COMPUTOLOGY: Journal of Applied Computer Science and Intelligent Technologies
Volume 4, Issue 1, Jan-Jun 2024, pp. 20-40
DOI: 10.17492/computology.v4i1.2402
https://www.journalpressindia.com/cjacsit

Cognitive Impairment Diagnosis - A Game Theoretic Approach of Cognitive Machine Learning

Shardul Shastri, Sriya Mukkavilli** and Hetal Thaker****

ABSTRACT

Machine Learning (ML), a subset of Artificial Intelligence (AI) is used to simulate human-like intelligence in machines. Cognitive Machine Learning is a specialised branch of ML that focuses on developing algorithms to induce human-like reasoning and decision- making without explicit programming. Game theory explores mathematical models, used with statistical and behavioural analysis to predict outcomes in a social situation. Game theory can aid AI/ML training models with strategy identification and decision-making. To leverage insights from game theory and cognitive psychology, the analysis aims to emulate human-like decision-making and reason In AI systems. In this research paper, we analyse existing models, particularly focusing on Logistic Regression and Random Forest Algorithm to find an early diagnosis of cognitive impairment. By performing the comparative analysis, we aim to optimise the model selection process. New possibilities for intelligent systems are explored in this paper, demonstrating the synergy between AI, game theory, and cognitive psychology in tackling real-world challenges like cognitive impairment diagnosis.

Keywords: *Game Theory, Machine Learning, Random Forest Algorithm, Cognitive Impairment Diagnosis, Logistic Regression.*

1.0 Introduction

With technological advancement and improved technological advancement and improved healthcare facilities, the Average Lifespan of people has increased significantly.

**Corresponding author; BCA Student, School of Computer Science, Engineering and Applications, D. Y. Patil International University, Pune, Maharashtra, India (E-mail: shastrishar@gmail.com)*
***BCA Student, School of Computer Science, Engineering and Applications, D. Y. Patil International University, Pune, Maharashtra, India (E-mail: sriya.mukkavilli@gmail.com)*
****Assistant Professor, School of Computer Science, Engineering and Applications, D. Y. Patil International University, Pune, Maharashtra, India (E-mail: hetal.jica@gmail.com)*

This coupled with decreasing fertility rates has resulted in a rising population in the elderly population. This ageing of the population is expected to be accompanied by a dramatic increase in the prevalence of dementia. (Ravindranath & undarakumar, 2021). Some studies predict that by 2050, around 19% of the Indian population will consist of people above the age of 60. Cognitive Impairment disorders such as Alzheimer's and other kinds of dementia are prevalent in the elderly. The projections are specifically high in South Asian countries such as India and China. (Revathi *et al.*, 2022)

Dementia is marked by the decline of cognitive functions in humans. It may lead to loss of memory, reasoning, and judgement and impact an individual's capability to perform everyday tasks. WHO studies estimate that over 50 million individuals globally have dementia. This number is projected to grow every year with increments of 10 million new diagnoses.

Most research on dementia primarily focuses on treatment and care after the diagnosis of the disease. However, the progression of this degenerative disease can be prevented or delayed with early diagnosis.

Current analysis and diagnosis of these tests depend upon the inclinations of the psychologist and hence, human error cannot be avoided. A machine-learning approach can remove these inconsistencies significantly.

Machine Learning (ML) is a subcategory of artificial intelligence (AI) which focuses on developing algorithms which help computers to gather data and from the gathered data make predictions without vast amounts of programming. Machine Learning (ML) is an indispensable tool for revealing complex questions in the healthcare industry. ML takes advantage of unsupervised, supervised and reinforcement learning algorithms to create prediction models for early symptoms.

The concept of Machine Learning was introduced by Arthur Samuel in 1969. He was a pioneer of computer gaming and artificial intelligence During the 1950's and 1960's researchers explored many forms of machine learning such as the perceptron algorithm, introduced by Frank Rosenblatt. In the late 20th century as the rise of internet and new advance computers were introduced, new ML techniques were introduced as well ranging from deep learning to ensemble methods. Nowadays as technology continues to advance more than before, ML has become a key factor for multiple industries like economics, healthcare, transportation etc. It will keep on going through continuous evolution as eventually it will deliver more sophisticated algorithms in the future. Different ML Algorithms can be used for the purpose of early diagnosis. Some prevalent algorithms are logistic regression (LR), random forests Algorithms (RFL), and Support vector machines (SVMs). To evaluate these models, a case-study is employed along with a comparative analysis. The results can be optimized by employing more than one ML Technique. There are three stages to Dementia diagnosis (So *et al.*, 2017). After the three stages, the patients are prescribed various tests using MRI or CT scans and Blood Tests. Depending on the result, the severity is classified into one of three categories: normal, Mild Cognitive Impairment (MCI) and Severe Cognitive Impairment I.e. Dementia. The stages of dementia span from no cognitive decline to severe decline.

Game Theory is a branch of mathematics and theoretical science which provides players with the best strategic actions that players can make during a game. Depending on the individual choices made by players during a game, it provides a suitable framework to analyze potential outcomes based on moves, incentives and preferences and potential outcomes. Initially ideas of Game Theory were proposed during the 16th and 17th century but it wasn't until 1928 the birth of modern game theory was formulated. John von Neuman can be widely considered as the father of modern game theory who had published the *paper On the Theory of Games of Strategy*. After its publication in 1940 John von Neuman and Oskar Morgenstern published a book together called *Theory of Games and Economics Behavior* which provided an innovative foundation of how game theory can be used in the field of economics.

Eventually game theory expanded into various fields such as social sciences, biology and artificial intelligence. Game Theory plays an crucial role in the realm of Artificial intelligence (AI). Game environments provide a training mechanism to facilitate reinforcement learning in Machine learning. Furthermore, multi-agent systems are used with game theory to make quick and efficient decisions. Game theory can be applied to navigate complex situations which can lead to more robust and well-planned solutions in real world applications.

One such concept of Game Theory is Combinatorial Games. These games are straightforward games where players have all the information they need, there's no luck involved, and someone always wins or loses. It is a turn-by-turn system where the positions set, starting from the initial position determines the winner. By using a similar idea, one can classify the types of dementia more efficiently. Using the Random Forest classifier to move through different possibilities, we can speed up the process of dementia type-detection and its severity in a patient more efficiently. A study by Wei Ying Tan highlighted the use of a ML Approach for Early Diagnosis of Cognitive Impairment Using Population-Based Data. In this, Tan highlighted the viability of using multiple domains of data for reliable diagnosis of early cognitive impairment.

By taking Tan's model as our case study, we aim to build upon this model and analyze it with other learning algorithms. This approach evaluates the modified model along with combinatorial game theory and decision trees. To achieve this, we will utilize 50% of available data for training and negative reinforcement, while the remaining 50% will be reserved for testing purposes. This analysis seeks to highlight strengths and weaknesses of each approach in early diagnosis, thereby contributing to healthcare diagnostics.

2.0 Review of Literature

There is limited research on Cognitive impairment diagnosis since it is a relatively new field of study in the healthcare industry. Only in 2003 was the term Mild Cognitive Impairment (MCI) introduced. As ML techniques are still developing, it has become a highly relevant and growing field of study. Advanced Machine Learning on Cognitive Computing for Human Behavior Analysis by Zhihan Lv, Liang Qiao and Amit Kumar Singh propose the idea of using clustering for diagnosis.

The gradient-boosted tree ensemble method is used by Robert P. Adelson to Improve Longitudinal Prediction of Progression from Mild Cognitive Impairment to Alzheimer's Disease (Adelson *et al.*, 2024). Game theory is proposed by many as a prediction tool, however, there is limited research conducted on its applications. Using it in healthcare diagnosis is a fresh perspective that needs to be explored further.

3.0 Methodology

This study analyses previously published works and uses them as a foundational base to build upon our theoretical models. This can be perfectly achieved by following a case study approach. By using this method, we aim to assess how well the newly developed models may theoretically perform in real-world scenarios. Through this methodology, existing models can be compared with each other to identify the strengths and weaknesses of each approach.

Our proposed theoretical model uses methodologies of random forest numbers with combinatorial game theory and positive reinforcement learning. In the first phase of investigation, existing literature was explored to deem relevant factors of cognitive science that potentially can be included with our model to potentially enhance traditional diagnostics models. In Phase Two, Combinatorial Game Theory (CGT) and Positive Reinforcement are added to further improve accuracy. Visual presentation of concepts is included to further facilitate our interpretation.

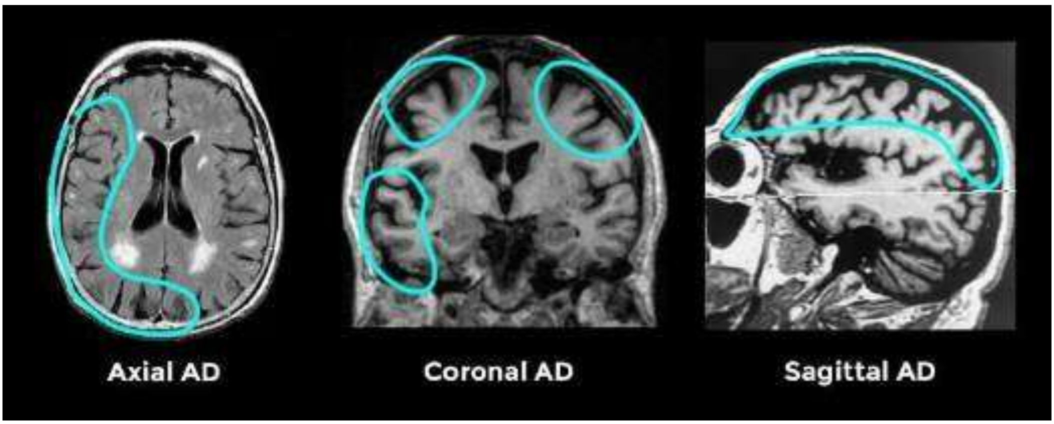
3.1 Dementia classification

Dementia is a clinical diagnosis requiring new functional dependence based on progressive cognitive decline (*Dementia - PMC [1]*, n.d.). Dementia is an umbrella term that is used to define a group of degenerative neurodegenerative disorders which result in a cognitive functioning decline. It reduces the ability of an individual to reason and worsens with age. The most commonly found instance of dementia is Alzheimer's Disease. When considering several other symptoms and parts of brain affected, a patient can have other types of dementia present. With this, the type of dementia can be classified into 4 main types (Dauwan *et al.*, 2016).

3.1.1 Alzheimer's disease

Alzheimer's disease (AD) is a disorder that causes degeneration of the cells of the brain, which is characterised by a decline in thinking and independence in personal daily activities (Breijyeh & Karaman, 2020). It is one of the most common instance of dementia, widely affecting persons over the age of 65. As the damage spreads, cells of the brain lose their ability to function and may eventually die. This causes irreversible harm, the longer it goes undiagnosed.

Figure 1: MRI Scans of Alzheimer's Affected Areas



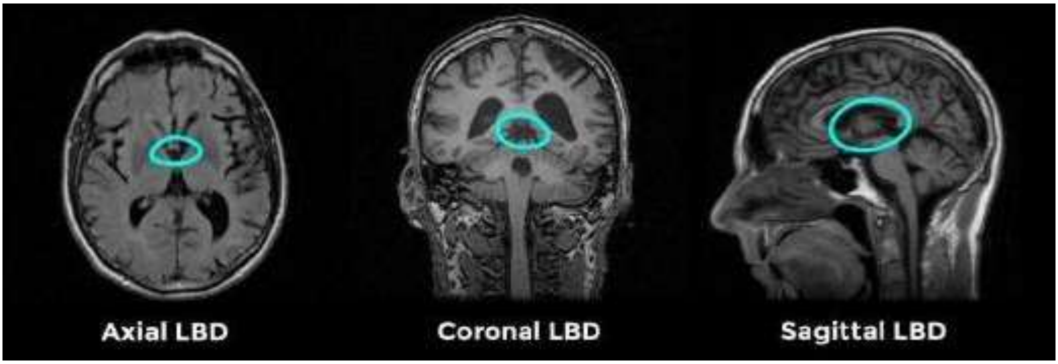
Source: Scheltens P. (2009). Imaging in Alzheimer's disease.

This destruction and failure of neurons leads to memory failure, changes in personality and problems carrying out other daily activities. With ageing, it is quite normal for the brain to shrink, while not necessarily signifying loss of neurons. In contrast, Alzheimer's damages connections among neurons in parts of brain such as the hippocampus and entorhinal cortex (Figure 1). Most studies show a tendency to develop plaques and tangles with age. In the case of Alzheimer's, the development is far more widespread and a predictable pattern

3.1.2 Lewy body dementia

Lewy Body is a type of protein called alpha-synuclein. Lewy body dementia (LBD) is caused by abnormal deposits of Lewy Bodies in the brain. LBD affects the chemical balance of the brain leading to unpredictable mood, behavior, and movement. Diagnosis of LBD is challenging as in early stages, symptoms are mild however, as it advances, a decline in thinking and rationale is observed. Dementia with Lewy bodies (DLB) is thought to account for up to 20% of cases of dementia (Frank, 2003).

Figure 2: MRI Scans of Lewy Body Dementia Affected Areas



Source: Gaillard F, Lewy body dementia. Case study, Radiopaedia.org

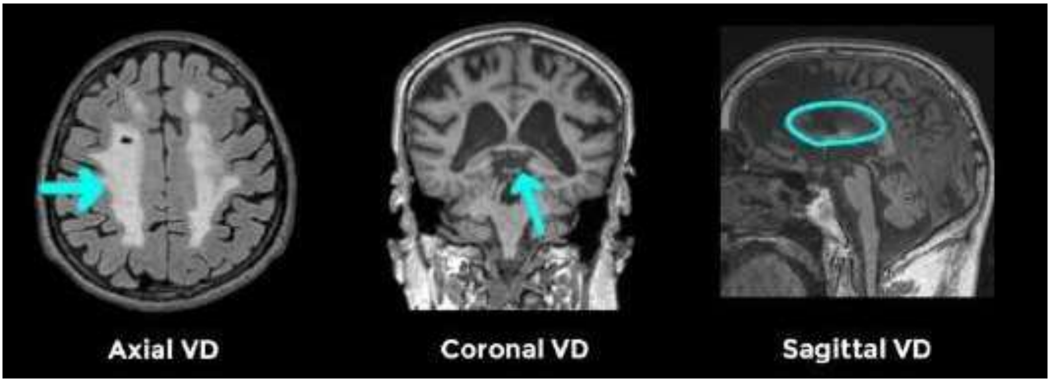
LBD may affect different regions of the brain. It is most noticeable in Cerebral cortex which is responsible for perception, language, and information processing (Figure 2)

3.1.3 Vascular dementia

Vascular Dementia is triggered by irregular blood flow to the brain. Due to a series of minor strokes, a decline in cognitive abilities of the brain is noticed. It is a mix of cerebrovascular illnesses that cause structural changes of brain. . It may result from multiple cortical infarctions due to cerebral large vessel pathologies or to subcortical ischemic changes such as leukoaraiosis (Lee, 2011).

Unlike Alzheimer’s, Vascular Dementia does not affect the personality of the patient. They show normal levels of emotional responsiveness even in the latter stages of the disease. MRI scans may show affected areas and blood clots (Figure 3).

Figure 3: MRI Scans of Vascular Dementia Affected Area

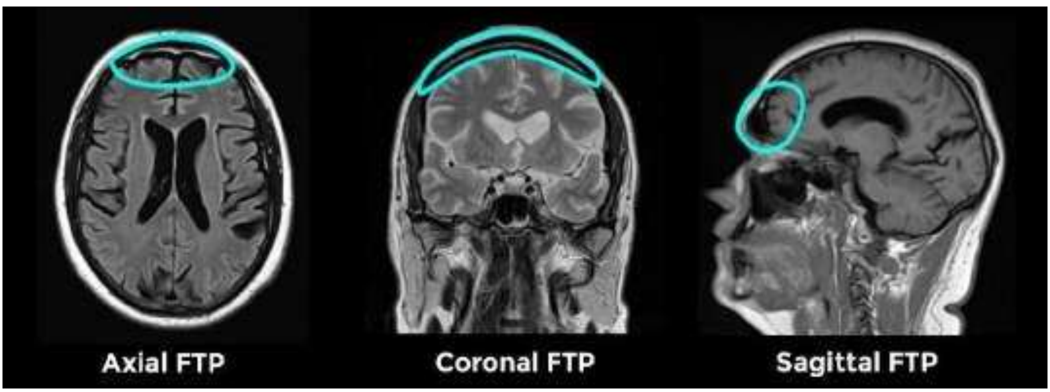


Source: Matta G, Vascular dementia. Case study, Radiopaedia.org

3.1.4 Frontotemporal dementia

Frontotemporal dementia (FTD) mainly affects the temporal and frontal lobes of the brain, hence the name. These parts of brain are associated with personality, behaviour, and language. FTD is quite like Alzheimer’s.

Figure 4: MRI Scans of Frontotemporal Dementia Affected Areas



Source: Gaillard F, Frontotemporal dementia. Case study, Radiopaedia.org

It occurs in the front lobe of the brain (Figure 4). As it advances, parts of the lobes shrink. This is known as Atrophy. Symptoms of the patient may depend on the affected areas. FTD was previously thought to be a rare disease. However, researchers report that FTD is the third most common form of dementia (Cycyk & Wright, 2008). There is very little literature available on automated discrimination between the types of dementia, with ensemble ML models such as RFLs this is made possible. We further optimise the computation by opting for a Combinatorial Game theoretic approach.

3.2 Tan's ensemble model

Ensemble methods refer to the aggregation of two or more methods to improve accuracy and optimize results. Tan utilizes these methods along with population-based datasets taking age, ethnicity, education, and neuroimaging markers as predictors to determine Cognitive Impairment. This model achieved an F1 score of 0.87 and AUC and 0.80 with a sensitivity of 0.86. Instead of using population-based data with MR imaging data, the results can be further optimized. This will also increase the scope of study beyond the multi-ethnic Asian population.

3.3 Logistic regression

Logistic Regression (LR) is a supervised-based ML technique used for classification of tasks. It is an important statistical and data mining technique employed by statisticians and researchers for the analysis and classification of binary and proportional response data sets (Maalouf, 2011). It analyses the relationship between two different variables. It is mainly used as binary classification where it takes two different input variables and produces a probability between 1 and 0. The logistic function was invented in the 19th century for the description of the growth of populations and the course of autocatalytic chemical reactions (Cramer, 2002).

3.4 Random forests algorithm

Random Forest Classifiers (RFCs) are used as ensemble ML algorithms that are constructed based on decision tree algorithms. Instead of relying on a single decision tree, a collection of trees (i.e. forest) is used to predict. Random forests are a scheme proposed by Leo Breimanq for building a predictor ensemble with multiple decision trees that grow in randomly selected subspaces of data (Biau, 2012). RFCs use hyperparameters to enhance performance. Hyperparameter tuning is the process of selecting features and datasets that can be passed through this model for training purposes. Hyperparameters are like settings that are yet to be learned by the learning model. They directly affect accuracy of models and their computational efficiency. Some factors to consider are the number of individual decision trees, the maximum depth of each tree and the size of the Random Feature subset selection. In this context, the MRI images that store data regarding lobes of the brain are hyperparameters.

3.5 Comparing Logistic Regression with Random Forest

Table 1: Difference between Logistic Regression and Random Forest Algorithm

Random Forest Algorithm	**Linear Regression**
More accurate than Linear Regression with large datasets	Less accurate than Random Forest with large datasets
Less prone to outliers (false positives)	More prone to outliers (false positives)
Less prone to overfitting	More prone to overfitting
Stores data in non-linear relationship which helps to model complex data	Stores data in linear relationships which can cause it to oversimplify complex data
Prediction time of Random Forest is slower than Linear Regression with large datasets	Prediction time of Linear Regression is faster than Random Forest with large datasets
Space-Time Complexity is O(mnlog(n)) m- number of branches n- depth of tree	Space-Time Complexity is O(n2)

3.6 Tree pruning

RFL is optimal for classifying large datasets, however, it does so by demanding significant time and computational resources. As the model grows more complex, it risks overfitting and becomes difficult to interpret and less efficient. In order to tackle this problem, it becomes necessary to explore algorithms that can optimize the model without sacrificing accuracy. Tree Pruning is one such technique. Size of the decision tree is reduced by removing redundant branches which are unlikely to result in optimal solutions, therefore increasing the overall efficiency. A pruned tree is easier to interpret and reduces the risk of overfitting. In simpler words, pruning can optimize RFL models by improving the computational focus. To further strategize the Pruning Technique, some principles of Game Theory can be applied.

3.7 Combinatorial game theory

Combinatorial Game Theory (CGT) is a branch of mathematics and theoretical sciences with focus on sequential games that have perfect information. CGT first came into the spotlight concerning impartial games. Impartial Games states that all players must always have all information present. Games which come under Combinatorial Game Theory are known as Combinatorial Games, in which two players take turns making moves both have complete information about what has happened in the game so far and what each player's options are from each position (Crash Course on Combinatorial Game Theory, n.d.). Few examples of Combinatorial games are chess, tic-tac-toe, checkers etc. By considering single decision trees as 'players' and each traversal as the 'moves', the logic behind CGT can be applied to RFL.

3.8 Reinforcement learning

Reinforcement Learning (RL) is a decision-making machine learning algorithm. Its primary focus is on learning optimal behaviour to achieve maximum reward. It's a way of training where learning outcomes determine what action needs to be taken. A Sample of works that applied reinforcement learning to natural language processing includes DeepMind work using reinforcement learning to compute representations of natural language sentences by learning tree-structured neural networks (Hammoudeh & Rlr, n.d.). Reinforcement Learning has come a long from the 1950s till the 1990s which was a time of theoretical advancements for this algorithm to the 2010s when it began being used in the field of business. There are two ways to achieve Reinforcement Learning:

3.8.1 Negative reinforcement learning

Negative Reinforcement Learning (NRL) is a reinforcement-based Machine Learning Algorithm which involves avoiding the frequency of unpleasant behaviour to decrease the likelihood of that behaviour in the future. The process of negative reinforcement typically involves the removal, reduction, postponement, or prevention of stimulation these operations strengthen the response on which they are contingent (Iwata, 2007).

3.8.2 Positive reinforcement learning

Positive Reinforcement Learning (PRL) is a reinforcement-based ML Algorithm that is very similar to NRL with one key difference. While NRL focuses on decreasing the likelihood of unpleasant behaviour, PRL focuses on increasing the frequency of desirable behaviour to build up the likelihood of such behaviour in the future.

3.9 MRI-based imaging data and feature selection

Feature Selection is a procedure of identifying and selecting relevant data as 'features' to improve predictive analysis by focusing on the most important attributes that contain the most identifiable information. For the feature extraction the MRI scans are taken from the Axial plane, Sagittal plane, and Coronal plane. The cerebrum will be segregated into Frontal, temporal, parietal and occipital lobes for monitoring.

Alzheimer's disease is noticeable via the decay of the hippocampus and atrophy which are in the temporal lobe. Abnormal traces of alpha-synuclein when found in parietal, frontal or occipital lobes may signify presence of Lewy body dementia. Alpha-synuclein can be found as traces of soft white deposits on the parietal, frontal or occipital lobes.

To qualify for vascular dementia, Leukoaraiosis (neuroimaging of white matter) of more than 25% should be present on the frontal. This pathological appearance of white matter in the brain has been believed to be caused by perfusion disturbances within the arterioles perforating through the deep brain structures (Marek *et al.*, 2018). The shrinking of any lobe is known as atrophy. For frontotemporal dementia, the frontal or the temporal lobes of the brain shrink.

3.10 Dataset description

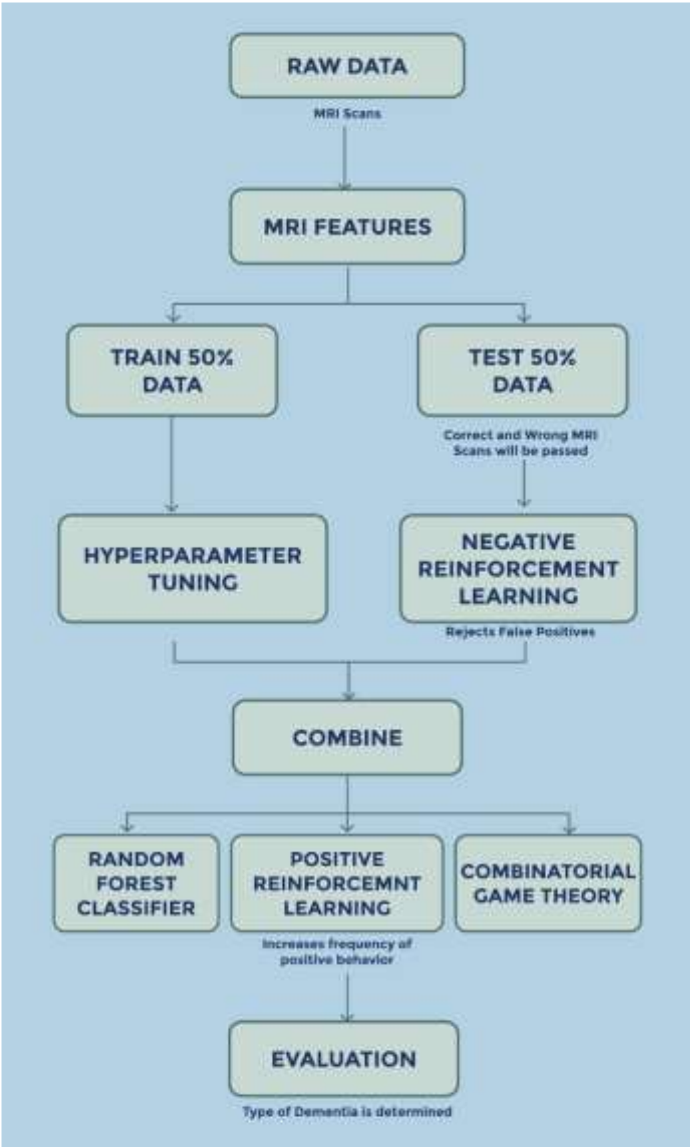
The types of dementia dataset were collected from Radiopaedia, an open-source and peer-reviewed educational radiology resource with 50,000+ cases of different types of diseases and dementia from all around the world. MRI scans collected from multiple cases and were classified into Alzheimer's disease (AD), Lewy Body Dementia (LBD), Vascular Dementia (VD) and Frontotemporal Dementia (FTD). With an average of 68.5kb, the images are available in 630x630 dimensions.

4.0 Proposed Model

4.1 Framework for Learning, Analysis and Prediction in Early Diagnostics (FLAP-ED) model

Machine Learning based decision support tools have been used in healthcare for multiple tasks which range from neurodegenerative conditions to chronic conditions. An opportunity to provide earlier treatment in case of diagnosis of any condition is made possible by these types of tools. The Framework for Learning, Analysis and Prediction in Early Diagnostics (FLAP-ED) is a Machine Learning training model which utilises these tools to accurately perform type detection of dementia a patient may suffer from. (Figure 5). It utilizes the concepts of Random Forest Algorithm (RFL), Positive Reinforcement Learning (PRL) and Combinatorial Game Theory (CGT) as its classifiers.

Figure 5: FLAP-ED Model



Source: Self

4.2 A layered approach

The model is segregated into three layers to simplify its working (Figure 6).

4.2.1 Input layer

Raw data is obtained in the form of MRI scans. Once the data is obtained features of the MRI scans are determined.

4.2.2 Hidden layer

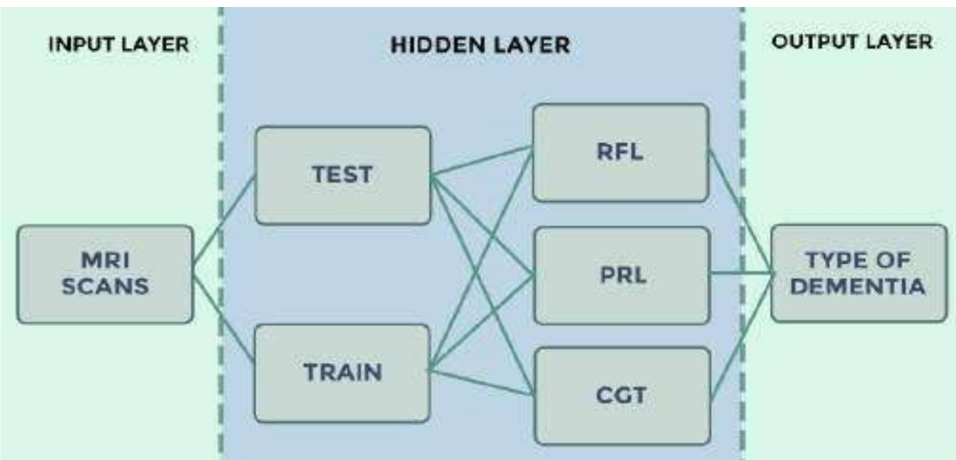
This layer is the main processing layer of the model. Once the features are determined the scans are sent for training and testing. Training and testing of data are done independently and parallel to each other. Data is split equally 50-50% for training and testing data. While training the model, hyperparameters are tuned. Scans are then sent to the 3 main classifiers, Random Forest Algorithm (RFL), Positive Reinforcement Learning (PRL) and Combinatorial Game Theory (CGT).

The FLAP-ED model utilises RFL to classify the types of dementia depending on the hyperparameters set when training the model. When RFL is in progress CGT is applied after reaching a certain level of nodes in each decision tree. This allows for the model to reach a conclusion before completing the traversal of redundant nodes

4.2.3 Output layer

Type detection is performed in this layer. When correct results are derived, positive feedback is returned to the machine, ensuring that the same behaviour is encouraged.

Figure 6: Layers of the FLAP-ED Model

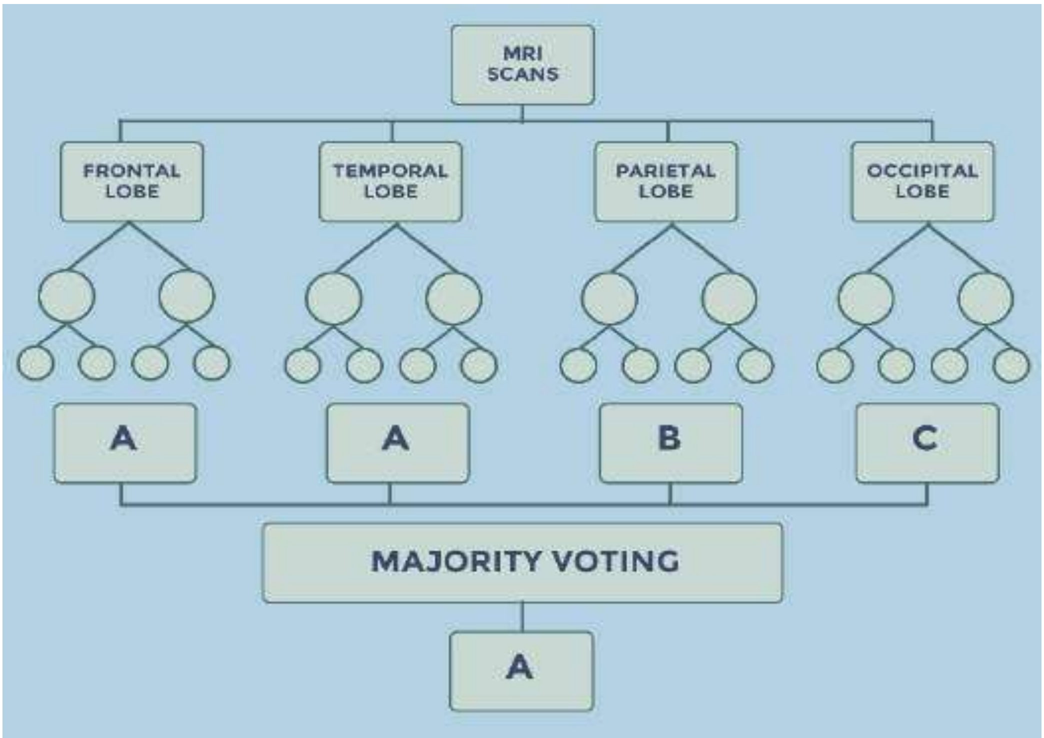


Source: Self

4.3 Type detection

Parameters inside the classifiers are set per the hyperparameters. In contrast to direct, first-level model parameters, which are determined during training, these second-level tuning parameters often have to be carefully optimised to achieve maximal performance (Probst & Bosch, 2019) (Figure 7).

Figure 7: Type Detection of Dementia using Random Forest Algorithm



Source: Self

In RFL the number of trees and the maximum depth of trees is established by the hyperparameters. Combinatorial Game Theory's focus is achieving a winning position. It will consider a single decision tree at a time present inside the RFL. A winning position is achieved when it reaches a node where it can directly determine the correct type of dementia without having to move further down the decision tree. Once the outcome is determined for the decision tree it will move onto the next decision tree and repeat the process. It will iteratively perform this step until all decision trees are executed and the majority voting begins. The result of voting is our final output. Impartial Game Logic is to be applied. In this type of game, the set of allowable moves depends only on the position of the game and not on which of the two players is moving (Li & Gymrek, 2009).

4.4 Normal play games

Normal Play games are a subset of CGT. It suggests that the player with the last viable move is the winner. The outcome must be a win or a loss, draws aren't permitted in any scenario. Even though Normal Play Game Logic is commonly used with a two-entity system, the FLAP- ED model attempts to apply it to multiple entities. Entities being the four types of dementia in context of the study.

4.4.1 Position and types

New notations are defined in this section and later classified based on different positions in the decision tree to facilitate determining the guaranteed winning positions i.e. definitive type of dementia. Positions are denoted as *a, b, c* and *d*. These positions emphasize the current node

This can be formulated as:

$$P = \{a1,\ a2,\ ...an \mid b1,\ b2,\ ...bn \mid c1,\ c2,\ ...cn \mid d1,\ d2,\ ...dn\}$$

Where,

P - Current Position

$P \rightarrow a_i$: All possible path for Alzheimer's disease

$P \rightarrow b_i$: All possible path for Vascular dementia

$P \rightarrow c_i$: All possible path for Lewy Body dementia

$P \rightarrow d_i$: All possible path for Frontotemporal dementia

P is considered the current position. From this position, there are four possible moves: a, b, c, and d. Each move corresponds to a type of dementia, and the formulas above represent these connections. For e.g. the notations a_1, a_2, ..., an indicate the possible moves of Alzheimer's disease. However, these notations differ from the winning positions. From current node the possible winning strategies are:

A → It is definitely Alzheimer's disease

B → It is definitely Vascular dementia.

C → It is definitely Lewy body dementia.

D → It is definitely Frontotemporal dementia.

E → Current position is on winning strategy.

F → Current position is not on a winning strategy.

The letters A, B, C, D, E, and F represent definitive winning strategies. For instance, while a_n represents a series of moves leading to the conclusion that Alzheimer's disease is likely, A is a winning move that definitively confirms Alzheimer's disease as the diagnosis, ruling out other types of dementia. Determination of possible strategies tells us what position the game is in. The game in this scenario is the individual decision tree which is present inside the random forest algorithm. We can tabularize possible moves from each node (Table 2).

Table 2: Decision Table to Determine the Possible Winning Positions

	a_i = B, C, D, E	b_i = A, C, D, E	c_i = A, B, D, E	d_i = A, B, C, E	All a_i = A,F	All b_i = B, F	All c_i = C, F	All d_i = D, F
a_i = A, E	F(A)	B	C	D	A	A, B	A, C	A, D
b_i = B, E	A	F(B)	C	D	A, B	B	B, C	B, D
c_i = C, E	A	B	F(C)	D	A, C	B, C	C	C, D
d_i = D, E	A	B	C	F(D)	A, D	B, D	C, D	D
All a_i = B, C, D, F	B, C, D	A, B, C, D	A, B, C, D	A, B, C, D	E	A	A	A
All b_i = A, C, D, F	A, B, C, D	A, C, D	A, B, C, D	A, B, C, D	B	E	B	B
All c_i = A, B, D, F	A, B, C, D	A, B, C, D	A, B, D	A, B, C, D	C	C	E	C
All d_i = A, B, C, F	A, B, C, D	A, B, C, D	A, B, C, D	A, B, C	D	D	D	E

4.4.2 Sum of positions

Sum of positions is calculated in each Decision Tree after it has been executed. If the starting position is $b1$ and in the next node it moves towards $b2$ then in game terms we can say that $b1$ moves to $b2$:

$P = \{b1, b2\}$

Sum of positions can also be derived in a tabular format (Table 3). Table 3 is used to determine the sum of positions. Sum of positions dictates that all results (A+B+C+D) are possible from starting position (P) i.e. the root node. The more we traverse down the tree, the clearer the type of dementia determined becomes. For example, if after the root node it traverses towards position A, which can be considered a winning position for Alzheimer's disease, then it is moving from P(A+B+C+D) to $P \rightarrow a_i$ similarly it can move to $P \rightarrow b_i$, $P \rightarrow c_i$ and $P \rightarrow d_i$. The question mark (?) in Table 03 indicates that the sum of positions is not fixed, and no type of dementia has been determined yet, hence suggesting to continue traversal until sum of positions is achieved. Our model primarily uses MRI scans as input. {a_1, a_2, a_3} correspond to different stages of the MRI data. To illustrate further:

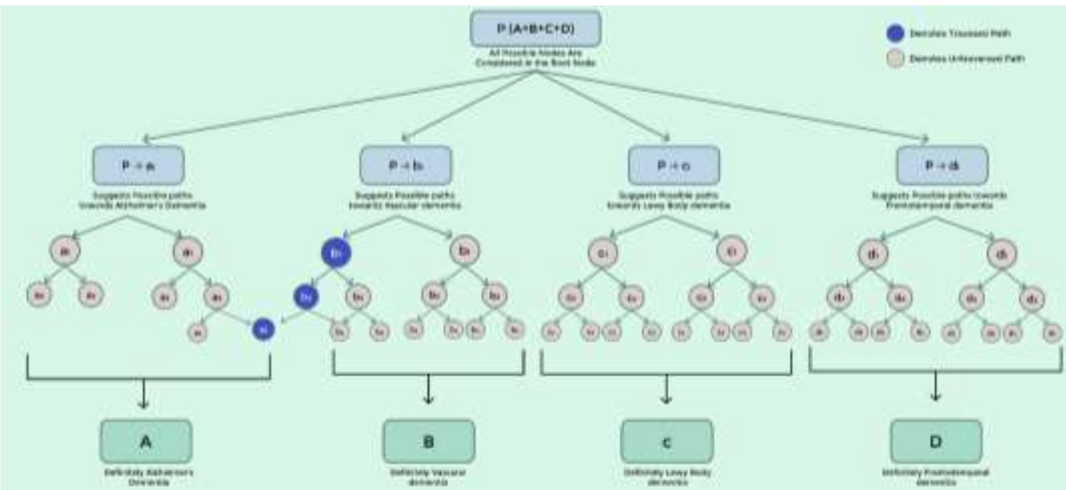
Table 3: Decision Table to Determine the Sum of Positions

	A	**B**	**C**	**D**	**E**	**F**
A	A	?	?	?	A	?
B	?	B	?	?	B	?
C	?	?	C	?	C	?
D	?	?	?	D	D	?
E	A	B	C	D	E	F
F	?	?	?	?	F	?

a_1: This is where the initial feature extraction begins. It focuses on specific regions like the hippocampus.
a_2: Might consider patient history, genetic disorders etc.

This sequence will continue as a_n will keep looking at various other factors until it can determine Alzheimer's is the definite winning position i.e. A. Figure 8 is derived to illustrate the working of a single decision tree's traversal using Decision Table (Table 2) and Sum of Positions (Table 3). It will be a Non-Binary Decision tree in our case since we are considering 4 Possible moves initially i.e. a_i, b_i, c_i and d_i.

Figure 8: Combined Game Tree of A+B+C+D



In Figure 8, the concept of decision streaming is illustrated. The traversal starts at the root node and initially follows the path $P \rightarrow b_i$. However, as the traversal progresses down the tree, the model shifts from the path $P \rightarrow b_i$ to $P \rightarrow a_i$. This adjustment leads to the conclusion that the type of dementia is Alzheimer's disease, rather than vascular dementia. Although the initial path suggested vascular dementia, the model modified its course because a possible winning position (Table 1) had not yet been determined, and the sum of positions was varying and undetermined (Table 2). As the traversal continued, the model ultimately identified Alzheimer's as the correct diagnosis, reaching a definitive winning position.

5.0 Results/Findings

Through this study, there is a new method developed to optimise the RFL workflow. Logistic Regression is a common technique which is used in the field of early diagnosis. R comes under Regression Analysis which focuses on the internal relationship between dependent and independent variables (Lv *et al.*, 2021). RFL makes voting decisions based on the decision of multiple smaller decision trees while LR uses logistic functions to conclude.

Table 4: Difference between FLAP-ED model and Tan's Model

Aspect	FLAP-ED Model	Tan's Model
Accuracy	93.97%	93.74%
Sensitivity	80.60%	76.82%
Specificity	95.29%	95.89%
Classification Types	Default Parameters	Binary Classification
Scalability	Can scale up to introduce more types of dementia	Cannot scale up because it is not determining type of dementia
Dataset Selection	MRI Scans	Multimodal Data
Model Performance	Less prone to overfitting	More prone to overfitting

Note: Accuracy, Sensitivity and Specificity values are Adapted from Han et al., (2019). Random forest can accurately predict the development of end-stage renal disease in immunoglobulin a nephropathy patients.

Both LR and RFL are types of supervised Learning, however, RFL can provide a higher accuracy than LR. LR is observed to perform better when:

1. no. of noise variables $\leq$ no. of explanatory variables
2. RFL returns high rate of true and false positive with increase in explanatory variables

Many times, datasets can be imbalanced due to which in LR data can sometimes be biased. RFL solves this issue by randomly selecting subsets to consider per split.

Drawbacks lie in its complexity when considering computational speed due to the huge amounts of data it must handle. This issue can be solved with the integration of Combinatorial Game Theory.

6.0 Conclusion

Our proposed FLAP-ED model is theoretically more efficient than Tan's model [Table 4] for early diagnosis of cognitive impairment, as we have implemented elements of Combinatorial Game Theory in Machine Learning which eliminates the complexity when considering computational speeds due to the vast data that must be handled. By incorporating NPG logic into the Random Forest Algorithm, data processing speeds are enhanced. Initially, each node of the model calculates a winning strategy. Once the strategy is determined, the conclusion is derived directly without the need to traverse through redundant nodes. Since reinforcement learning is used, favourable behaviour is encouraged when correct results are produced. This makes the model more efficient over time as it processes more data.

These Advantages of FLAD-ED Model make it scalable to larger datasets. This makes it particularly valuable in context of Big Data Environments in Healthcare Industry. By introducing it to a broader range of cognitive disorders, this model has the potential to extend beyond the four types of dementia discussed in this paper. Training FLAP-ED on Individual Patient Profiles and genetic predispositions can help it evolve into a Personalized Diagnostic Tool.

References

Adelson, R. P., Garikipati, A., Maharjan, J., Ciobanu, M., Barnes, G., Singh, N. P., Dinenno, F. A., Mao, Q., & Das, R. (2024). Machine learning approach for improved longitudinal prediction of progression from mild cognitive impairment to Alzheimer's disease. *Diagnostics*, *14*(1), 13.

Biau, G. (2012). Analysis of a random forests model. *The Journal of Machine Learning Research*, *13*(1), 1063-1095.

Breijyeh, Z., & Karaman, R. (2020). Comprehensive review on Alzheimer's disease: Causes and treatment. *Molecules*, *25*(24), 5789. Retrieved from https://doi.org/10.3390/MOLECULES25245789

Cramer, J. S. (2002). *The Origins of Logistic Regression.*

Cycyk, L. M., & Wright, H. H. (2008). Frontotemporal dementia: Its definition, differential diagnosis, and management. *Aphasiology*, *22*(4), 422–444. Retrieved from https://doi.org/10.1080/02687030701394598

Dauwan, M., van der Zande, J. J., van Dellen, E., Sommer, I. E. C., Scheltens, P., Lemstra, A. W., & Stam, C. J. (2016). Random forest to differentiate dementia with Lewy bodies from Alzheimer's disease. *Alzheimer's and Dementia: Diagnosis, Assessment and Disease Monitoring*, *4*, 99–106. Retrieved from https://doi.org/10.1016/j.dadm.2016.07.003

Frank, C. (2003). Dementia with Lewy bodies. Review of diagnosis and pharmacologic management. *Canadian Family Physician*, *49*(10), 1304-1311.

Hammoudeh, A., & Rlr, A. H. (n.d.). A concise introduction to reinforcement learning. Retrieved from https://www.studocu.com/in/document/amity-university/clinical-psychology/aconcise-introductionto-reinforcement-learning-2018-0209-research-gate/21000238

Han, X., Zheng, X., Wang, Y., Sun, X., Xiao, Y., Tang, Y., & Qin, W. (2019). Random forest can accurately predict the development of end-stage renal disease in immunoglobulin a nephropathy patients. *Annals of Translational Medicine*, *7*(11).

Iwata, B. (2007). Negative reinforcement. Retrieved from https://homepages.se.edu/cvonbergen/files/2013/01/Negative-Reinforcement-in-Applied-Behavior-Analysis_An-Emerging-Technology.pdf

Lee, A. Y. (2011). Vascular Dementia. *Chonnam Medical Journal*, *47*(2), 66. Retrieved from https://doi.org/10.4068/cmj.2011.47.2.66

Li, J., & Gymrek, M. (2009). Theory of impartial games. *Massachusetts Institute of Technology (MIT)*, *3*.

Lv, Z., Qiao, L., & Singh, A. K. (2021). Advanced machine learning on cognitive computing for human behavior analysis. *IEEE Transactions on Computational Social Systems*, *8*(5), 1194–1202. Retrieved from https://doi.org/10.1109/TCSS.2020.3011158

Maalouf, M. (2011). Logistic regression in data analysis: An overview. In *International Journal of Data Analysis Techniques and Strategies,* 3(3), 281–299. Retrieved from https://doi.org/10.1504/IJDATS.2011.041335

Marek, M., Horyniecki, M., Frączek, M., & Kluczewska, E. (2018). Leukoaraiosis - New concepts and modern imaging. In *Polish Journal of Radiology* (Vol. 83, pp. e76–e81). Medical Science International. Retrieved form https://doi.org/10.5114/pjr.2018.74344

Probst, P., & Bischl, B. (2019). Tunability: Importance of Hyperparameters of Machine Learning Algorithms. In *Journal of Machine Learning Research* (Vol. 20). Retrieved from http://jmlr.org/papers/v20/18-444.html.

Ravindranath, V., & Sundarakumar, J. S. (2021). Changing demography and the challenge of dementia in India. *Nature Reviews Neurology, 17*(12), 747–758. Retrieved from https://doi.org/10.1038/s41582-021-00565-x

Revathi, A., Kaladevi, R., Ramana, K., Jhaveri, R. H., Rudra Kumar, M., & Sankara P.K.M. (2022). Early detection of cognitive decline using Machine Learning algorithm and cognitive ability test. *Security and Communication Networks*. Retrieved from https://doi.org/10.1155/2022/4190023

So, A., Hooshyar, D., Park, K. W., & Lim, H. S. (2017). Early diagnosis of dementia from clinical data by machine learning techniques. *Applied Sciences, 7*(7), 651.