression	78.95%	75%
Support vector machine	81.58%	70%
Decision Tree	81.58%	65%
Random Forest	84.21%	80%
Adaboost	84.21%	65%

From Table 11, we can see that the Random Forest model achieves the highest accuracy at 84.21% and a recall of 80%, indicating superior performance in correctly identifying positive cases of AD. Support Vector Machine and Decision Tree models also show competitive accuracy but have lower recall rates.

Table 12. Result of the second reference paper [4].

Model	Accuracy	Precision	Recall	F1-score
Decision Tree	80.46%	80%	79%	78%
Random Forest	86.92%	85%	81%	80%
Support vector machine	81.67%	77%	70%	79%
XG Boost	85.92%	85%	83 %	85%
Voting classifier	85.12%	83%	83%	85%

From Table 12, we can see that the Random Forest model shows the highest accuracy at 86.92%, with a precision of 85% and a recall of 81%, indicating strong overall performance. XG Boost also performs well, achieving 85.92% accuracy with a total of 85% and a recall of eighty-three.

Table 13. Result of the third reference paper [5].

Model	Accuracy
K-nearest neighbor (KNN)	73.10%
Decision tree	76.43%
Rule induction	92.47%
Naïve Bayes	79.44%
Generalized linear model	92.75%
Deep learning	78.79%

From Table 13, we can see that. The Rule induction model achieves the highest accuracy at 92.47%, indicating it is the most accurate among the listed models.

4.2 | Comparison

In the first reference paper, they used 5 Machine learning models: Logistic regression, Support vector machine (SVM), Decision Tree, Random Forest, and Adaboost. They achieved the highest accuracy using random forest and adaboost, which is 84.21%.

In the second reference paper, they used 5 Machine learning models: Support vector machine (SVM), Decision Tree, Random Forest, XG Boost, and Voting classifier. They achieved the highest accuracy using Random Forest, which is 86.92%.

In the third reference paper, they used 6 Machine learning models: K-nearest neighbor (KNN), Decision Tree Rule induction, Naïve Bayes, generalized linear model, and Deep learning. They achieved the highest accuracy using the Generalized linear model, which is 92.75%.

We used five models of Gaussian NB, decision tree, Random Forest, logistic regression, and Adaboost. Our model achieved the highest accuracy using Random Forest with an accuracy of 92.55%, which is higher than the reference paper, and our model achieved higher accuracy using logistic regression, Adaboost, and Decision Tree than the reference paper [12].

5 | Results

These experiments utilize five machine learning models, including Gaussian Naive Bayes, Decision Tree, Random Forest, Logistic Regression, and AdaBoost, and two sets of hybrid models: combinations of two and three classifiers for the prediction performance based on the applied dataset. All the evaluation metricsincluding accuracy, precision, recall, and F1-score- were used to assess the effectiveness of the models. The Random Forest classifier had the maximum accuracy at 92.55%, precision of 92.82%, recall of 92.55%, and an F1-score of 92.41%, proving to be better than other algorithms. In the hybrid of combining two algorithms, Random Forest and Adaboost achieved maximum accuracy at 92.55%, precision of 92.82%, recall of 92.55%, and an F1-score of 92.41%. In the hybrid of combining three algorithms, Decision Tree, AdaBoost, and logistic Regression achieved maximum accuracy at 95.46%, precision of 91.59%, recall of 91.62%, and an F1-score of 91.57%.

6 | Conclusion

This work, therefore, underlines the capability of machine learning algorithms in predicting AD through an analysis of unique features of patients. Among the tested models, Random Forest and AdaBoost showed the best performance; this would prove that ensemble methods do well because of complexities in the dataset, resulting from the possibility of capturing non-linear relationships. These findings then strengthen the viability of the application of machine learning in the diagnosis of AD and provide useful support for early intervention strategies. Future studies can explore other advanced models, including deep learning, and make use of more substantial and varied datasets to validate these findings further and enhance predictive performance.

Acknowledgments

The author is grateful to the editorial and reviewers, as well as the correspondent author, who offered assistance in the form of advice, assessment, and checking during the study period.

Author Contribution

All authors contributed equally to this work.

Funding

This research has no funding source.

Data Availability

The datasets generated during and/or analyzed during the current study are not publicly available due to the privacy-preserving nature of the data but are available from the corresponding author upon reasonable request.

Conflicts of Interest

The authors declare that there is no conflict of interest in the research.

Ethical Approval

This article does not contain any studies with human participants or animals performed by any of the authors.

References

[1] Alzheimer's Association. (2023). "Alzheimer's Disease Facts and Figures." Alzheimer's & Dementia, 19(3), 1-76. Available at: https://www.alz.org/alzheimers-dementia/facts-figures.

[2] Liu, Y., Chen, P. C., & Krause, J. (2020). "How to Read Machine Learning Papers: A Primer." JAMA, 324(10), 1015-1017. Available at: https://doi.org/10.1001/jama.2020.11728

[3] A. Rajak and A. Kumar Shrivastava, "Diagnosis of Alzheimer Disease using Machine Learning Approaches," Int. J. Adv. Sci. Technol., vol. 29, no. 04, pp. 7062–7073, 2020, [Online]. Available: https://www.researchgate.net/publication/342847161

[4] C. Kavitha, V. Mani, S. R. Srividhya, O. I. Khalaf, and C. A. Tavera Romero, "Early-Stage Alzheimer's Disease Prediction Using Machine Learning Models," Front. Public Heal., vol. 10, no. March, pp. 1–13, 2022, Doi: 10.3389/fpubh.2022.853294.

[5] M. Shahbaz, S. Ali, A. Guergachi, A. Niazi, and A. Umer, "Classification of Alzheimer's disease using machine learning techniques," DATA 2019 - Proc. 8th Int. Conf. Data Sci. Technol. Appl., no. Data, pp. 296–303, 2019, Doi: 10.5220/0007949902960303.

[6] James, G., Witten, D., Hastie, T., & Tibshirani, R. (2013). "An Introduction to Statistical Learning: With Applications in R." Springer.

[7] Rish, I. (2001). "An empirical study of the naive Bayes classifier." IJCAI 2001 Workshop on Empirical Methods in AI, 41-46. [4] Escudero, J., et al. (2013). "Machine learning-based method for personalized and cost-effective detection of Alzheimer's disease." IEEE Transactions on Biomedical Engineering, 60(1), 164-168.

[8] Breiman, L. (2001). "Random Forests." Machine Learning, 45(1), 5-32. https://doi.org/10.1023/A:1010933404324

[9] Quinlan, J. R. (1986). "Induction of Decision Trees." Machine Learning, 1(1), 81-106. https://doi.org/10.1007/BF00116251

[10] Freund, Y., & Schapire, R. E. (1997). "A decision-theoretic generalization of on-line learning and an application to boosting." Journal of Computer and System Sciences, 55(1), 119-139. https://doi.org/10.1006/jcss.1997.1504

[11] Powers, D. M. (2011). "Evaluation: From Precision, Recall and F-Measure to ROC, Informed Ness, Markedness and Correlation." Journal of Machine Learning Technologies, 2(1), 37-63.

[12] Sokolova, M., & Lapalme, G. (2009). "A systematic analysis of performance measures for classification tasks." Information Processing & Management, 45(4), 427-437. https://doi.org/10.1016/j.ipm.2009.03.002

[13] Fawcett, T. (2006). "An introduction to ROC analysis." Pattern Recognition Letters, 27(8), 861-874. https://doi.org/10.1016/j.patrec.2005.10.010