AI Driven Digital Cognitive Twin for Brain Health Monitoring

Jawahar Jonathan K., G Prakash Babu., Ooha E., Preethi K., Preethi Nityanand Shetty., Vikas Gowda D

Computer Science & Engineering Acharya Institute of Technology Bengaluru, India

DOI: https://doi.org/10.51244/IJRSI.2026.1310000079

Received: 09 October 2026; Accepted: 14 October 2026; Published: 01 November 2026

ABSTRACT

Cognitive health is usually checked in clinics. That makes it hard to keep track of changes over the time. This paper introduces CogTwin, a web-based platform where a person can complete five cognitive tasks—reaction time, memory recall, pattern recognition, attention span and decision making—and see how their own results change over the time. Of comparing the user to population averages the platform creates the personal baseline using data from the user's past sessions. This baseline acts like the cognitive twin. Each new session is compared to this baseline using a z-scores. A regression line shows the trend and a TensorFlow.js neural network predicts a what the session's scores might be. A language model gives explanations in the everyday language following rules that keep it from making any diagnostic claims. The system is built using React, Node.js, Express and the MongoDB Atlas. User accounts are secured with the bcrypt and JWT. At this point CogTwin has only been tested as the software. No real users have been involved yet so the paper does not say that it measures the cognition accurately. The model and the rules for detecting an anomalies are described in full. A pilot study protocol is also included, covering test–retest reliability, comparison, with the MoCA and prediction error compared to the baseline. CogTwin is meant to help with monitoring a changes not to diagnose any condition.

Index Terms: digital twin, cognitive monitoring, brain health, TensorFlow.js, z-score, linear regression, personalized healthcare

INTRODUCTION

Changes in memory, attention, reaction time and decision making are often among the first signs that a person's brain health is shifting. Cognitive assessments, however, are usually carried out in a clinic by trained staff, so they happen rarely and are difficult to repeat often enough to show change over time. Researchers have therefore examined machine learning on non-invasive data such as speech, facial expression and movement [1]. Lin et al. built a computer-based visual memory test that uploads its results to cloud servers, which removes errors caused by manual data entry [2]. Gait patterns have been classified with LSTM networks to flag cognitive impairment [3], and the voice has been studied as a possible biomarker [4].

Digital twin technology offers another route to personalized care. A digital twin is a virtual model of a person or a system that is kept up to date with new data, and it can be combined with AI, IoT and cloud services for monitoring and prediction [5]–[7]. Open problems remain, including data quality, privacy, scalability and real-time updating [6], [7].

I read this paper. It describes CogTwin, a web platform that brings the digital twin idea into everyday cognitive testing. A user completes five tasks. The results create a baseline and later sessions are compared to that baseline rather than to other people. The CogTwin platform combines five tasks with personal-baseline tracking, anomaly flags, trend lines and a short-term prediction model. The paper provides a description of the baseline, the anomaly rule and the prediction model so that the approach can be reproduced and tested. The paper presents a safety and privacy design for the AI assistant, which is limited to explaining results and may not make statements. The paper outlines a validation plan, for a pilot study because the current evaluation covers software behavior. I note that Section II reviews related work. Section

III describes the method. Section IV discusses the implementation and Section V covers safety and privacy. Section VI reports the evaluation done far and the planned pilot study. Section VII lists limitations and Section

ISSN No. 2321-2705 | DOI: 10.51244/IJRSI |Volume XIII Issue X October 2026

VIII concludes.

Related Work

Several studies have used AI and digital twins in healthcare. Khan and colleagues introduced a twin for tracking strokes that uses UWB radio sensing, autoencoders and a finely tuned KNN classifier [5]. Dan and others talked about how AI and digital twins can help with diagnosing and treating health issues by creating personalized patient models [6]. Chen and others looked at AI-powered human digital twins in IoT healthcare explaining how data is gathered, shared and processed [7]. Momand and others developed a twin system for elderly care based on physiological data using LSTM and Bi-LSTM models along, with generative AI for feedback [8]. Blasch studied twins for better understanding of situations [9].

Other research focuses on assessment directly. Alsuhaibani and others examined non-invasive machine learning techniques that use speech, language, facial and movement features [1]. Lin and others tested a visual memory assessment system that uses cloud data transfer [2].. Jung et al. estimated the risk of cognitive impairment from gait with LSTM networks [3], and Pacheco-Lorenzo et al. analyzed voice biomarkers [4]. Shubar et al. studied cost-effective model selection for early cognitive impairment prediction [10]. Irfan et al. proposed a voting-based ensemble of classifiers, trained on selected cognitive features, for early detection of Alzheimer's disease [11]. Baytas predicted progression from mild cognitive impairment to Alzheimer's dementia with a neural network framework that uses adversarial attacks [12]. These studies show that AI can help with assessment and prediction, but each one depends on a particular kind of data or a particular condition.

Table I Comparison Of Existing Approaches

Ref.	Focus / methods	Main limitation
[1]	Review of ML on speech, facial and movement data	Data and standardization issues
[2]	Embedded system for visual memory assessment	Needs special hardware
[3]	Gait analysis with ML and LSTM for cognitive risk	Depends on gait sensor data
[4]	Voice-based cognitive assessment with ML	Small study sample
[5]	RF sensing and AI-based digital twin for stroke	Limited to stroke monitoring
[6]	AI and digital twin for mental health	Needs good-quality data
[7]	Generative AI with human digital twins	Privacy and data-handling concerns
[8]	Digital twin with LSTM models for elderly care	Complex system
[9]	AI-based digital twin for situation awareness	Limited real-world use
[10]	Cost-effective ML for cognitive impairment prediction	Depends on suitable training data
[11]	Voting ensemble on cognitive features for Alzheimer's detection	Depends on suitable training data
[12]	Neural network for MCI-to-Alzheimer's progression	Needs neuroimaging and genetic data

Research gap. Most earlier systems look at a single condition or a single kind of signal, such as stroke, speech, gait or visual memory. Few combine several simple cognitive tasks with continuous tracking against the user's own history, and many need special sensors or clinical data. CogTwin addresses part of this gap with a low-cost, browser-based tool that follows five cognitive areas against a personal baseline.

ISSN No. 2321-2705 | DOI: 10.51244/IJRSI |Volume XIII Issue X October 2026

PROPOSED METHODOLOGY

A. Overview

CogTwin is a client–server web application. The user works in the browser, while the back end handles login, stores data, runs the statistical and machine learning steps, calls the AI service and talks to the database. The digital cognitive twin is updated after every session, so the system follows how each person changes and not how they compare with the general population. Fig. 1 shows the architecture and Fig. 2 the workflow.

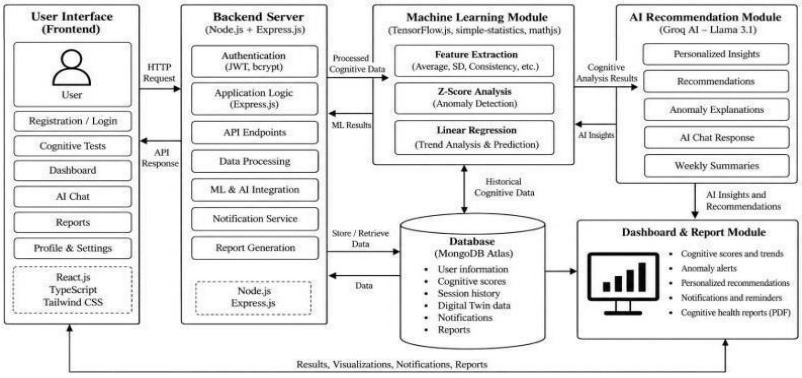


Fig. 1. System Architecture of CogTwin

Fig. 1. System architecture of CogTwin.

B.System Modules

1) *User interaction and authentication.* User. Logs in through this module. User passwords are hashed with bcrypt and User sessions are kept with JSON Web Tokens (JWT). After login User can reach the tests and a personal dashboard.

2) *Cognitive assessment*. Cognitive assessment module runs five tasks: reaction time, memory recall, pattern recognition, attention span and decision making. Each task returns a score. Scores are saved with the session record. Scores are later used to set the baseline and to follow change over time.

3) *Data processing and machine learning*. For each task Data processing and machine learning module computes features: score, standard deviation, consistency, learning rate, attention variability and improvement trend. Z-scores flag sessions linear regression gives the trend and a neural network predicts the next session.

4) *Digital twin and AI recommendations.* Digital twin and AI recommendations stores a model of User: session history, baseline statistics, regression trend and trained predictor together with times when User took the tests. A language model (Llama 3.1 through the Groq SDK) receives a summary of these results. Produces explanations for flagged sessions, suggestions, answers, in the chat and weekly summaries.

5) *Dashboard, monitoring and reports.* Dashboard shows scores, trends, predictions and flags. Dashboard module also sends reminders when an assessment is overdue. Creates a downloadable PDF report. All data are stored in MongoDB Atlas.

The simple forecast is also used as one of the comparison methods, for the network.

E. Neural Network Predictor

A small neural network built with the TensorFlow.js predicts the user's score on each task from their recent history. Table II lists its specification. The entries in brackets must be completed from the implementation, before submission..

ISSN No. 2321-2705 | DOI: 10.51244/IJRSI |Volume XIII Issue X October 2026

Table Ii Neural Network Specification

Item	Description
Inputs	I list eight normalized features and these eight normalized features are: memory score, reaction-time score, pattern-recognition score, attention score, decision-making score normalized session index, normalized day of the week and normalized hour of the day.
Output	Predicted score for each of the five tasks in the next session
Architecture	A sequential neural network with eight input features, a dense hidden layer containing 32 ReLU units, a second dense hidden layer containing 16 ReLU units, and one sigmoid output unit.
Loss, optimizer	Mean squared error (MSE) loss with the Adam optimizer and a learning rate of 0.005.
Training data	User-specific completed assessment sessions, with up to 50 sessions retrieved for training. At least three complete sessions with valid input and output values are required.
Schedule	Training is initiated in the background after a completed session when sufficient session history is available. The model uses 100 epochs, a batch size of up to four samples, and shuffled training data. No separate validation split is configured in the training function.
Evaluation	Mean absolute error (MAE) on held-out later sessions, compared with (a) carrying forward the last score and (b) the linear regression forecast

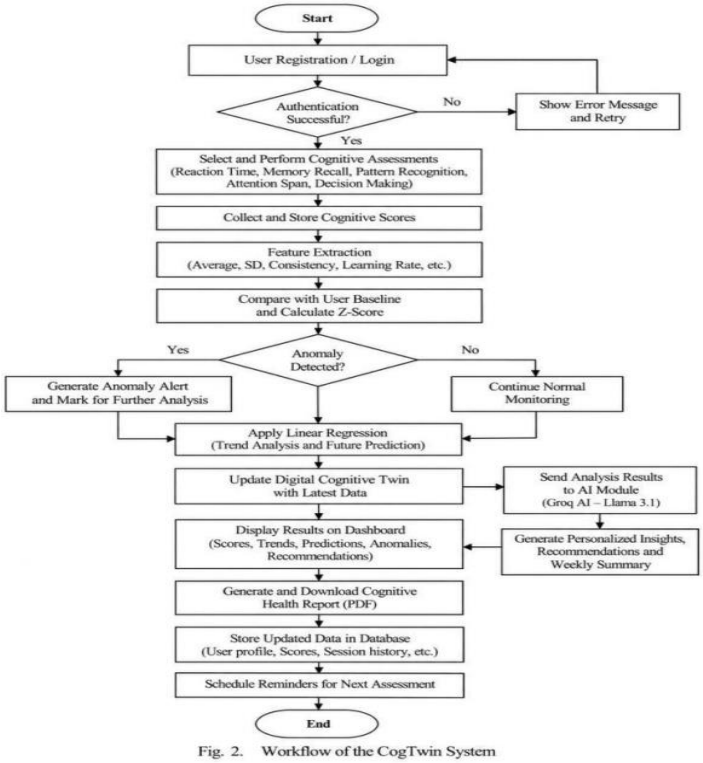


Fig. 2. Workflow of the CogTwin System

Fig. 2. The Workflow of the CogTwin system.

C. Personal Baseline and Anomaly Detection

CogTwin builds a personalized performance baseline from the user's completed assessment sessions. After three completed sessions CogTwin uses the assessment history to calculate the average score for each cognitive task. Up to ten recent completed sessions are considered when retrieving records. The resulting averages become

ISSN No. 2321-2705 | DOI: 10.51244/IJRSI |Volume XIII Issue X October 2026

reference values for examining changes in assessment results.

For each assessment score CogTwin calculates a scaled deviation, from the corresponding baseline mean using a fixed divisor of 10. The calculation is expressed as follows:

$$z = (x - \mu) / 10 \quad (1)$$

I want to explain how we handle scores. Here (x) represents the recorded score and (mu) denotes the mean of the historical scores that have been collected before. The constant 10 is used as a scaling factor in the current implementation. The constant 10 does not represent the deviation of the historical scores. The system categorizes a deviation as mild when the absolute value of the deviation exceeds 1.5, when the absolute value of the deviation exceeds 2.0 and severe when the absolute value of the deviation exceeds 2.5. These thresholds help identify score variations that may warrant review. The resulting flags support the users review of performance history. Are not intended to indicate or diagnose a medical condition. Changes in scores may also be influenced by sleep, stress, distractions, device differences and familiarity, with repeated assessments.

D.Trend Analysis

For each task ordinary least-squares linear regression is fitted to the score sequence against the session number t:

$\hat{y}t = a + b \cdot t$ (2)

I observe that the slope b tells whether the user is improving, declining or staying flat. The ordinary least-squares linear regression line, when extended by one step gives a forecast.

A prediction model is only worth including if it beats very simple alternatives, and the simplest of these is to assume that the next score equals the last one. Section VI-C describes how the network will be compared with that rule and with the regression forecast. Until this comparison is done, the predictions shown to users should be treated as indicative only.

F. Algorithm

Algorithm 1: CogTwin session processing

Input: new assessment scores and the stored history. Output: anomaly status, trend, predicted scores and explanations. Step 1: The user logs in and the stored records are retrieved. Step 2: The user completes the five assessments.

Step 3: The scores are saved in the database.

Step 4: Features are extracted: mean, standard deviation, consistency, learning rate, attention variability and improvement trend.

Step 5: If the user has than N sessions, for a task only the baseline is updated and the algorithm moves to Step 9.

Step 6: The baseline μ and σ are derived from the user's history.

Step 7: The z-score of the session is calculated for each task. Step 8: A task is marked as an anomaly if |z| reaches the threshold; otherwise the task remains normal.

Step 9: Regression shows the trend and the neural network predicts the scores.

Step 10: The digital cognitive twin is updated with the scores and session details.

Step 11: A summary of the results is sent to the AI module. Step 12: The AI module returns insights, anomaly explanations and the weekly summary.

ISSN No. 2321-2705 | DOI: 10.51244/IJRSI |Volume XIII Issue X October 2026

Step 13: The dashboard shows scores, trends, predictions, flags and suggestions.

Step 14: A report is. A reminder, for the next assessment is scheduled.

Step 15: The updated results are stored for later baseline comparison.

Implementation

CogTwin follows a client–server design. The React front end handles the user interface, and the Node.js and Express back end handles application logic and API calls. The statistical and machine learning code runs in the back end. User details, test results, session records, twin data, notifications and reminder preferences are stored in MongoDB Atlas. Table III lists the technology stack and Fig. 3 the implementation modules.

Table Iii Technology Stack

Component	**Technology**	**Purpose**
Front end	React.js, TypeScript, Tailwind CSS	Interface and visualization
Back end	Node.js, Express.js	API and application logic
Database	MongoDB Atlas	Data storage and retrieval
Machine learning	TensorFlow.js	Score prediction
Statistics	simple-statistics, math.js	Feature and z-score calculation
Authentication	JWT, bcrypt.js	Secure login
AI	Groq SDK (Llama 3.1)	Insights, explanations, chat
Visualization	Recharts, Chart.js	Trend charts
Testing	Jest, Postman	Software and API testing
Development	VS Code, GitHub	Coding and version control

Safety, Privacy And Ethics

Not a diagnostic tool. CogTwin tracks change in a person's own task scores. A low score or an anomaly flag can have many everyday causes, such as tiredness, stress, distraction, a different device, or simply practice effects from repeating a task. The platform therefore does not diagnose and must not replace a clinical assessment.

Limits on the chatbot. In CogTwin the language model follows instructions that are set up for types of user requests. When the language model answers health-related questions the instructions tell the chatbot to avoid giving diagnoses and to advise users to consult a healthcare professional if they have medical concerns. The instructions, for explaining assessment results ask the language model to use clear and reassuring language. Separate prompts help the language model generate insights, recommendations and weekly summaries.

The back end also deals with JSON-formatted answers for some insight and recommendation features to keep the data setup.. This handling does not fully protect against medical issues. The chat interface that was checked does not add a warning to every answer automatically. Because of that the system can't say for sure that all the information it gives is correct or right in every case. More work should involve testing with questions that relate to health ways to check the answers and more ways to stop wrong medical statements. CogTwin is meant to let users look at their thinking skills, over time. Should not be used instead of a real doctors checkup.

ISSN No. 2321-2705 | DOI: 10.51244/IJRSI |Volume XIII Issue X October 2026

Data protection. CogTwin uses bcrypt to hash passwords and JWTs to control access, to protected API routes. MongoDB Atlas stores user. Cognitive assessment records. The Groq AI service receives selected assessment information to generate insights and explanations; some requests may also include the users name. Future improvements should reduce the sharing of information and strengthen privacy controls.

Ethics for the pilot study. The study in Section VI-C will start only after approval from an institutional ethics committee, with written informed consent from every participant. Data will be anonymized before analysis.

Evaluation

A. Software Verification

So far the evaluation has checked that the software behaves as designed. It has not tested whether CogTwin measures cognition. The tests covered units (authentication, score calculation, anomaly detection, AI recommendation and report generation), integration of the front end, back end, ML engine, AI module and database, user functions (login, assessments, dashboard, notifications, AI chat and reports), database storage and retrieval, responsiveness, and the user interface on supported browsers and devices. Table IV shows selected test cases.

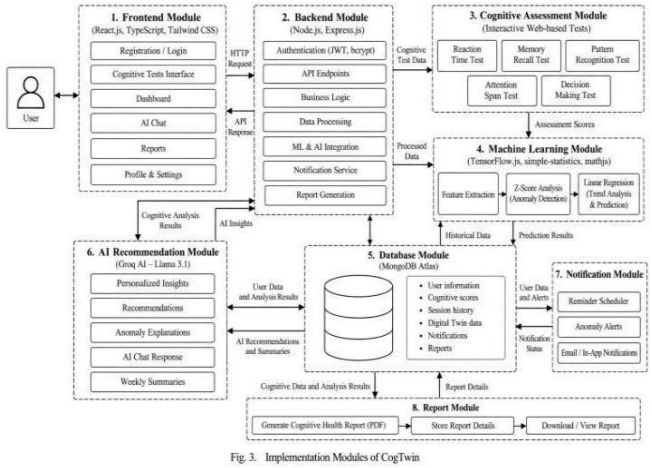


Fig. 3. Implementation Modules of CogTwin

Fig. 3. Implementation modules of CogTwin.

Table Iv Selected Test Cases And Observed Results

Test case	**Condition and expected result**	**Observed**
Reminder	Assessment left incomplete for 7 days; a reminder is generated	Reminder generated
Test case	**Condition and expected result**	**Observed**
Dashboard	Cognitive data in the database; scores, charts and trends are shown	Displayed correctly
AI chat	User asks a health-related question; a relevant response is returned	Response generated
Authentication	Valid credentials; authorized access is granted	Access granted
Assessment	User completes a test; the score is calculated and stored	Score stored

These tests show that the components work together. Response times, throughput and prediction accuracy were not recorded, so no quantitative performance claims are made and the system is described as functional only.

ISSN No. 2321-2705 | DOI: 10.51244/IJRSI |Volume XIII Issue X October 2026

B.What the Current Evaluation Does Not Show

No participants have used the system in a controlled study. The following are therefore unknown: how consistent a person's scores are from one session to the next, whether the scores relate to established cognitive tests, how well the neural network predicts compared with simple methods, and how often the z-score rule raises a false alarm. The results in Section VI-A must not be read as evidence on any of these points.

C. Planned Pilot Study

The following study is planned, subject to ethics approval.

- *Participants:* about 30 healthy adult volunteers, with age, education and device type recorded.
- *Sessions:* each participant completes all five tasks at least four times over three to four weeks, at similar times of day where possible, on their own devices.
- *Reliability:* for each task, the intraclass correlation coefficient (ICC) between sessions with 95% confidence intervals. Practice effects are reported separately.
- *Validity:* in the first week, participants also take the Montreal Cognitive Assessment (MoCA) under supervision, and Spearman correlations between CogTwin task scores and MoCA scores are reported. Healthy volunteers score near the ceiling on the MoCA, so the study will also state what it cannot show for clinical populations.
- *Prediction:* the neural network is trained on early sessions and tested on later ones. Its MAE is compared with (a) carrying forward the last score and (b) the linear regression forecast.
- *Anomaly threshold:* because the volunteers are stable, the flags they receive estimate the false-alarm rate directly. It is reported for several thresholds so that a sensible one can be chosen.

Limitations

The main limitation is that CogTwin has not been validated with users. The z-score rule relies on a small number of baseline sessions, and the scores are affected by practice effects, device differences (for example touch screen against mouse) and the user's state on the day. The neural network has not yet been shown to improve on simpler methods. The AI assistant depends on an external language model whose behavior can change and whose answers cannot be guaranteed correct, so its output is supportive information and not advice. Finally, the tasks are short browser tasks and are no substitute for neuropsychological testing.

CONCLUSION AND FUTURE WORK

This paper presented CogTwin, a web platform that follows a person's cognitive task results against their own baseline, flags unusual sessions with z-scores, shows trends with linear regression, predicts the next session with a small neural network and explains the results in plain language. The software has been built and tested for correct operation. It has not been tested with participants, and the paper does not claim that it measures cognition accurately. The next step is the pilot study in Section VI-C, which will provide reliability, validity and prediction results and will allow the anomaly threshold and the prediction model to be justified with data. Later work could add a larger and longer study, wearable or EEG data and a mobile app. Any use of the system to support clinicians must leave the final judgment with qualified medical professionals.

REFERENCES

1. M. Alsuhaibani, A. Pourramezan Fard, J. Sun, F. Far Poor, P. S. Pressman, and M. H. Mahoor, "A review of machine learning approaches for non-invasive cognitive impairment detection," IEEE Access, vol. 13, pp. 56355–56384, 2025, doi: 10.1109/ACCESS.2025.3555176.
2. C.-L. Lin, T.-C. Chu, M.-H. Wu, M.-Y. Deng, W.-C. Chiu, C.-H. Chen, P.-S. Sung, W.-L. Tsao, and T.-

C. Lin, "Evaluation of novel cognitive assessment system for testing visual memory of the elderly," IEEE Access, vol. 9, pp. 47330–47337, 2021, doi: 10.1109/ACCESS.2021.3065684.

3. Jung, J. Kim, M. Kim, C. W. Won, and K.-R. Mun, "Classifying the risk of cognitive impairment using sequential gait characteristics and long short-term memory networks," IEEE J. Biomed. Health Inform., vol. 25, no. 10, pp. 4029–4040, 2021, doi: 10.1109/JBHI.2021.3073372.
4. M. R. Pacheco-Lorenzo et al., "Analysis of voice biomarkers for the detection of cognitive impairment," IEEE Access, 2024, doi: 10.1109/ACCESS.2024.3442431.
5. S. Khan, U. Anwar, A. Khan, and T. Arslan, "RF-based sensing and AI decision support for stroke patient monitoring: A digital twin approach," IEEE Access, 2025.
6. S. Dan, J. Ghosh, B. Sarkar, and J. Bhattacharya, "AI and digital twins: A breakthrough approach to mental health diagnosis and therapy," IGI Global, 2025.
7. J. Chen, Y. Shi, C. Yi, H. Du, J. Kang, and D. Niyato, "Generative-AI-driven human digital twin in IoT healthcare: A comprehensive survey," IEEE Internet Things J., 2024.
8. Z. Momand, P. Mongkolnam, J. H. Chan, N. Charoenkitkarn, and D. Pal, "Building digital twins for elderly care: An end-to-end framework from data acquisition to modeling," IEEE Access, 2025.
9. Blasch, "Digital twins for cognitive situation awareness," in Proc. IEEE CogSIMA, 2024.
10. G. Shubar, K. Ramakrishnan, and C.-K. Ho, "Optimizing machine learning models for accessible early cognitive impairment prediction: A novel cost-effective model selection algorithm," IEEE Access, vol. 12, pp. 180792–180814, 2024, doi: 10.1109/ACCESS.2024.3505038.
11. M. Irfan, S. Shahrestani, and M. Elkhodr, "Early detection of Alzheimer's disease using cognitive features: A voting-based ensemble machine learning approach," IEEE Eng. Manag. Rev., vol. 51, no. 1, pp. 16–25, Mar. 2023.
12. M. Baytaş, "Predicting progression from mild cognitive impairment to Alzheimer's dementia with adversarial attacks," IEEE J. Biomed. Health Inform., vol. 28, no. 6, pp. 3750–3760, 2024.