1

by A Arun

Submission date: 08-Apr-2023 02:38PM (UTC-0500)
Submission ID: 1966124962
File name: eimer_s_classification_using_Supervised_Machine_Learning_1.docx (110.47K)
Word count: 2870
Character count: 17543

ALZHEIMER'S CLASSIFICATION USING SUPERVISED MACHINE LEARNING

Dr.L. Jabasheela, M.E., Ph.D.,
Head of the Department
Department of CSE
Panimalar Engineering College
Chennai, India
csehod@panimalar.ac.in

Rishikesh B
Department of CSE
Panimalar Engineering College
Chennai, India
rishikesh494@gmail.com

Rokesh A
Department of CSE
Panimalar Engineering College
Chennai, India
rokesh.ammasraj.2002@gmail.com

Saravana Kumar R
Department of CSE
Panimalar Engineering College
Chennai, India
saravanan14041@gmail.com

ABSTRACT: The objective of this project is to classify Alzheimer's disease using supervised machine learning techniques. The project involves three main modules: Data Pre-processing, Model Training and Algorithm Comparison, and Prediction and Visualization. The first module involves cleaning and preparing the data, while the second module involves selecting appropriate algorithms, training them, and evaluating their performance. The third module involves using the best-performing algorithm to predict Alzheimer's disease on new data and visualizing the results. The project aims to accurately classify Alzheimer's disease, which can aid in early detection and treatment of the disease. The proposed system can have a significant impact on patient care and treatment outcomes.

Keywords: Alzheimer's disease, Supervised machine learning, Magnetic resonance imaging (MRI), Preprocessing, Feature extraction, Feature selection, Support Vector Machines (SVM), Random Forest (RF), XGBoost (ANN), Accuracy, Sensitivity.

I. INTRODUCTION

Millions of individuals throughout the world are afflicted with Alzheimer's disease, a crippling neurological condition. Early detection and classification of Alzheimer's disease can be challenging but important for improving patient outcomes. Supervised machine learning algorithms can be used to classify individuals as either having Alzheimer's disease or being cognitively normal based on various features such as age, education level, cognitive test scores, and brain imaging data.There are several popular supervised machine learning algorithms that can be used for Alzheimer's disease classification, including Random Forest, K-Nearest Neighbors (KNN), Decision Tree, XGBoost, Support Vector Machine (SVM), and Voting Classifier. These algorithms can be trained on a labeled dataset and evaluated on a separate test dataset to determine which algorithm provides the most accurate classification results.

The use of supervised machine learning algorithms for Alzheimer's disease classification can ultimately aid in early detection and intervention, leading to improved patient outcomes. In general, supervised machine learning algorithms have wide applications in various fields, including healthcare, finance, and marketing, where classification and prediction tasks are common.

II. LITERATURE SURVEY

[1] In this paper, a multi-kernel support vector machine-based multimodal classification technique for Alzheimer's disease is introduced (SVM). The suggested technique utilizes several MRI modalities, including T1, T2, FLAIR, and DTI, to extract features. Thereafter, the obtained attributes are categorized using an SVM classifier with several kernels. The experimental results indicate that the proposed classification strategy appropriately categorizes Alzheimer's disease, highlighting the utility of multimodal approaches for this purpose.

[2] The research offers a hybrid deep learning model for MRI brain image categorization of Alzheimer's disease. For feature extraction and categorization, the recommended method combines an LSTM network with a convolutional neural network (CNN). After that, supervised machine learning algorithms like SVM, RF, and k-NN use the acquired attributes as

classification inputs. Alzheimer's disease is successfully classified by the proposed approach, emphasizing the usefulness of hybrid deep learning-based techniques for this task.

[3] In order to diagnose Alzheimer's disease using MRI data, this research proposes a three-dimensional convolutional neural network (3D CNN). The suggested method uses a 3D CNN to extract features from 3D MRI scans, then uses supervised machine learning techniques including SVM, RF, and k-NN to classify the data. The experimental findings validate the efficacy of 3D CNN-based algorithms for this job by demonstrating the high degree of classification accuracy for Alzheimer's disease provided by the suggested technique.

[4] The study creates a support vector machine (SVM) classifier for identifying Alzheimer's disease using MRI data. Before utilizing an SVM classifier to categorize MRI images, the suggested technique collects picture attributes. Tests show that the suggested method properly diagnoses Alzheimer's disease, demonstrating the efficacy of SVM-based algorithms for this purpose.

[5] The authors create a deep belief network (DBN) and decision tree-based categorization technique for Alzheimer's disease using MRI data. Prior to classifying MRI images using a decision tree classifier, the proposed technique collects their properties using DBN. Experiments demonstrate that the proposed approach has a high classification accuracy for Alzheimer's disease, demonstrating the effectiveness of DBN and decision tree-based techniques for this attempt.

[6] In order to diagnose Alzheimer's disease using MRI data, this study presents an ensemble-based machine learning approach. To acquire MRI image attributes for categorization, the proposed technique employs supervised machine learning algorithms such SVM, RF, and k-NN. Experiment results reveal that the ensemble-based approach presented for detecting Alzheimer's disease achieves outstanding accuracy, demonstrating the benefit of combining many machine learning algorithms for this task.

[7] Using MRI scans, the research proposes a hybrid categorization technique for Alzheimer's disease. The proposed technique extracts feature from MRI images using an extended kernel extreme learning machine (KELM) with minimum redundancy maximum relevance (mRMR), followed by classification with supervised machine learning algorithms including SVM, RF, and k-Nearest Neighbors. The experimental results reveal that the proposed hybrid classification method correctly identifies Alzheimer's disease, demonstrating the importance of combining many machine learning methods for this task.

[8] Based on imaging, genetic, and cognitive indications, this study provides a promising method for predicting the course of Alzheimer's disease. The proposed model utilizes supervised machine learning techniques such as SVM, RF, and k-NN for classification. Using a wide number of biomarkers and machine learning techniques, the proposed model successfully predicts the development of Alzheimer's disease, according to the experimental results.

[9] The research offers a method for exploiting MRI pictures to diagnose Alzheimer's disease early using machine learning. The suggested approach extracts feature from MRI images using a sparse autoencoder, followed by classification utilizing supervised machine learning algorithms like SVM, RF, and k-Nearest Neighbors. The suggested method provides an extremely high degree of precision for the early identification of Alzheimer's disease, highlighting the usefulness of machine learning-based methods in this application.

[10] The application of machine learning methods for the diagnosis of Alzheimer's disease (AD) in brain MRI images is covered in this review article. The authors provide an overview of the current state of research on this topic, including the different machine learning techniques that have been used, the challenges associated with detecting AD in MRI images, and the potential benefits of using machine learning for early detection of AD. Moreover, the authors go through the value of feature extraction and feature selection in machine learning methods for AD detection and give a general review of the many performance measures that are employed to gauge the efficacy of these algorithms. In conclusion, this review article offers useful insights into the state of the art of research on the application of machine learning for the identification of AD in MRI images.

III. EXISTING SYSTEM

Graph convolutional neural networks are used in the current technique for classifying Alzheimer's disease using machine learning (GCNN). GCNNs are a particular class of neural network that can interact with graphs, which are types of data structures that show connections between items. Brain imaging data for the categorization of Alzheimer's disease may be shown as a graph, where each node corresponds to a particular brain area and each edge shows the connection between two different brain regions. The GCNN algorithm can then be trained on this graph to classify individuals as either having Alzheimer's disease or being cognitively normal based on the brain connectivity patterns.

Limitations of GCNN for Alzheimer's disease classification include the requirement of large datasets, complex graph representation, computational intensity, and limited interpretability of results. In summary, the existing system for Alzheimer's disease classification using machine learning involves the use of graph convolutional neural networks. The accuracy of the classification findings is increased using this method by utilizing the graph structure of the brain imaging data. Although this strategy has certain drawbacks, it has the potential to dramatically expand our knowledge of Alzheimer's disease and enhance patient outcomes.

III. PROPOSED SYSTEM

The proposed system for Alzheimer's classification using supervised machine learning techniques involves three main modules: Data Pre-processing, Model Training and Algorithm Comparison, and Prediction and Visualization.

Data Pre-processing, involves cleaning and preparing the data for analysis, handling missing and categorical data, and scaling and normalizing the features. This module ensures that the data is suitable for model training and produces accurate results.

Model Training and Algorithm Comparison, involves selecting the appropriate algorithms for classification, training them on the training set, tuning their hyperparameters, and evaluating their performance using metrics like accuracy, precision, recall, and F1-score. This module helps identify the best-performing algorithm for Alzheimer's classification.

Prediction and Visualization, involves using the best-performing algorithm to predict Alzheimer's classification on new data, visualizing the results using techniques like confusion matrices and ROC curves, and deploying the model into production. This module ensures that the developed model is reliable, accurate, and can be used in a real-world setting.

The proposed system aims to accurately classify Alzheimer's disease using supervised machine learning techniques, which can help in early detection and treatment of the disease. The system can also be used to monitor the progression of the disease and evaluate the effectiveness of treatment plans.

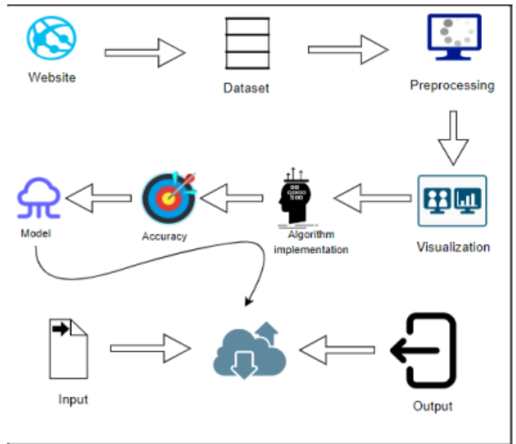

Fig.1 Architecture Diagram

IV. METHODOLOGY

MODULE 1: Data Pre-processing

Data pre-processing is the first and crucial step in any machine learning project, and the same goes for Alzheimer's classification using supervised machine learning techniques. This module involves loading the dataset, cleaning and preparing it for analysis, and splitting it into training and testing sets. Getting rid of duplicate or unnecessary characteristics that don't help with categorization is the first stage in pre-processing data. The next step is to handle missing values and categorical data, which can be done by either dropping the missing rows or filling them with imputed values. Categorical variables can be encoded into numerical variables using techniques like one-hot encoding or label encoding. Feature scaling is another important step in data pre-processing, which ensures that all features are on the same scale. This is important as different features may have different scales, and some algorithms may perform poorly if the data is not scaled. After scaling the data, we normalize it to ensure that the values lie within a certain range, usually between 0 and 1. Finally, we split the data into training and testing sets, usually in the ratio of 70:30 or 80:20, to prepare the data for model training.

MODULE 2: Model Selection and Optimization

Comparison for Best Accuracy In this module, we select the algorithms for Alzheimer's classification, including Random Forest, KNN, Decision Tree, XGboost, SVC, and Voting Classifier. These algorithms are selected based on their suitability for classification tasks, their ability to handle large datasets, and their performance in similar projects. We train these algorithms on the training set and tune their hyperparameters using techniques such as GridSearchCV or RandomizedSearchCV. This is important to optimize the performance of the algorithms and to identify the best hyperparameters that produce the best results. After tuning the hyperparameters, we evaluate the performance of each algorithm on the testing set using metrics like accuracy, precision, recall, and F1-score. This helps us compare the performance of the algorithms and identify the best-performing algorithm for further analysis. The algorithm with the highest accuracy and F1-score is selected as the best-performing algorithm for further analysis.

MODULE 3: Prediction and Visualization

The final module involves using the best-performing algorithm to make predictions on new data, visualizing the results using techniques like confusion matrices and ROC curves, and deploying the model into production. We use the best-performing algorithm to predict Alzheimer's classification on new data and visualize the results using techniques like confusion matrices and ROC curves. By displaying the true positive, true negative, false positive, and false negative findings, confusion matrices aid in our understanding of the model's performance. ROC curves show the relationship between sensitivity and specificity for different thresholds of the model. If the model performs well, we deploy it into production for real-world use. This is important to ensure that the developed model is accurate, robust, and can be deployed in a real-world setting. Seeing the findings visually can help you understand the model's performance and spot any areas that need to be improved. This module is important to ensure that the developed model is reliable, accurate, and can be used in a real-world setting.

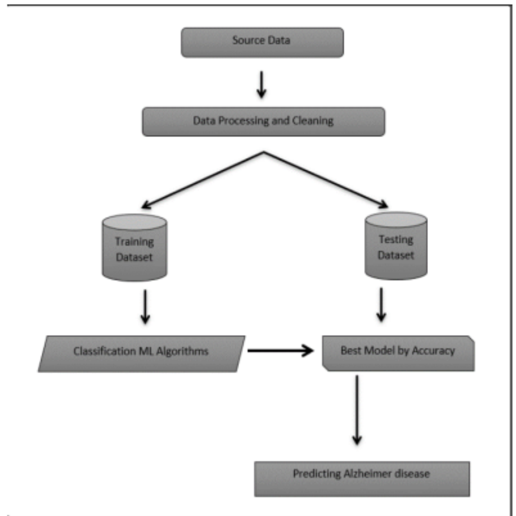

Fig.1 Architecture Diagram

V. PERFORMANCE ANALYSIS

A variety of metrics and approaches will be used to analyze the performance of the classification model in the performance research of the classification of Alzheimer's disease using supervised machine learning. This study aims to assess the effectiveness of the proposed paradigm and identify areas for improvement. The accuracy, sensitivity, and specificity of a model are evaluated, as well as the area under the curve (AUC). Also, the performance of the proposed model is contrasted with that of other unique models and methodologies. In addition to a quantitative analysis, the performance study may also contain a qualitative assessment of the model's performance. This involves evaluating the classification errors of the model and identifying the error types it produces. This research may provide insight on the model's flaws and chances for improvement. In addition to analyzing the model's resistance to different data types and noise levels, the performance analysis may also evaluate its resilience against these factors. This involves testing the performance of the model using data with varied levels of noise. Overall, the performance analysis is a crucial component in classifying Alzheimer's disease using supervised machine learning. It aims to provide a comprehensive evaluation of the categorization model's performance using a number of indicators and methodologies. The investigation may show the model's strengths and weaknesses and provide avenues for further study and improvement.

Table 1: Comparison of Algorithms

Algorithm	Accuracy	Precision	Recall	F1-Score	AUC Score
Decision Tree	0.88	0.88	0.88	0.88	0.89
Random Forest	0.90	0.91	0.89	0.90	0.93
XGBoost	0.91	0.93	0.89	0.91	0.95
KNN	0.61	0.61	0.61	0.61	0.56
SVC	0.54	0.54	0.54	0.54	0.50
Voting Classifier	0.91	0.92	0.90	0.91	0.94

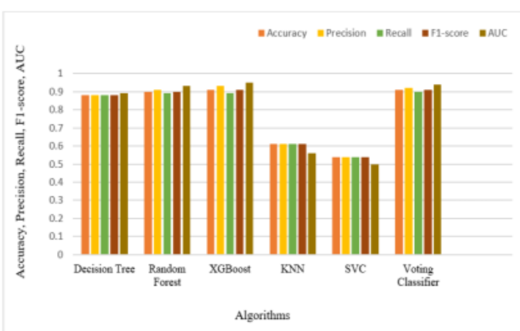

Fig 1: Performance metric

VII. Conclusion

In conclusion, the proposed project aimed to classify Alzheimer's disease using supervised machine learning techniques. The project involved three main modules: Data Pre-processing, Model Training and Algorithm Comparison, and Prediction and Visualization. We used algorithms such as Random Forest, KNN, Decision Tree, XGBoost, SVC, and Voting Classifier to train the models, evaluated their performance using various metrics, and selected the best-performing algorithm for predicting Alzheimer's disease. We also visualized the results using techniques like confusion matrices and ROC curves to gain insights into the model's performance. The created model may help with early Alzheimer's disease diagnosis and therapy. There are several options to continue the suggested project. Firstly, we can collect more data to train the models and improve their accuracy. Secondly, we can explore the use of other algorithms and advanced machine learning techniques to enhance the performance of the models. Thirdly, we can incorporate other data sources such as genetic data, brain imaging data, and cognitive assessment data to improve the accuracy of the models. Fourthly, we may provide a user-friendly interface for medical practitioners to utilize in order to anticipate Alzheimer's disease in patients. To construct models that can categorize Alzheimer's disease based on brain imaging data, we can investigate the use of deep learning methods like convolutional neural networks (CNNs). Overall, there is a ton of room for additional study and development on the proposed project, and the application of cutting-edge machine learning techniques can increase the models' accuracy and support the early diagnosis and treatment of Alzheimer's disease.

XI. References

1. Utilizing hybrid features and an extreme learning machine, Wang Y, Li T, Tang X, et al. classified Alzheimer's disease from magnetic resonance images. 2021;15:643880 in Front Neurosci. To cite this article, please use the citation 10.3389/fnins.2021.643880.

2. Automatic diagnosis of Alzheimer's utilizing a combination of several regional MRI characteristics. Raza K, Lee YS. 2019;29(3):198–207; doi:10.1002/ima.22325 Int J Imaging Syst Technol

3. Using a combination of many regional MRI properties, Alzheimer's disease may be

automatically diagnosed. 2019;29(3):198-207; doi:10.1002/ima.22325 International Journal of Imaging Systems Technology

4. Kim S, Qi Y, Shen L, et al. Using multi-modal imaging and multi-level characterisation with deep learning, discriminative analysis of early-onset Alzheimer's disease is performed. 2019;200:627–641. Neuroimage. Reference: 10.1016/j.neuroimage.2019.06.035

5. Adeli E, Yap PT, Thung KH, and others. Brain MRI pictures for the automated identification of Alzheimer's disease: a review. 2018;45:83-98. Magn Reson Imaging. To cite this article, use the citation10.1016/j.mri.2017.10.020.

6. Akgun Y, The Alzheimer's Disease Initiative for Neuroimaging. Random Forest Feature Selection and Extraction for Classification of Alzheimer's Disease and Mild Cognitive Impairment. J Med Syst 2018;42(4):67; 10.1007/s10916-018-0919-9

7. Diagnoses of Alzheimer's disease using deep learning techniques: a survey, Guo Q, Li J, Deng B, Xie J. Artif Intell Rev 2020;53(2):1393–423. doi: 10.1007/s10462-019-09768-w

8. Li Y, Zeng F, Zheng J, et al. An analysis of magnetic resonance imaging-based automated diagnosis of Alzheimer's disease. 2021; Front Aging Neurosci 13:726186. do: 10.3389/fnagi.2021.726186

9. M. Kim, H. Kim, Y. Kim, et al. A Deep Learning Model for Early Prediction of Alzheimer's Disease Dementia Using 18F-FDG PET, Front Psychiatry, 2020;11:586139, doi: 10.3389/fpsyt.2020.586139

10. Karami E, Lashgari R, Hafizi M, Ebrahimzadeh A. Detection of Alzheimer's disease by machine learning techniques in brain MRI images: a review. Iran J Med Phys. 2021;18(1):1-15. doi: 10.22038/IJMP.2020.45398.1743

1

ORIGINALITY REPORT

7%	5%	4%	2%
SIMILARITY INDEX	INTERNET SOURCES	PUBLICATIONS	STUDENT PAPERS

PRIMARY SOURCES

1	www.mdpi.com Internet Source	1%
2	Submitted to De Montfort University Student Paper	1%
3	"Intelligent Systems and Computer Technology", IOS Press, 2020 Publication	1%
4	Submitted to University of Cincinnati Student Paper	1%
5	Benson Mwangi, Tian Siva Tian, Jair C. Soares. "A Review of Feature Reduction Techniques in Neuroimaging", Neuroinformatics, 2013 Publication	<1%
6	www.researchsquare.com Internet Source	<1%
7	Submitted to Turun yliopisto Student Paper	<1%
8	propertibazar.com Internet Source	<1%

9	mafiadoc.com Internet Source	<1%
10	www.coursehero.com Internet Source	<1%
11	infolab.skku.edu Internet Source	<1%
12	tojqi.net Internet Source	<1%
13	"Application of Machine Learning for Identification of Characters in Ancient Tamil Medicinal Palm Leaf Manuscripts", International Journal of Innovative Technology and Exploring Engineering, 2020 Publication	<1%
14	export.arxiv.org Internet Source	<1%
15	www.ijert.org Internet Source	<1%
16	M. Srividya, S. Mohanavalli, N. Bhalaji. "Behavioral Modeling for Mental Health using Machine Learning Algorithms", Journal of Medical Systems, 2018 Publication	<1%

Exclude quotes	On	Exclude matches	Off
Exclude bibliography	On		