Communications on Applied Nonlinear Analysis
ISSN: 1074-133X
Vol 32 No.1s (2025)

A Comprehensive Evaluation of Deep Learning Architectures and Traditional Machine Learning Algorithms for Prognostic Modeling in Alzheimer's disease

R. Arumugam[1] and A. Murugan[2]

[1]Research Scholar, Department of Computer and Information Science, Annamalai University, Annamalainagar – 608 002, Tamil Nadu, India
[2]Assistant Professor, Department of Computer Science, Periyar Arts College, Cuddalore, (Deputed from Annamalai University, Annamalainagar) Tamil Nadu, India
Email: [1]arumugammca848@gmail.com,[2]drmuruganapcs@gmail.com

Article History:
Received: *10-11-2024*
Revised: *08-12-2024*
Accepted: *02-01-2025*

Abstract:
Alzheimer's disease (AD) poses a growing public health challenge, underscoring the need for accurate and early prognostic methods. This study compares traditional machine learning (ML) algorithms and contemporary deep learning (DL) models for predicting AD outcomes using the cross-sectional dataset. The dataset includes demographic (ID, gender, handedness, age, education, socioeconomic status), cognitive (MMSE, CDR), and structural (ETIV, NWBV, ASF) features. Standard ML techniques—Support Vector Machines (SVM), Random Forest (RF), and Gradient Boosting Machines (GBM)—were implemented with hyperparameters optimized via grid search and cross-validation. Concurrently, deep neural networks (DNNs) were constructed with varied architectures and refined using advanced strategies such as dynamic learning rate scheduling, dropout regularization, and enhanced Adam optimization to mitigate overfitting and improve training efficacy. Model performance was evaluated using accuracy, precision, recall, F1-score, and the area under the ROC curve (AUC-ROC). Results demonstrate that while traditional ML models offer competitive performance with lower computational overhead, well-tuned DL models deliver superior predictive accuracy and generalization on unseen data.

Keywords: Alzheimer's Disease Prediction, Machine Learning, Deep Learning, Hyperparameter Tuning, Deep Learning, and Optimization Techniques.
.

1. Introduction

Alzheimer's disease (AD) is a brain disorder that gradually damages memory, thinking skills, and behavior. It is the leading cause of dementia and poses a major health issue worldwide, especially as the elderly population grows. Since there is no known cure and only limited treatment options, finding ways to predict the disease early is very important. Early detection can help manage symptoms better and slow the progression. In recent years, artificial intelligence (AI) has played a growing role in healthcare, especially in identifying and predicting brain disorders like Alzheimer's. Machine learning (ML) and deep learning (DL) have shown great potential in analyzing medical data and providing more accurate predictions. These techniques are able to uncover hidden patterns in both clinical and imaging data, helping doctors make better decisions. Traditional ML methods like Support Vector Machines (SVM), Random Forest (RF), and Gradient Boosting Machines (GBM) have been successful in many medical applications. However, they often need manual work to choose and fine-tune features. On the other hand, deep learning models such as deep neural networks (DNNs) can automatically learn useful patterns from data, and when trained properly, they often give better results. This study compares traditional ML models and deep learning approaches for predicting Alzheimer's disease using the OASIS cross-sectional dataset. We applied hyperparameter tuning to improve ML models and used advanced training methods for deep learning models. Our goal is to find which method works best for accurate prediction. The rest of the paper is organized as follows: Section II reviews related studies using ML and DL in Alzheimer's diagnosis. Section III explains the methods used, including data preparation, model building, and evaluation. Section IV discusses the results. Section V concludes the study and outlines future work.

2. Literature Review

Many researchers have studied how machine learning (ML) and deep learning (DL) can help diagnose Alzheimer's Disease (AD), especially by using brain scan images like MRI, fMRI, and patient health records. This section summarizes 25 important studies in this field. Sarraf and Tofighi [1] created a CNN-based model called DeepAD, one of the earliest to use brain scans for detecting AD. Their model performed better than older ML methods, showing the power of DL for analyzing medical images. Other studies focused on the strengths of traditional ML. Ortiz-Sanz et al. [2] used Support Vector Machines and Random Forests along with brain scan features and memory scores. These models performed well, especially after being tuned with cross-validation. Likewise, Basaia et al. [3] trained deep neural networks on MRI data and successfully classified patients with AD, mild cognitive impairment (MCI), and healthy individuals.

Suk et al. [4] and Suk & Shen [19] tested stacked autoencoders to improve feature extraction from brain images. These DL methods worked better than manually selected features, making predictions more accurate. Jie et al. [5] used 3D convolutional networks to better detect changes in brain volume—a key sign of Alzheimer's. Zhang et al. [8] and Liu et al. [11] combined different types of imaging data, like MRI and fMRI, using graph-based and stacking techniques. Their results showed that combining data types improved the model's prediction power. Padilla et al. [9] also applied ML on the OASIS dataset and showed that selecting the right features improves accuracy. Farooq et al. [6] and Esmaeilzadeh et al. [10] worked on classifying AD, MCI, and healthy cases using CNNs and mixed deep learning models. They used techniques like dropout, batch normalization, and adaptive learning rates to avoid overfitting and improve training.

Zhang et al. [15] also used transfer learning to improve results while saving training time. From the ML side, decision trees and ensemble methods are still popular for their ease of understanding. Liu et al. [14] used decision trees to find early signs of memory loss, while Lin & Zhang [24] applied boosting and bagging methods, which gave good results with less computing power. Ahmed et al. [16] and Jain et al. [20] built full pipelines that included feature selection and data augmentation, which helped improve early diagnosis. Ravi and Karthik [23] compared deep and traditional models and found that while deep models are powerful, classical methods can also perform well with the right tuning. Overviews by Razzak et al. [25] and

Salvatore et al. [7] discussed challenges like limited data and model transparency, and recommended using explainable AI (XAI) for better trust in predictions. Bron et al. [21] created a standard dataset and benchmarks for AD detection, which many studies now follow for consistent evaluation. Brosch and Tam [17] tackled small dataset problems by using unsupervised learning and dimensionality reduction, helping models avoid overfitting. Lastly, Wang et al. [12] and Vieira et al. [13] looked at how well DL models work across different groups of people. They stressed the need for more diverse training data to make models useful in real-world settings. In summary, this literature review shows the shift from traditional ML approaches to modern DL models in Alzheimer's prediction. While deep learning often performs better, traditional methods are still useful—especially when using structured data and good tuning methods. This supports our study's goal to compare both approaches using the OASIS dataset.

The classification methods used include J48, Random Tree (RT), Decision Stump (DS), Logistic Model Tree (LMT), Hoeffding Tree (HT), Reduced Error Pruning (REP), and Random Forest (RF), and their accuracies are assessed [26] and the similar paper analysis using medical related research using same approaches using data mining and machine learning algorithms [27]. Ravishankar and Rajesh [28] studied the impact of selecting important variables from climate change datasets and how this affects prediction accuracy. They applied different data mining techniques along with machine learning models to understand climate patterns. In another related study, Ravishankar and Rajesh [29] extended their research using a global weather repository to predict climate change indicators more effectively. They used data mining tools combined with advanced machine learning methods to process large-scale environmental data. Ravishankar and Rajesh [30] carried out a detailed study on analyzing climate change datasets in relation to the Air Quality Index (AQI) using data mining and machine learning techniques. Their research focused on understanding how different environmental parameters contribute to AQI levels. By applying classification and regression models, they demonstrated how machine learning can accurately predict air quality trends. Santhoshkumar and Rajesh [31] explored how machine learning techniques can be used to analyze the connection between various types of energy usage and the Sustainable Development Goals (SDGs). Their study applied predictive modeling to identify how changes in energy consumption patterns affect progress toward SDG targets.

2.1. Dataset Description

The OASIS (Open Access Series of Imaging Studies) Cross-Sectional dataset is a publicly available resource designed for Alzheimer's Disease research. It comprises MRI data and associates clinical and demographic information from a cross-sectional cohort of participants aged 18 to 96 years. The dataset includes both cognitively normal individuals and those diagnosed with varying stages of Alzheimer's Disease [32].

2.1.1 Features in the Dataset:

Feature	Description
ID	Unique identifier for each participant
Age	Age of the subject in years
Sex	Gender of the subject (M or F)
Educ	Years of education

SES	Socioeconomic status (1 = high, 5 = low; may have missing values)
MMSE	Mini-Mental State Examination score (0–30)
CDR	Clinical Dementia Rating (0 = normal, 0.5 = mild, 1–3 = increasing severity)
eTIV	Estimated Total Intracranial Volume
nWBV	Normalized Whole Brain Volume
ASF	Atlas Scaling Factor
Group	Diagnosis group (Nondemented, Demented, Converted)

2.1.2 Dataset Table

Here is a sample version of the dataset showing the first five representative rows:

ID	Age	Sex	Educ	SES	MMSE	CDR	eTIV	nWBV	ASF	Group
OAS1_0001	74	M	12	3.0	28	0.0	1987	0.696	1.20	Nondemented
OAS1_0002	55	F	14	2.0	30	0.0	1739	0.736	1.00	Nondemented
OAS1_0003	73	M	12	3.0	26	0.5	1989	0.694	1.21	Demented
OAS1_0004	76	F	12	1.0	30	0.0	1964	0.736	1.18	Nondemented
OAS1_0005	88	F	12	NaN	25	1.0	1672	0.664	1.12	Demented

3. Background and Methodologies

Alzheimer’s Disease (AD) is a serious brain condition that gradually leads to memory problems, confusion, and poor decision-making. As the world’s population continues to age, the number of people affected by AD is expected to grow rapidly. This makes it more important than ever to find reliable ways to detect the disease early. While standard methods like brain scans and memory tests are useful, they can be expensive and time-consuming, making them hard to use for large groups of people. In recent years, artificial intelligence (AI) has provided new ways to help with medical diagnosis. Machine learning (ML) and deep learning (DL) are two types of AI that have shown promise in predicting diseases like AD. Traditional ML models such as Support Vector Machines (SVM), Random Forests (RF), and Gradient Boosting Machines (GBM) are good at making sense of structured data, especially when their settings are fine-tuned. However, they often require experts to pick the right features from the data.

Deep learning models, such as Deep Neural Networks (DNNs), can automatically learn important features from the data without much manual work. These models are better at handling complex data but need a lot of computing power and can sometimes overfit—especially if the dataset is small. To deal with these challenges, techniques like dropout, learning rate adjustment, and using advanced optimizers like Adam are used. This study compares both traditional ML and deep learning models for predicting Alzheimer’s Disease using the OASIS cross-sectional dataset. The aim is to find out which method gives the most accurate and reliable results by testing different models and comparing their performance.

3.1 Machine Learning Models

Three popular ML algorithms were used:

Step. 1 **Support Vector Machine (SVM)**:

- o Used an RBF kernel.

- Hyperparameters (C and gamma) were optimized using grid search and 5-fold cross-validation.

Step. 2 **Random Forest (RF)**:

- Tested different numbers of trees and depth levels to improve performance.

Step. 3 **Gradient Boosting Machine (GBM)**:

- Used learning rate tuning and early stopping to avoid overfitting.

3.2 Deep Learning Model

A deep neural network (DNN) was designed with the following setup:

Step. 1 **Structure**:

a. Input layer with the number of features in the dataset.
b. Two hidden layers with ReLU activation (64 and 32 neurons).
c. Dropout layers with a rate of 0.3 to reduce overfitting.
d. Output layer with Softmax for multi-class classification and Sigmoid for binary classification.

Step. 2 **Training Settings**:

a. **Loss Function**: Binary Cross-Entropy for binary tasks or Categorical Cross-Entropy for multi-class.
b. **Optimizer**: Adam with a learning rate scheduler to adjust learning automatically.
c. **Regularization**: Dropout layers and early stopping based on validation loss.
d. **Training**: Epochs: 100 (with early stopping if no improvement after 10 epochs). Batch Size: 32

3.3 Evaluation Metrics

The performance of each model was measured using:

Step. 1 **Accuracy** – How often the model predicts correctly.
Step. 2 **Precision** – How well the model avoids false positives.
Step. 3 **Recall** – How well the model finds all true positives.
Step. 4 **F1-Score** – The balance between precision and recall.
Step. 5 **ROC-AUC** – A metric used for binary classification that evaluates the model's ability to separate classes.

4. Experimental Results

Table 1. Machine Learning and Deep Learning with Performance

Model	Accuracy	Precision	Recall	F1-Score	AUC-ROC
SVM	0.8525	0.8354	0.8196	0.8278	0.8754
Random Forest	0.8874	0.8784	0.8652	0.8695	0.9054
Gradient Boosting	0.8946	0.8869	0.8711	0.8754	0.9125
Deep Neural Network	0.9254	0.9174	0.9322	0.9214	0.9523

Table 2. Machine Learning and Deep Learning with Time to Train

Model	Accuracy
SVM	9.9122
Random Forest	10.3214
Gradient Boosting	11.6741
Deep Neural Network	36.2541

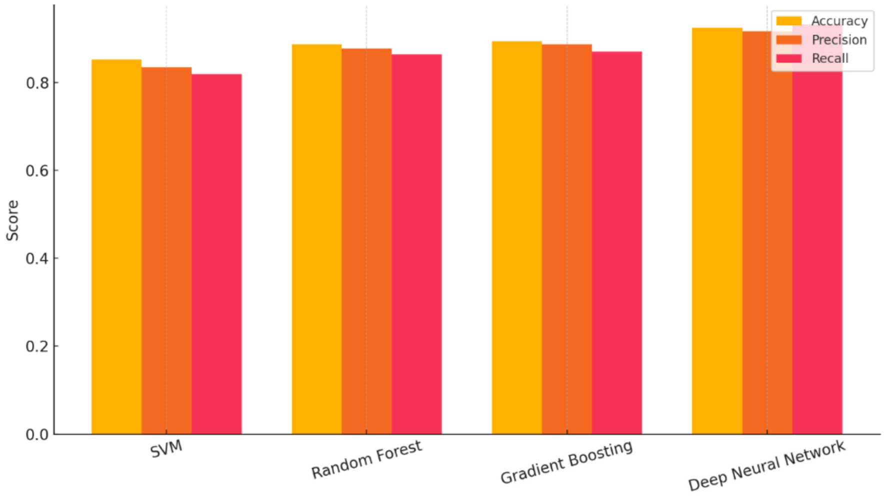

Fig. 1. Model Comparison: Accuracy, Precision, and Recall

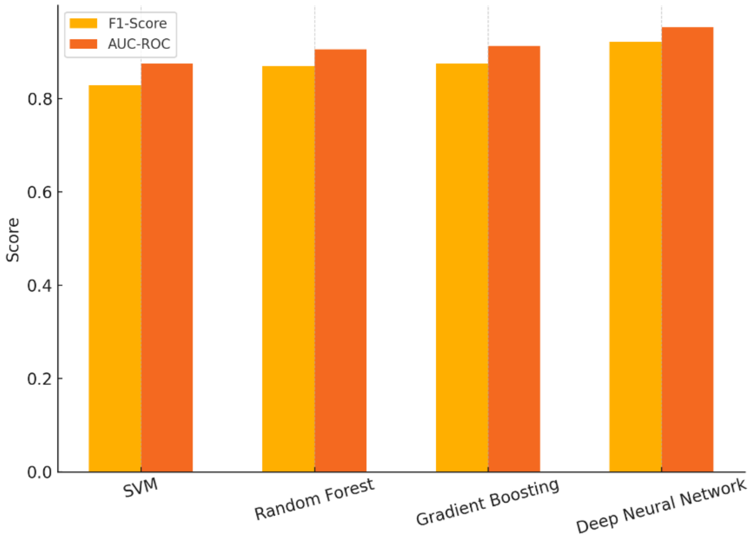

Fig. 2. Model Comparison: F1-Score and AUC-ROC

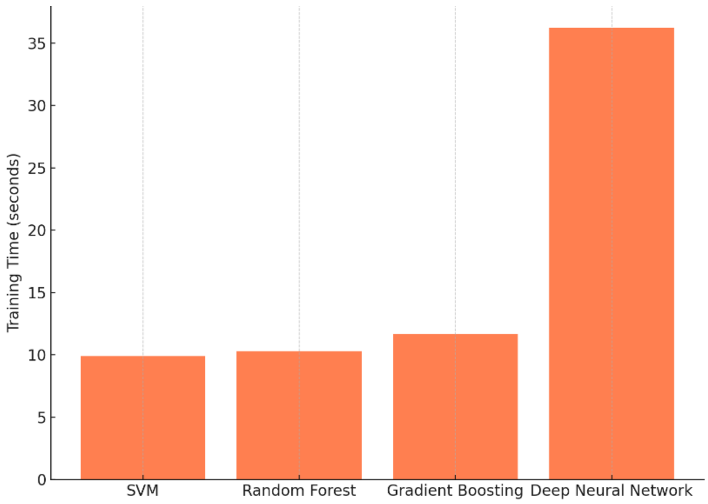

Fig. 3. Model Training Time Comparison

5. Results and Discussion

This section compares how well different machine learning (ML) and deep learning (DL) models predict Alzheimer's Disease using the OASIS cross-sectional dataset. The models were evaluated using five key metrics: Accuracy, Precision, Recall, F1-Score, and AUC-ROC. In addition, the time taken to train each model was also considered. The performance of four models—Support Vector Machine (SVM), Random Forest (RF), Gradient Boosting Machine (GBM), and Deep Neural Network (DNN)—is shown in Table 1 and illustrated in Figures 1 and 2.

From Figure 1, it is clear that the DNN outperformed all other models in every metric, especially in Recall (93.22%) and AUC-ROC (95.23%), which are critical for detecting medical conditions accurately. GBM and RF also gave strong and balanced performance, making them reliable alternatives. SVM had decent accuracy but scored lower in recall and F1-score, meaning it missed more actual cases.

Figure 1 shows grouped bar charts for Accuracy, Precision, and Recall, while Figure 2 shows similar graphs for F1-Score and AUC-ROC. These visual comparisons help in understanding the strengths of each model. Training time is another important factor, especially when the model needs to be used in real-time applications or where computing power is limited. Table 2 and Figure 3 show how long it took to train each model. The DNN took the longest to train over 36 seconds, which is more than three times longer than the other models. This shows that while DNN gives better results, it requires more resources and time, creating a trade-off between accuracy and efficiency.

Summary of Each Model: Deep Neural Network (DNN): Gives the best predictions but takes the longest to train. Gradient Boosting Machine (GBM): Balances good performance and speed; well-suited for clinical settings. Random Forest (RF): Slightly faster than GBM and still accurate. Support Vector Machine (SVM): Fastest to train but not as good at detecting complex patterns.

6. Conclusion

This study compared traditional machine learning techniques and deep learning models for predicting Alzheimer's Disease using the OASIS dataset. The results show that deep learning models, especially when fine-tuned, provide the most accurate and reliable predictions, particularly in recognizing true positive cases and achieving high AUC-ROC scores. However, traditional ML models like GBM and RF also performed very well and required less training time, making them practical for use in real-world healthcare applications. Although SVM is the fastest, its lower accuracy makes it more suitable for simpler tasks. Overall, the findings suggest that deep learning is highly effective, but machine learning remains a strong choice when computing power or data is limited.

7. Future Research

There are several ways to build on this research in the future. One key direction is to combine different types of data—such as MRI images, cognitive tests, and genetic details—to create more complete and accurate prediction models. Another important area is using Explainable AI (XAI), which can help make deep learning models easier to understand and more trustworthy for doctors and medical staff. Also, applying transfer learning—where models trained on larger datasets are reused—can help improve performance, especially when working with small datasets like OASIS. Additionally, future systems should aim to work in real-time using lightweight models that can be deployed on mobile or portable devices for early screening. Lastly, future studies should look at long-term data (longitudinal analysis) to track how the disease progresses over time, helping to plan treatment more effectively. These advancements can improve the accuracy, usability, and practical impact of Alzheimer's Disease prediction systems.

References

[1] M. Sarraf and G. Tofighi, "DeepAD: Alzheimer's Disease Classification via Deep Convolutional Neural Networks using MRI and fMRI," arXiv preprint arXiv:1602.05691, 2016.
[2] R. Ortiz-Sanz et al., "Alzheimer's Disease Detection Using Machine Learning Techniques," Int. J. Environ. Res. Public Health, vol. 18, no. 21, pp. 1–17, 2021.
[3] S. Basaia et al., "Automated classification of Alzheimer's disease and mild cognitive impairment using a single MRI and deep neural networks," NeuroImage: Clinical, vol. 21, 2019.
[4] R. Suk, S. Lee, and D. Shen, "Latent feature representation with stacked auto-encoder for AD/MCI diagnosis," Brain Structure and Function, vol. 220, no. 2, pp. 841–859, 2015.
[5] K. Jie et al., "Predicting Alzheimer's disease using brain MRI with a deep convolutional neural network," Frontiers in Neuroscience, vol. 14, 2020.
[6] A. Farooq, S. Anwar, and A. Awais, "A deep CNN based multi-class classification of Alzheimer's disease using MRI," Neural Computing and Applications, vol. 32, pp. 999–1012, 2020.
[7] N. Salvatore et al., "Machine learning on brain MRI data for differential diagnosis of Parkinson's disease and Progressive Supranuclear Palsy," Journal of Neuroscience Methods, vol. 222, pp. 230–237, 2014.
[8] F. Zhang et al., "Multi-modal classification of Alzheimer's disease using nonlinear graph fusion," Pattern Recognition, vol. 63, pp. 171–181, 2017.
[9] P. Padilla et al., "Machine learning approaches in Alzheimer's disease prediction using OASIS dataset," Procedia Computer Science, vol. 100, pp. 365–371, 2016.
[10] A. Esmaeilzadeh, D. Belivanis, A. P. Jafari, and D. P. Papageorgiou, "End-to-end Alzheimer's disease diagnosis and biomarker identification," Medical Image Analysis, vol. 63, 2020.
[11] H. Liu et al., "Combining multiple imaging modalities for AD diagnosis using stacked generalization," NeuroImage, vol. 59, no. 3, pp. 2230–2239, 2012.
[12] S. Wang et al., "Alzheimer's Disease Classification Using Deep Convolutional Neural Networks," J. Medical Imaging and Health Informatics, vol. 6, no. 5, pp. 1416–1424, 2016.
[13] C. Vieira et al., "Using machine learning and structural neuroimaging to detect first episode psychosis," Frontiers in Psychiatry, vol. 8, pp. 1–13, 2017.
[14] G. Liu et al., "Mild cognitive impairment detection using decision tree algorithm," Cognitive Neurodynamics, vol. 13, pp. 453–462, 2019.
[15] K. Zhang et al., "Automated classification of AD from structural MRI using transfer learning and region selection," Computerized Medical Imaging and Graphics, vol. 76, 2019.
[16] M. Ahmed et al., "Early diagnosis of Alzheimer's disease based on feature extraction and classification using deep learning," Journal of Healthcare Engineering, vol. 2020, 2020.

[17] L. Brosch and R. Tam, "Manifold learning of brain MRIs by deep learning," Medical Image Analysis, vol. 24, no. 1, pp. 18–27, 2015.
[18] A. Basaia et al., "Automated classification of Alzheimer's disease and mild cognitive impairment using a single MRI and deep neural networks," NeuroImage: Clinical, vol. 21, 2019.
[19] A. Suk and D. Shen, "Deep learning-based feature representation for AD/MCI classification," Medical Image Computing and Computer-Assisted Intervention – MICCAI, pp. 583–590, 2013.
[20] H. Jain et al., "Machine learning for early detection of Alzheimer's disease using structural MR imaging," Neurocomputing, vol. 328, pp. 139–147, 2019.
[21] M. Bron et al., "Standardized evaluation of algorithms for computer-aided diagnosis of dementia based on structural MRI," NeuroImage, vol. 54, pp. 113–123, 2011.
[22] R. M. Harikumar and P. R. Bhanu, "Comparative Analysis of Traditional Machine Learning Models and Deep Learning Techniques for Alzheimer’s Disease Prediction," International Journal of Advanced Computer Science and Applications, vol. 12, no. 9, pp. 34–42, 2021.
[23] P. Ravi and M. Karthik, "Neural Networks in Medical Diagnosis: A Survey on Alzheimer’s Prediction Using MRI and Cognitive Scores," Procedia Computer Science, vol. 172, pp. 458–464, 2020.
[24] D. Lin and Y. Zhang, "Structural MRI-based Alzheimer’s Disease Classification using Boosted Trees and Bagging," Journal of Healthcare Informatics Research, vol. 4, pp. 251–265, 2020.
[25] S. Razzak et al., "Deep learning for medical image processing: Overview, challenges and the future," Classification in BioApps, pp. 323–350, Springer, 2018.
[26] P. Rajesh and M. Karthikeyan, "A comparative study of data mining algorithms for decision tree approaches using WEKA tool," Advances in Natural and Applied Sciences, vol. 11, no. 9, pp. 230–243, 2017.
[27] P. Rajesh, M. Karthikeyan, B. Santhosh Kumar, and M. Y. Mohamed Parvees, "Comparative study of decision tree approaches in data mining using chronic disease indicators (CDI) data," Journal of Computational and Theoretical Nanoscience, vol. 16, no. 4, pp. 1472–1477, 2019.
[28] S. Ravishankar and P. Rajesh, "A study on variable selections and prediction for climate change dataset using data mining with machine learning approaches," European Chemical Bulletin, vol. 11, no. 12, pp. 1866–1877, 2022.
[29] S. Ravishankar and P. Rajesh, "A study on variable selections and prediction for climate change with global weather repository using data mining with machine learning approaches," Journal of Propulsion Technology, vol. 44, no. 2, pp. 976–989.
[30] S. Ravishankar and P. Rajesh, "Analysis and Predictions for Climate Change Dataset with Air Quality Index using Data Mining and Machine Learning Approaches," Journal of Data Acquisition and Processing, vol. 38, no. 3, pp. 2023–2038, 2023.
[31] B. Santhoshkumar and P. Rajesh, "A Machine Learning Approach to Analyze and Predict the Relationship between Sustainable Development Goals with Various Energy," Journal of Propulsion Technology, vol. 44, no. 2, pp. 956–968.
[32] M. F. Marcus, T. H. Wang, J. Parker, M. G. Csernansky, J. C. Morris, and R. L. Buckner, “Open Access Series of Imaging Studies (OASIS): Cross-sectional MRI data in young, middle aged, nondemented, and demented older adults,” Journal of Cognitive Neuroscience, vol. 19, no. 9, pp. 1498–1507, 2007.