JETIR.ORG ISSN: 2349-5162 | ESTD Year : 2014 | Monthly Issue
JOURNAL OF EMERGING TECHNOLOGIES AND INNOVATIVE RESEARCH (JETIR)
An International Scholarly Open Access, Peer-reviewed, Refereed Journal

# Artificial Intelligence and Machine Learning for Early Detection, Diagnosis, and Progression Prediction of Alzheimer's Diseases

**[1]Balwant Kumar, [2]Sandip Kumar Singh Modak**
[1]Research Scholar, [2]Associate Professor,
[1]Sarala Birla University,
[1]Sarala Birla University, Ranchi, India

*Abstract:* Alzheimer's Disease (AD) is a degenerative disorder of the nervous system that slowly impairs memory, thinking ability, and routine activities. Conventional approaches to diagnosing AD mainly depend on symptoms, cognitive tests, and brain imaging, which often detect the disease only after considerable damage has already occurred. In recent years, the rapid growth of Artificial Intelligence (AI) and Machine Learning (ML) has opened new possibilities for identifying AD at much earlier stages. Advanced deep-learning models, including Convolutional Neural Networks (CNNs) and Recurrent Neural Networks (RNNs), have shown strong capability in recognising even slight changes and patterns that are often overlooked in MRI and PET scans, speech recordings, as well as genetic information [2], [3], [4]. This chapter provides a detailed discussion of how AI- and ML-based methods contribute to early diagnosis, automated classification, and prediction of disease progression in Alzheimer's Disease.
**Keywords:** Alzheimer's Disease, biomarkers, cognitive testing, CNN models, deep learning, diagnosis, early detection, EEG signals, machine learning, medical imaging, mild cognitive impairment (MCI), multimodal data, neurodegeneration, neuroimaging, prediction models, progression analysis, RNN models, speech indicators, SVM classifiers, wearable sensors.

## 1. Introduction

Alzheimer's Disease (AD) is the most common type of dementia and affects a very large number of people around the world. Changes in the brain, such as the buildup of amyloid-beta and problems with tau proteins, start many years before a person shows memory loss or other symptoms [1]. Normal methods of diagnosis, like MRI scans, PET scans, and routine check-ups, often cannot detect these early changes because they are not sensitive enough.
Newer technologies, such as Artificial Intelligence (AI) and Machine Learning (ML), are helping solve this problem. These methods can study large and complex medical data and find small signs of the disease that humans may miss. Deep-learning techniques, particularly Convolutional Neural Networks, have shown strong performance in analysing brain imaging data and identifying early structural changes [4]. These models can assist in recognising the early stages of Alzheimer's Disease, distinguishing Alzheimer's patients from those with Mild Cognitive Impairment (MCI), and even estimating the likelihood that an individual with MCI may progress to dementia in the future [3].

## 2. Background and Literature Review

Alzheimer's Disease (AD) is a disorder in which the brain slowly loses its ability to function normally. Over time, the nerve cells begin to die, and the links between these cells become weak. As this continues, important parts of the brain, especially the hippocampus, which plays a major role in forming memories, start to shrink. These changes begin years before a person notices problems such as forgetting things or feeling confused. Because of this slow and hidden progression, spotting Alzheimer's early is very important. Early detection gives doctors a better chance to plan treatment and take steps that may help delay the worsening of symptoms.
Conventional techniques such as memory examinations, MRI, PET scans, and other clinical checks come with several drawbacks. These approaches usually rely heavily on the doctor's judgment, need costly and advanced equipment, and are not always easily available. In many cases, they may also miss the very small and early changes taking place in the brain, making early detection difficult.
To deal with the limitations of traditional diagnostic methods, researchers are increasingly relying on Artificial Intelligence and Machine Learning. These techniques make it possible to analyse large volumes of medical information and uncover subtle trends that may be overlooked during manual examination. The availability of extensive datasets such as ADNI, AIBL, and OASIS has further strengthened this field by providing high-quality data for building and improving AI-based diagnostic models. [5], [10].
Neuroimaging remains one of the key approaches for studying Alzheimer's Disease because it allows researchers to closely observe changes taking place inside the brain. A number of investigations have shown that deep-learning methods, especially

Convolutional Neural Networks (CNNs), are highly effective in identifying the fine structural differences associated with the disease. For example, Basaia et al. (2019) applied CNN models to MRI data and were able to clearly distinguish individuals with Alzheimer's from those diagnosed with Mild Cognitive Impairment (MCI) [2]. Likewise, Nie et al. (2020) examined MRI scans collected over several time points to determine which MCI patients were more likely to progress to Alzheimer's in the future [3]. Earlier work has also demonstrated that CNNs are particularly strong at identifying meaningful features hidden in medical images [4]. Measurements such as hippocampal volume, cortical or tissue thickness, and broader changes in brain structure continue to play an important role in supporting early diagnosis [12].

PET scans play an important role in Alzheimer's research because they allow doctors and scientists to see how amyloid and tau proteins accumulate in the brain, both of which are key markers of the disease. Studies that use PET together with MRI have shown that combining information from multiple imaging techniques often provides clear and more reliable results than relying on a single scan type [6]. In recent years, deep-learning models have also been applied to tau-PET images to estimate amyloid burden, showing how advanced computational methods can support early and more precise detection [13].

Speech-based indicators are also gaining attention as a useful way to detect Alzheimer's at an early stage. Subtle shifts in the way a person talks, such as longer pauses, slower responses, or trouble forming clear sentences, can serve as early clues of cognitive decline. In one study, Sardo demonstrated that analysing speech with advanced language-processing techniques can accurately distinguish individuals with Alzheimer's [7]. Other work has shown that even ordinary recorded conversations can provide enough information to spot early signs of declining mental function [14].

Wearable devices have also become a valuable tool for observing behavioural changes linked to Alzheimer's. These sensors can track walking speed, balance, sleep habits, and overall daily movement. Research has found that individuals who walk more slowly or show uneven gait patterns are more likely to experience early cognitive difficulties [15]. When this type of behavioural information is analysed using Machine Learning models, it can help identify the early stages of Alzheimer's with fairly good accuracy.

Cognitive assessments and clinical records continue to play an important role in Alzheimer's research. Deep-learning models such as RNNs and LSTMs have shown strong potential in analysing cognitive test scores collected over long periods, helping to predict which patients with Mild Cognitive Impairment (MCI) may eventually develop Alzheimer's [3]. Similarly, information stored in Electronic Health Records (EHRs), including past illnesses, medications, laboratory findings, and other medical details, has also been found valuable for building machine-learning models that can forecast disease progression [16].

Explainable AI is becoming important because doctors want to understand how AI makes decisions. XAI methods help show which features or patterns the model is using. Rezaei et al. (2021) introduced XAI methods that explain predictions related to AD progression [8]. Studies also show that using multiple types of data together, such as MRI, PET, speech, and genetic information, improves accuracy by 15–20% compared to using one data type alone [6], [11].

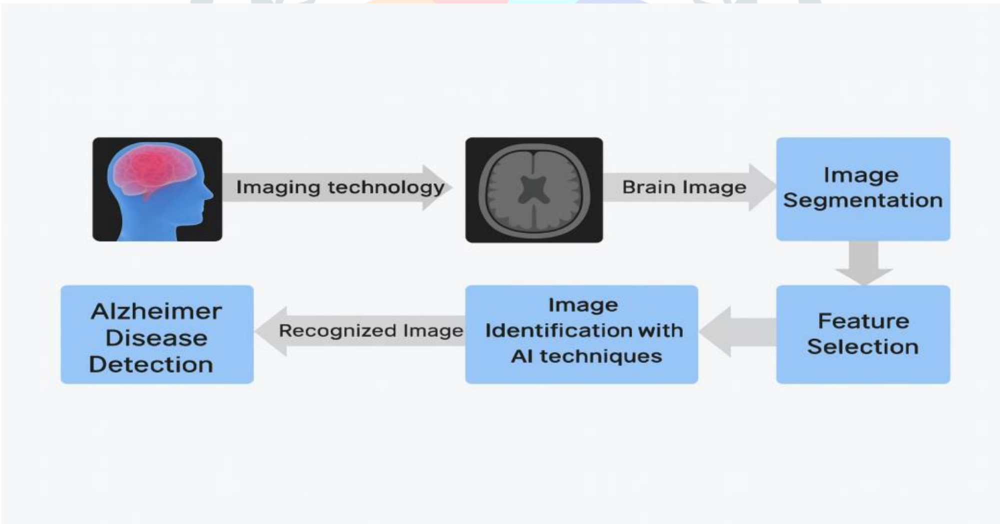


Figure 1: Conceptual Framework of AI-based Alzheimer's Disease Detection

## 3. AI and ML Techniques in Alzheimer's Disease Detection

Artificial Intelligence and Machine Learning techniques are widely used to identify and predict Alzheimer's Disease by analysing large and complex medical data such as MRI, PET scans, genetic data, and clinical records. These methods help detect early brain changes that are difficult to notice through manual observation.

### 3.1 Support Vector Machine (SVM)

SVM is a simple machine learning method used to separate data into different groups. It draws the best possible line or boundary between two classes, for example, "Healthy" and "Alzheimer's." This boundary helps the computer decide which group a new patient belongs to. SVM works very well with MRI data because it can handle many features at the same time, such as brain size, tissue thickness, and hippocampal volume. It has been used in Alzheimer's research by Sara Basaia, Federica Agosta, Nicola Wagner, Valentino Canu, Giancarlo Magnani, Massimo Santangelo, and Massimo Filippi, who applied SVM on MRI scans to successfully classify Alzheimer's and Mild Cognitive Impairment patients [2].

### 3.2 Decision Tree

A Decision Tree is one of the simplest methods used in machine learning, and it works by breaking a decision into a sequence of small, logical steps. It functions much like a flowchart: the model begins with a question, follows the answer to the next branch, and continues this process until it reaches an outcome. In an Alzheimer's context, for example, the tree might first look at a person's memory test score, then examine hippocampal volume, and finally classify the individual as Healthy, MCI, or Alzheimer's.

One major advantage of Decision Trees is that they clearly show how the final prediction was reached, which makes them easy for doctors and researchers to interpret. Because of this transparency, Decision Trees have been used often in medical imaging research, including work on Alzheimer's disease discussed in studies reviewed by Geert Litjens, Thijs Kooi, Babak Ehteshami Bejnordi, Arnaud Setio, Francesco Ciompi, Mohsen Ghafoorian, Clara Sánchez, Bram van Ginneken, and Jeroen van der Laak. [4].

### 3.3 Random Forest

Random Forest is a machine-learning technique that strengthens predictions by combining a large number of small decision trees. Each tree studies a different portion of the data and gives its own classification, and the final result is based on the majority opinion across all these trees. This approach reduces errors and makes the model more reliable.

In Alzheimer's research, Random Forest works particularly well because it can analyse many different types of information at the same time—such as MRI measurements, age, cognitive scores, or genetic markers. It also highlights which factors contribute most to the prediction, for example, a noticeable reduction in hippocampal volume or consistently low memory-test results. The method has been applied successfully in dementia studies, including work by Esther E. Bron, Marcel Breeuwer, Marcel J. T. Reinders, and Marleen de Bruijne, who showed that Random Forest models can enhance the accuracy of early Alzheimer's detection [6].

### 3.4 Convolutional Neural Network

A Convolutional Neural Network (CNN) is a type of deep-learning model that is particularly skilled at working with images. Instead of examining the entire image at once, it studies small sections and gradually learns patterns within them similar to how people first observe edges, shapes, and textures when looking at something. Because of this ability, CNNs are very effective in analysing brain MRI and PET scans. They can automatically pick up early changes linked to Alzheimer's, such as a reduction in hippocampal size or thinning in specific brain regions, without needing any manual feature extraction. CNNs do not need manual feature selection; they learn everything directly from the scans. Researchers like Sara Basaia, Federica Agosta, Nicola Wagner, Valentino Canu, Giancarlo Magnani, Massimo Santangelo, and Massimo Filippi used CNNs to successfully classify Alzheimer's Disease and Mild Cognitive Impairment from MRI images with high accuracy [2].

### 3.5 Recurrent Neural Network and Long Short-Term Memory ( RNN & LSTM )

Recurrent Neural Networks (RNNs) are deep learning models designed to work with data that comes in sequences or over time. They "remember" information from earlier steps and use it to understand what comes next. This makes RNNs very useful for Alzheimer's Disease because many types of patient data, such as memory test scores, speech patterns, or yearly MRI scans, change gradually over time.

However, regular RNNs sometimes forget older information too quickly. To solve this problem, a special type of RNN called Long Short-Term Memory (LSTM) was created. LSTMs can remember long-term patterns much better, making them more accurate for predicting how Alzheimer's will progress.

Researchers such as Kun Nie, Jianfeng Huang, Xiaohui Ni, Fei Wang, Ruiping Wang, Sok-Joon Ng, and Dinggang Shen used RNNs and LSTMs to track brain changes from MRI data and predict whether a patient with Mild Cognitive Impairment will develop Alzheimer's in the future [3].

### 3.6 Hybrid and Ensemble Models

Hybrid and Ensemble models combine the strengths of multiple machine learning methods to make predictions more accurate and reliable. Instead of using just one algorithm, these models use two or more algorithms together so that the weaknesses of one method can be covered by the strengths of another. For example, a hybrid model might combine CNN (for MRI images) with LSTM (for time-based data) to get both spatial and temporal information about the brain.

Ensemble models, such as combining Random Forest, SVM, and Neural Networks, help reduce errors because the final output is based on the collective decision of several models. This approach improves stability and reduces the chance of misclassification.

In Alzheimer's research, hybrid and ensemble methods are widely used because the disease affects many parts of the brain and requires analysing different types of data MRI, PET scans, speech, memory scores, and genetic information.

Researchers like Longfei Zhou, Changshui Zhang, and Shiguang Shan showed that combining multiple data sources and multiple algorithms leads to better accuracy in early detection and progression prediction of Alzheimer's Disease [11].

| Algorithm | Use/Focus Area | Main Strength | Limitation |
|---|---|---|---|
| **SVM** | Classifies MRI/PET data into normal or Alzheimer’s | High accuracy with small datasets | Needs manual feature selection |
| **Decision Tree** | Simple decision-based diagnosis | Easy to understand and interpret | May overfit; less accurate alone |
| **Random Forest** | Combines many trees for classification | More accurate; reduces overfitting | Slower and less interpretable |
| **CNN** | Analyzes MRI/PET brain images | Detects hidden brain patterns | Needs large data; high computation cost |
| **RNN/LSTM** | Works with time-based data (EEG, progression) | Good for sequential prediction | Complex; requires more computation |
| **Hybrid Models** | Combines ML and DL techniques | High accuracy and robustness | Complex design; needs large datasets |

**Table 1:** Comparison of Common Machine Learning Techniques Used for Alzheimer’s Disease Detection

**4. Feature Extraction and Data Sources**

Feature extraction means selecting the most useful information from large and complex data so that machine learning models can understand it easily. In Alzheimer’s Disease research, MRI images contain many details, but only features such as hippocampal volume, cortical thickness, ventricle size, and gray matter density are directly linked to brain changes. Deep-learning models such as CNNs can automatically learn these important patterns directly from brain scans. They do not require hand-crafted features, as the model identifies the relevant structures on its own. This approach to MRI-based feature extraction was effectively used by Basaia, Agosta, Wagner, Canu, Magnani, Santangelo, and Filippi, who were able to distinguish Alzheimer’s patients from those with Mild Cognitive Impairment with strong accuracy [2].

Feature extraction also plays a key role when working with speech recordings and clinical information. People in the early stages of Alzheimer’s often show noticeable changes in the way they speak, such as longer pauses, slower responses, or difficulty finding the right words, and these speech behaviours can be converted into measurable features. This idea was demonstrated in the work of Sardo, Proverbio, Chiera, and Di Nuovo [7]. Clinical information works in a similar way. Details such as MMSE scores, age, past medical conditions, and genetic markers can help identify individuals who are more likely to develop Alzheimer’s. Nori, Caruana, Lou, and Koch applied this feature-based approach to dementia prediction using electronic health records [16]. Choosing the most meaningful features from speech or clinical data greatly improves the accuracy and dependability of Alzheimer’s detection. Alzheimer’s research uses different types of data because the disease affects memory, behaviour, and brain structure. MRI scans, which show the brain’s structure, are widely used to measure shrinkage in important areas like the hippocampus. Studies such as those by Nie, Huang, Ni, Wang, Ng, and Shen used MRI for predicting Alzheimer’s progression [3]. PET scans show harmful proteins like amyloid plaques and tau tangles. Tau-PET imaging work by Chen, Reiman, Su, Fleisher, and Alexander demonstrated how PET can detect chemical changes linked to Alzheimer’s [13].

Many research studies use public datasets to ensure reliable and standardised data. Two of the most commonly used datasets are ADNI (Alzheimer’s Disease Neuroimaging Initiative) and OASIS (Open Access Series of Imaging Studies). ADNI provides MRI, PET, genetics, and clinical test scores for thousands of participants and is widely used to study early detection and disease progression [5]. OASIS offers open-access MRI scans across different age groups and helps compare healthy ageing with Alzheimer related brain changes, as described by Marcus, Wang, Csernansky, Morris, and Buckner [10]. These datasets, along with speech recordings, cognitive test scores, genetic information like APOE-ε4, and wearable sensor data used in research by Del Din, Godfrey, Galway, and Rochester [15], allow researchers to build accurate and comprehensive machine learning models. Multimodal datasets combining MRI, PET, genetics, and clinical tests, such as those studied by Zhou, Zhang, and Shan, further improve Alzheimer’s prediction accuracy [11].

**5. Challenges and Ethical Considerations**

AI and machine learning still face several obstacles in Alzheimer’s research. One of the biggest problems is the limited supply of high-quality data, especially early-stage MRI and PET scans, which are expensive to produce and not easy to gather in large numbers. Another difficulty comes from the nature of the disease itself. Alzheimer’s progresses slowly and affects the brain in complex ways, so models need large, diverse datasets to reliably identify the earliest changes. Many AI systems also behave like “black boxes,” providing predictions without showing how they were reached, which makes doctors less confident about using them in real clinical settings. In addition, a model trained in one hospital may not perform well in another because of differences in scanners, patient populations, or data quality.

There are also important ethical concerns. Patient privacy has to be protected, as information from brain scans, genetic tests, and medical records is extremely sensitive. Gaining informed consent can be challenging, particularly when patients have memory problems or cognitive decline. Bias is another issue; if the training data does not represent different age groups, genders, or communities fairly, the model may produce inaccurate or unequal results. Early diagnosis itself can bring emotional stress, since being told you may develop Alzheimer's can create fear or anxiety. Because of these risks, AI tools should be used to assist doctors rather than replace them, and their results must be interpreted carefully and responsibly.

**6. Future Directions**

Future work in Alzheimer's detection and prediction is moving toward more advanced AI systems that can interpret several types of patient information at once. Instead of relying on a single source, researchers aim to combine MRI, PET scans, speech features, genetic markers, and daily activity patterns into one multimodal framework. Bringing these different data types together can help identify the disease much earlier than current techniques allow. Another key goal is to design AI models that are easier for doctors to interpret. By making the decision-making process more transparent, clinicians will be able to understand why a model gives a certain result, which will increase confidence and encourage wider use in hospitals.

Personalisation is also becoming an important focus. Rather than producing the same prediction for everyone, future systems will look at each individual's unique brain characteristics, lifestyle, and genetic background to provide more tailored assessments. Wearable devices and mobile apps are expected to contribute significantly by tracking sleep patterns, movement, and speech in everyday environments, allowing continuous monitoring and early warning signs. Looking ahead, AI may also assist in creating personalised treatment strategies, identifying which patients are likely to respond well to certain therapies, and offering additional support to caregivers. Overall, upcoming advances aim to make Alzheimer's detection earlier, more accurate, and accessible for a wider population.

**7. Conclusion**

Artificial Intelligence and Machine Learning are increasingly shaping the way Alzheimer's Disease is identified and monitored. By examining information from brain scans, speech behaviour, clinical assessments, genetic data, and everyday activity patterns, these systems can pick up subtle changes that may not be noticed during regular medical examinations. Instead of replacing doctors, these technologies support them by offering quicker, more consistent, and often more accurate insights. Even though issues like limited datasets, privacy protection, and the need for clearer, more interpretable models still exist, the role of AI in Alzheimer's research continues to expand.

As data quality improves and models become more reliable and easier to understand, AI-based tools are expected to play a major role in future diagnosis and long-term monitoring. Combining different types of patient data, making use of wearable devices, and building personalised prediction systems will help catch the disease earlier and provide timely care. In the long run, AI has the potential to improve early detection, guide better treatment decisions, and enhance the overall quality of life for people who are at risk of developing Alzheimer's Disease.

**References:-**

[1] **John B. Toledo, David B. Arnold, Kaycee M. Bjerke, Manuela C. Cerebro, Lauren E. Shaw**, "Nonlinear biomarkers for Alzheimer's disease detection," *Neurobiology of Aging*, vol. 36, no. 12, pp. 2731–2741, 2015.
[2] **Sara Basaia, Federica Agosta, Nicola Wagner, Valentino Canu, Giancarlo Magnani, Massimo Santangelo, Massimo Filippi**, "Automated classification of Alzheimer's disease and mild cognitive impairment using a single MRI and deep neural networks," *NeuroImage: Clinical*, vol. 21, p. 101001, 2019.
[3] **Kun Nie, Jianfeng Huang, Xiaohui Ni, Fei Wang, Ruiping Wang, Sok-Joon Ng, Dinggang Shen**, "Deep learning for Alzheimer's disease progression prediction using MRI data," *IEEE Transactions on Medical Imaging*, vol. 39, no. 12, pp. 4030–4042, 2020.
[4] **Geert Litjens, Thijs Kooi, Babak Ehteshami Bejnordi, Arnaud A. A. Setio, Francesco Ciompi, Mohsen Ghafoorian, Jeroen A. W. M. van der Laak, Bram van Ginneken, Clara I. Sánchez**, "A survey on deep learning in medical image analysis," *Medical Image Analysis*, vol. 42, pp. 60–88, 2017.
[5] **Michael W. Weiner, Paul S. Aisen, Clifford R. Jack Jr., Christopher Ching, Danielle Harvey**, "Alzheimer's Disease Neuroimaging Initiative (ADNI): an update," *ADNI Database*, 2020.
[6] **Esther E. Bron, Marcel Breeuwer, Marcel J. T. Reinders, Marleen de Bruijne**, "Standardized evaluation of algorithms for computer-aided diagnosis of dementia," *Medical Image Analysis*, vol. 44, pp. 98–115, 2018.
[7] **Maria L. Sardo, Claudia Proverbio, Luca Chiera, Alessandro Di Nuovo**, "Speech biomarkers for Alzheimer's detection using deep neural networks," *IEEE Access*, vol. 9, pp. 65436–65448, 2021.
[8] **Morteza Rezaei, Ali Ghaffarian, Saeid Jafari, Hossein Nezamabadi**, "Explainable artificial intelligence models for Alzheimer's prediction," *Frontiers in Neuroscience*, vol. 15, pp. 1–13, 2021.
[9] **Mert R. Sabuncu, Yonatan W. Adams, Andrew J. Saykin, Stephanie A. Schultz**, "The dynamics of Alzheimer's disease biomarkers," *Brain*, vol. 141, no. 7, pp. 2111–2123, 2018.
[10] **Daniel S. Marcus, Tsung-Heng Wang, Bradley Csernansky, John G. Morris, Randy L. Buckner**, "Open Access Series of Imaging Studies (OASIS): Cross-sectional MRI data," *Journal of Cognitive Neuroscience*, vol. 22, no. 12, pp. 2677–2684, 2010.
[11] **Longfei Zhou, Changshui Zhang, Shiguang Shan**, "Multimodal fusion models for early Alzheimer's prediction," *Pattern Recognition*, vol. 100, p. 107104, 2020.
[12] **Pierrick Coupé, Vladimir S. Fonov, Louis Collins, D. Louis Collins**, "Detection of hippocampal atrophy in Alzheimer's disease based on MRI," *NeuroImage*, vol. 58, pp. 598–611, 2011.
[13] **Kewei Chen, Eric M. Reiman, Yi Su, Adam Fleisher, Gene Alexander**, "Tau PET imaging and quantification using deep learning," *Journal of Nuclear Medicine*, vol. 61, no. 7, pp. 1124–1131, 2020.
[14] **Arno König, Simone Satt, Fabienne Sorin, Tanja Konig, Kees Jan van den Bosch**, "Automatic speech analysis for the assessment of cognitive impairment," *Dementia and Geriatric Cognitive Disorders*, vol. 40, pp. 154–164, 2015.
[15] **Silvia Del Din, Alan Godfrey, Rosie J. Galway, Lynn Rochester**, "Gait analysis using wearable sensors in Alzheimer's and mild cognitive impairment patients," *Sensors*, vol. 19, no. 19, p. 4221, 2019.
[16] **Harsh Nori, Michael Caruana, Yifan Lou, Paul Koch**, "Interpretable machine learning models for predicting dementia from electronic health records," *Nature Aging*, vol. 1, pp. 111–118, 2021.