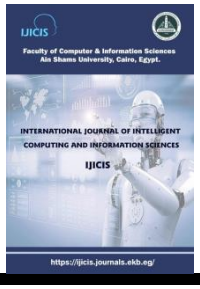
IJICIS, Vol.22, No.2, 97-111
DOI: 10.21608/ijicis.2022.116420.1153

International Journal of Intelligent Computing and Information Sciences

https://ijicis.journals.ekb.eg/



AN EFFICIENT HYBRID APPROACH FOR DIAGNOSIS HIGH DIMENSIONAL DATA FOR ALZHEIMER'S DISEASES USING MACHINE LEARNING ALGORITHMS

Nour Zawawi*

Department of information Systems,Faculty of Computers and Information Sciences, Ain Shams University, Cairo, Egypt
October University for Modern Sciences and Arts, Giza, Egypt
nour.zawawi@gmail.com

Heba Gamal Saber

Geratic Mediane Department,Faculty of Medicine, Ain Shams University, Cairo, Egypt
hebageasaber@gmail.com

Mohamed Hashem

Department of information Systems,Faculty of Computers and Information Sciences, Ain Shams University, Cairo, Egypt
mhashem100@yahoo.com

Tarek F.Gharib

Department of information Systems,Faculty of Computers and Information Sciences, Ain Shams University, Cairo, Egypt
tfgharib@cis.asu.edu.eg

Received 2022-01-15; Revised 2022-04-01; Accepted 2022-04-05

Abstract: *Alzheimer's disease (AD) is the most familiar type of dementia, a well-known term for memory loss and other cognitive disabilities. The disease is dangerous enough to interfere with ordinary life. Identifying AD in the early stages remains an extremely challenging task, meanwhile, the progression of it develops several years before observing any symptoms. The fundamental issue addressed during diagnosis is the high dimensionality of data. However, not all features are relevant for solving the problem, and sometimes, including some irrelevant ones may deteriorate the learning performance. Therefore, it is essential to do feature reduction by selecting the most relevant features.*
In this work, a hybrid approach Random Forest Partial Swarm Optimization(RF-PSO) for high-dimensional feature selection is proposed. The fundamental reason behind this work is to support geriatricians diagnose AD; by creating a clinically translatable machine learning approach. The dataset created by Alzheimer's Disease Neuroimaging Initiative (ADNI) was used for this purpose. The ADNI dataset contains 900 patients whose diagnostic follow-up is available for at least three years after the baseline assessment. The reason behind choosing is their strength in solving large-scale optimization problems with high data dimensionality. The Experiments show that RF-PSO outperforms most of the others found in the literature. It achieved high performance compared to them. The accuracy rate of this approach reached 95% for all the AD stages. In a comparison with Random Forest which achieve 86%, While Partial Swarm Optimization got 89%.

Keywords:*Alzheimer's disease, machine learning, feature selection, high dimensional*

***Corresponding Author**: Nour Zawawi

Department of information Systems,Faculty of Computers and Information Sciences, Ain Shams University, Cairo, Egypt

October University for Modern Sciences and Arts, Giza, Egypt

Email address: nour.zawawi@gmail.com

1. Introduction

Dementia is a broad word that refers to various symptoms that affect memory, daily activities, and communication. The most frequent type of dementia is Alzheimer's disease (AD). It worsens over time and impacts memory, language, and thought. As a result, increasing memory loss characterizes this neurodegenerative disease. It is responsible for more than 60% of dementia cases [1, 2]. Its victims frequently have a slew of symptoms. They include memory loss that worsens with time, linguistic problems, and disorientation. In general, there are multiple stages of Alzheimer's disease. The three stages are early, middle, and late (sometimes known as mild, moderate, and severe in medical terms) [3]. One of the significant concerns among experts on the early diagnosis of Mild Cognitive Impairment (MCI) is a stage in between health and Alzheimer's disease (AD). It demonstrates the possibility of continued progression toward Alzheimer's disease or other dementias. Although it does not affect daily activities, it is unusual for someone of this age and educational level. As a result, it does not fit the criteria for Alzheimer's disease.

According to recent research, only 20–40% of individual cases will transition to AD within three years [1]. It is a lower rate of exchange recorded in medical samples than in clinical samples. However, the progression of Alzheimer's disease begins several years before any symptoms appear and progress [1]. There are numerous medications in development because there is currently no cure for Alzheimer's disease [4]. As a result, developing medicines will have the most significant impact when used at the early stages of the disease. As a result, identifying high-risk patients who develop AD is crucial [1, 5]. As a result, early detection is critical for developing a treatment approach that would reduce the progression. It is when the sickness progresses from one symptom to the next. At the same time, current research is mainly focused on forecasting whether it will transition to a different stage.

The rise in prevalence of neurodegenerative disorders like Alzheimer's disease has piqued the interest of researchers worldwide, who are working to develop high-performing methodologies for diagnosis, treatment, prevention, and target drug discovery. The change rate of these variables could be used as an additional source of information in the risk assessment of conversion from MCI [2]. Using modern diagnostic methods and biomarker tests, researchers could diagnose Alzheimer's disease. By combining biomarkers, it achieves varying levels of accuracy [6]. Unfortunately, the present research focuses on using MRI to classify illness states at their current stage rather than combining various features. As a result, these studies function as proof of concept without being tested in the real world.

Predicting AD conversion is a significant research subject in today's artificial intelligence and machine learning technology age. The institutional use of machine learning techniques and the shift toward a personalized medicine concept, particularly in medical sectors, represents a chance to improve clinical results. Based on the subject's facts, it makes individual forecasts with a certain degree of certainty. It could help researchers and physicians make more informed and successful decisions [2, 7]. This research provides a hybrid strategy for high-dimensional data selection. The current project intends to

provide a therapeutically applicable machine learning approach for identifying the best features that can help diagnose Alzheimer's disease patients. The rest of the paper is organized as follows: Section 2 discusses the previous work related to the proposed approach. All details about the dataset, the number of subjects chosen, and conditions applied to choose this instance are described in section 3. The approach discussion and framework showed in section 4. Finally, Section 5 conducts the experiments made with results.

2. Related Work

Existing papers [8] classify machine learning techniques used to classify Alzheimer's disease. (1) human-engineered feature selection and traditional classification; (2) deep network feature selection and deep network classification; (3) deep network feature selection and traditional classification; and (4) human-engineered feature selection and deep network classification are all examples of human-engineered feature selection and deep network classification.

2.1. Human Engineered Feature Selection and Traditional Classification

Zeng et al. [9] presented an approach for the diagnosis of Alzheimer's disease that outperforms various SVM models as well as two additional state-of-the-art approaches with deep learning incorporated, making it a useful AD diagnosis method. The support vector machine (SVM) parameters are optimised using a new switching delayed particle swarm optimization (SDPSO) algorithm. Though classification techniques are commonly used to identify medical disorders, the lack or inaccuracy of labelled data might be an issue. Farouk and Rady[10] investigate the application of unsupervised clustering techniques in the early detection of Alzheimer's disease. When compared to a global whole brain study, the effect of selecting specific local regions of interest (ROIs). The findings reveal that the proposed method may accurately diagnose Alzheimer's disease at an early stage with a 76 percent accuracy rate.

2.2. Deep Network Feature Selection and Deep Network Classification

A performed a comprehensive search for articles that used deep learning techniques and neuroimaging data to diagnose Alzheimer's disease is described in[11]. Hong et al. [12] proposed a prediction model based on long short-term memory (LSTM), which could connect earlier information to the current task. As a result, an LSTM network is constructed to encode the temporal relationship between characteristics and the progression of Alzheimer's disease. Experiments demonstrate that the model outperforms the vast majority of other models.

Ramzan et al. [13] investigate the utility of resting-state functional magnetic resonance imaging for multi-class classification of Alzheimer's disease and its stages. It gives a deeper understanding of deep learning algorithms and how they might be used to classify Alzheimer's disease. It used residual neural networks to perform AD classification with an average accuracy of 97.92 percent and 97.88 percent. Building and validating Convolutional neural networks (CNNs) to predicting the individual diagnosis of Alzheimer's disease (AD) and mild cognitive impairment who will convert to AD (c-MCI) based on a single cross-sectional brain structural MRI scan discussed in [14]. It worked well without any prior feature engineering and across a wide range of imaging methods and scanners. Untrained operators can use it that it is generalized to previously unreported patient data.

2.3. Deep Network Feature Selection and Traditional Classification

Ebrahimighahnavieh et al.[15] illustrates a review of biomarkers and features for dealing with neuroimaging data from single-modality and multi-modality experiments. A deep learning-based feature representation with a stacked auto-encoder is proposed by [16]. The experiments show that the proposed method is 95.9%, 85.0%, and 75.8% accurate for AD, MCI, and MCI-converter diagnosis, respectively.

Lin et al.[17] proposed a deep learning approach based on convolutional neural networks (CNN). It uses magnetic resonance imaging (MRI) data to reliably forecast MCI-to-AD conversion. In leave-one-out cross-validations, it achieves an accuracy of 79.9% and an area under the receiver operating characteristic curve (AUC) of 86,1 %. CNN's prediction performance can be improved by using age correction and aided structural brain imaging characteristics.

2.4. Human Engineered Feature Selection and Deep Network Classification

Wen et al.[18] illustrate a review on over 30 papers that have to use convolutional neural networks (CNN) for AD classification from MRI. However, the classification performance is difficult to compare across studies due to variations in components. One of the issues discussed is that some of these papers may report a biased; due to inadequate or unclear validation or model selection procedures. Using four classes of a classification problem. Amoroso et al. [19] offer a classification technique based on Random Forest feature selection and deep Neural Network classification. Traditional feature selection, in addition to the deep network, can detect disease biomarker which affect its progression. Conventional classification approaches fail to capture the nonlinear relationship between features and disease. As a result, to develop the Alzheimer's Disease prediction model, we use method four; the classic feature selection approach; in conjunction with the deep network.

3. Material and Approach

3.1. ADNI Data

The data for this article was compiled using the Alzheimer's Disease Neuroimaging Initiative (ADNI) database (adni.loni.usc.edu). In 2003, the ADNI launched as a public-private collaboration. The goal is to find early detection and tracking indications for Alzheimer's disease (AD). It comprises patients from more than 50 different US and Canadian facilities and six-month follow-up exams. The data used in the suggested work is ADNIMERGE. Although it has nine classes, we only cover three in this paper (AD, MCI, and NL). It has 12612 instances and 90 characteristics.

3.2. Subjects

Eighteen months of 900 patients are extracted on each class (MCI, NL, AD) for each are available diagnostic follow-up assessments with at least eight years which described the time evolution of all variables in 3-month intervals. It covers 2700 records and 24 different neurological tests with corresponding MRIs. Each patient profile consisted of 24 data and seven image files (Table 1) classified as ordinal, continuous, or image. Detailed data processing steps are described in the following

subsection. Some features need to specify for all the participants. The following are related criteria of ADNI studies chosen in this research:

- Age between 55 and 90
- Education: 5 levels considered from elementary to graduate
- All ethnicity and races include
- Diagnosis at baseline are CN, AD, LMCI, EMCI

As mentioned before, the ADNI dataset contains different types of data. This work aims to help diagnose using other data by specifying the duration for test retake. For these reasons, the following are the types of data used.

- Sociodemographic characteristics: is the basics information about the patient
- Neurological Test: For the specialized to test in the brain and mental health conditions (neuropsychologist). The evaluation can include extensive tests to evaluate your memory and cognitive skills.
- Baseline: first tests and diagnosis made for the patients
- Brain image technology: ADNI dataset provides only two types of technologies (MRI and PET)

3.3. Data Preprocessing

Preparing (cleaning and arranging) raw data to make it appropriate for creating and training algorithms is referred to as data preparation. In simple terms, it is a data mining technique that converts unintelligible data into something that can be read. Dataset is confronted with two major challenges:
1) Missing Values 2) Discretization of data
Some of the ADNI dataset attributes contain missing values ranging from 50% to 90% of the original data. First, remove the one with more than 60% of its data absent. As a result, the dataset becomes 57 attributes instead of 87 original data. However, what about the features that contain 50% to 60% of their data missing. Two approaches depend on attribute type (numeric or nominal) selected [20]: 1) Impute missing values with mean 2) Multiple imputations. At the same time, they were replacing known features with their corresponding value [21, 22]. Data discretization is a technique for transforming many data values into smaller ones, making data interpretation and management more effortless. In other words, data is a technique for turning continuous data's attribute values into a finite collection of intervals with slight data loss. To change the data, refer to the ADNI data description and discuss with experts. More details are available in table 2.

Table 1: The model contains cognitive function variables, as well as MRI images.

Name	Type	Mean	Missing %
CDRSB	ordinal	2.425	1
MMSE	Ordinal	25.862	0
ADAS11	Ordinal	13.009	0
ADAS13	Ordinal	19.939	2
RAVLT immediate	continuous	31.223	1
RAVLT learning	Continuous	3.547	1
RAVLT forgetting	. continuous	4.067	1
RAVLT perc forgetting	Continuous	63.86	3
FAQ	Continuous	6.769	0
MOCA	. continuous	17.738	91
EcogPt (6 test)	Continuous	2.2	90
EcogPt Total	Continuous	2.689	90
Ventricles	image	----	2
Hippocampus	image	---	2
WholeBrain	image	---	1
Entorhinal	image	----	2
Fusiform	image	---	2
MidTemp	image	---	2
ICV	image	---	1

Table 2 Known Discretize Value

Feature Name	Discretization item
Age	≤60, ≤70, ≤80, ≤90, ≤100
Education	≤ 5: Elementary , ≤ 8:Middle, ≤ 11:High, ≤ 16: University, > 16: Graduate
CDRSB	0-6 A, 6-12 B, 12-18 C
ADAS11	0-17.5 Least , 17.5 - 35 Moderate , 35 - 52.5 Mild , 52.5 - 70 Most
ADAS13	0 - 20 Least , 20 -40 Moderate , 40 -60 Mild , 60 -80 Most
MMSE	0-10 Severe ,10-20 Moderate, 20-25 Mild, 25-30 Normal

3.4. Approach

A dataset's goal concepts in machine learning (ML) describe a group of features, with the expectation that the parts will contain as much usable information as possible and that the number of components will be as short as possible for solving tasks like classification. However, because there is typically little prior information on the dataset, determine which traits are helpful and which are not [23]. As a result, many features are frequently considered, including many that are useless or redundant. Unfortunately, the performance of machine learning will lower training efficiency [24].

Various feature selection methods propose removing the detrimental impact of irrelevant and redundant features. The main goal is to select significant sections from many attributes. Traditional search methodologies, such as Partial Swarm Optimization (PSO)[25], presented numerous meta-heuristics inefficiencies in tackling complex combination optimization issues. The PSO algorithm is a swarm-based stochastic optimization technique [25, 9].

The key advantages of the PSO algorithm over mathematical algorithms and other heuristic optimization techniques are its simple ideas, ease of implementation, robustness to control factors, and computational efficiency [25]. Also, A Simulator-driven nonconvex optimization framework included a few assumptions based on objective functions. As a result, machine learning optimization improves applied learning [26]. As a result, this paper introduces a PSO hybrid technique.

Random forest (RF) is a classification and regression ensemble method. It has significant advantages over other approaches in managing extremely non-linearly correlated data, noise robustness, tuning

simplicity, and parallel processing efficiency [27]. For disease prediction, it uses the relevant features that are supplied as inputs to the model. It also has another important feature: an intrinsic feature selection step performed before the classification task to decrease the variables' space by assigning an importance value to each feature [28]. RF was the best option for a hybrid strategy for all of these reasons.

4. Proposed Approach

This paper provides a hybrid strategy for selecting high-dimensional features that have been updated to be suitable for combinatorial optimization. The key benefit of this method is that it can cope with high-dimensional data and select the best characteristic to diagnose patients. At the same time, it accepts various data types as input and combines them into a single data type. Then it is wrapped up in a feature selection wrapper. Figure 1 shows the architecture of RF-PSO. It comprises four different categories of input data, each linked to a specific type of diagnostic data, as detailed in sections 3.1 and 3.2. (text, numbers, and images). Each piece of information is linked to its corresponding block.

It works simultaneously on four blocks and the result is related to a specific block. At the same time, it searches over only associated data. As a result, each block gives an output data that depends on the features included. The output is divide into four different blocks with the same data type (numbers). Finally, the classification approach (Neural network is used in this work – more in detail in the next section) takes results from the blocks simultaneously then make a diagnosis.

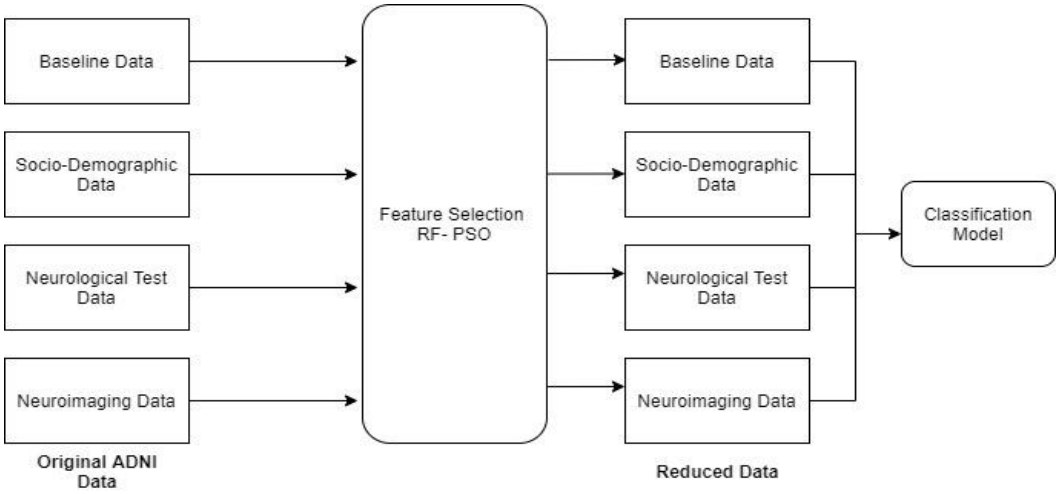


Figure. 1: Proposed Approach Framework.

The core idea is that a swarm of particles moves through a problem space, evaluating their positions using a fitness function. After defining a problem space, a collection of particles is produced, and their positions and velocities are iteratively updated. It separates each sample into a relevant dimension after organizing the data into several samples. With random placements and velocities, each dimension represents variables relevant to diagnosis criteria. The data is divided into multiples at this point, and each sample begins to create a tree. The algorithm starts with the development of a random forest. It

starts by randomly choosing k characteristics from each dimension. Then figure out which node reflects the best split. If the calculation number is large, it has a tangential relationship with class.

4.1. Feature Selection

The most practical tool for dealing with high-dimensional data difficulties is dimensionality reduction. Extraction and selection are the two most essential components. 1) Feature extraction converts the original high-dimensional properties to a low-dimensional feature space. 2) For approach construction, feature selection directly selects a subset of essential attributes [29]. They improve learning performance, computational efficiency, memory storage efficiency, and generalization models. Actual data has many features that are irrelevant, redundant, or noisy. By removing these by selecting attributes, space and money are saved while avoiding severe information loss and reducing learning performance.

After data preprocessing, the dataset includes 45 attributes, a significant number to analyze, and 2700 instances. As a result, a feature extraction and selection approach suggested choosing the best features. In contrast, these features need to be acceptable by medication standards. This article applies three experiments on RF-PSO to compare accuracy with original algorithms (RF, PSO). As a result, using PSO was done at the beginning.

Table 3 shows the number of attributes after applying three algorithms. PSO has only 13, while RF has only 27. It noticed that RF had the highest number of features, which takes time to establish and analyze. Simultaneously, RF-PSO includes 16 attributes. Discuss the difference between each approach and the accuracy explained in section 5.

Table 3 Feature Selection Comparison.

Original Dataset	PSO	RF	RF-PSO
Baseline (19 attribute)	5	14	3
Socio (9 attribute)	1	2	1
MRI (7 attribute)	1	3	5
Neurological (9 attribute)	5	7	5

5. Results and Discussion

Ten-fold cross-validation is used to train our model for high-dimension data selection. RF-PSO feature selection is an embedded approach that uses a random collection of features to create multiple decision trees over a random set of observations from the dataset. This technique can achieve good predictive performance with modest overfitting and noise resilience. However, one tree may be sensitive to noise; the correlation and noise sensitivity decrease when the result is averaged over numerous trees.

A neural network (NN) is a machine learning inspired by the human brain. While a neuron is the fundamental unit of the brain, a perceptron is the primary building component of NN. Given an ANN and an error function, it estimates the gradient of a loss function for all network weights accurately. Backpropagation is used to train the NN in this research. Some of the possible grounds for endorsing this strategy are as follows:

- It is quick, simple, and straightforward to program.
- It is adaptable and does not necessitate prior knowledge.
- It does not necessitate any specific description of the functions to be mastered.

The backpropagation network uses only one hidden layer with ten units. The loss function represents a squared error, and the activation function is Sigmoid. The following section contains a discussion of why ten units were chosen. A classification strategy (or "classifier") is described using a confusion matrix. A set of rate lists is computed from a confusion matrix. Let us start by defining the most basic terms:

- True positives (TP): They are predicted to have the disease and suffer from it.
- True negatives (TN): They are predicted not to have the disease and not suffer from it.
- False positives (FP): They are predicted to have the disease and not suffer from it.
- False negatives (FN): They are predicted not to have the disease and suffer from it.

Table 4 Relation between number of units and accuracy with Build time

Hidden Layer	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
Accuracy	91.8	94.3	93.9	93.5	94.3	93.7	94	94.7	94.6	95	94.5	94.1	93	92.5	91.9
Time (Second)	1.69	4.22	9.3	22.3	25.4	28.7	34.5	35	35.4	47.8	46	50	50.8	51.3	52

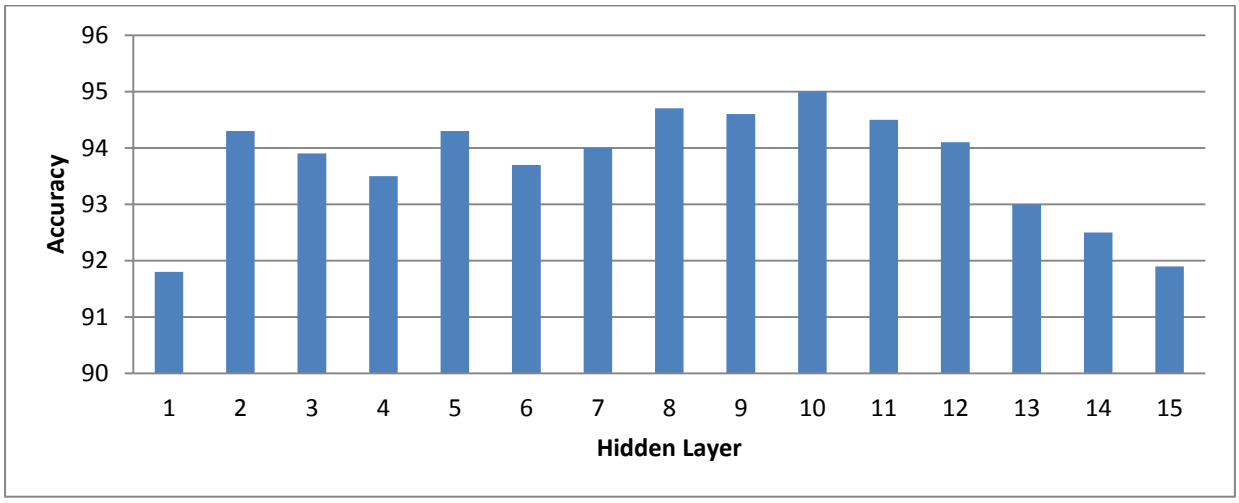


Figure. 2: Relation between accuracy and number of hidden layers.

Table 4 discusses the differences between layers and time with accuracy. Ten hidden layers got the most precise accuracy but not the lowest time. As a result, it needs to prefer which one is better 1)Low time 2) best accuracy. If the time is the only factor considered, choose one hidden layer where it had the worst accuracy. While, if the best accuracy is chosen, then the one with ten hidden layers is chosen. This work helps diagnose the patient. The time is essential but not the same as accuracy. Figure 2 illustrates The relation between the number of layers and RF-PSO accuracy. It shows that the accuracy is unstable from 1 to 6 hidden units. Periodically, it increased and in others decreased. Simultaneously, from layer 6 to 10, the value is rising; it dropped from layer 11. As a result, ten hidden layers were selected to match our objectives. A comparison is made with state-of-the-art approaches. The proposed approach obtains the best accuracy and extracts an average timing value.

Figure 3 illustrates the comparisons between our approach and others. Our approach got the best accuracy by 95%. At the same time, using RF only got 86%, the worst value. In contrast, the PSO approach got 89%. It shows that the hybrid approach decreased human interaction and increased accuracy percentage.

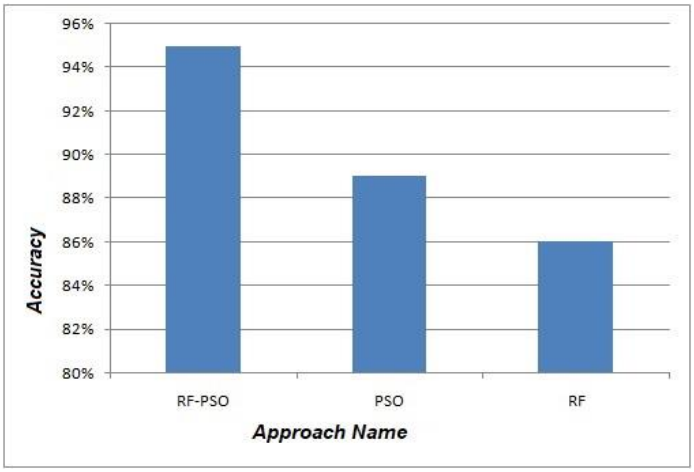


Figure. 3: Feature Selection Accuracy.

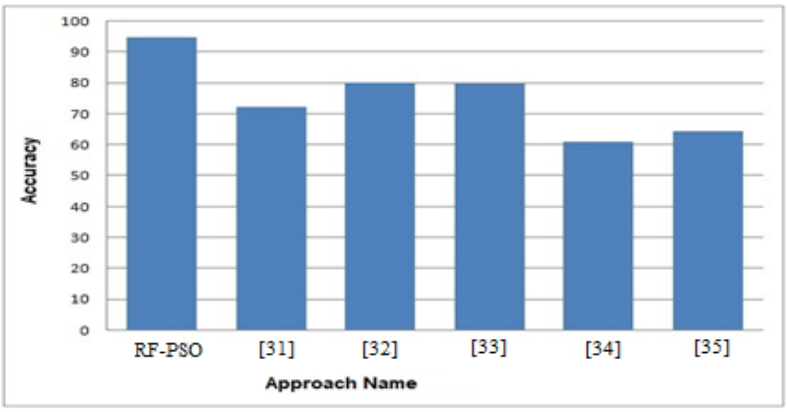


Figure. 4: RF-PSO vs. Other Approaches.

A ROC curve depicts the connection between clinical sensitivity and specificity for each probable patient. It is a graph with the following information: 1 – specificity (false-positive fraction = FP/(FP+TN)) is shown on the x-axis. The sensitivity (= genuine positive fraction= TP/(TP+FN)) is shown on the y-axis. It is the most effective balance strategy for evaluating the model [30]. The most significant number is 0,99, which belongs to the regular class. Dementia and MCI both have a score of 0,99 and 0,983, respectively.

In order to, compare RF-PSO a comparison is made with the following papers [31, 32, 33, 34, 35]. These specific works are chosen for their recent dates and objective similarity. The following features are standardized for comparison: 1) Number of Instances (2700) 2) Classification technique (Backprobagation NN) 3) Three Classes each one 900 records.

Figure 4 illustrates the difference between the proposed model and others accuracy values. The work [32], got 80,23% while [33], has 79,70 %. It followed by 72,3% for [31], and [35] 64,4% accuracy. Finally, the lowest number is 60,92% in work published by [34]. By dividing the data into different sets, the accuracy will increase. In real cases, AD cannot diagnose with MRI images only. Also, neurological test scores affect the final diagnosis. The final results connect between experiments and real scenarios. At the same time, it helps specialists diagnose patients with AD.

6. Conclusions

An efficient hybrid high-dimensional feature selection strategy (RF-PSO) is suggested based on a weighted rank average ensemble of several supervised machine learning algorithms. Its goal is to find the best characteristics in high-dimensional data. It supports a variety of data sets in order to improve diagnosis detection and increase the number of individuals in clinically transabled tests who are at continual risk of developing Alzheimer's disease. Our technique predicts Alzheimer's disease and detects the disease's relative brain ROI. Experiments demonstrate that our technique outperforms other models in terms of accuracy compared to other algorithms. For 3-class analysis, RF-PSO accuracy ranged from.93 to.95 across all modalities. There is a chance to look into how imaging modalities can help with multiclass accuracy, also, how our method can be utilized or extended to account for more than one type of AD prediction.

References

1. A. Association, 2019 alzheimer’s disease facts and figures, Alzheimer’s & Dementia 15 (3) (2019) 321–387. doi:https://doi.org/10.1016/j.jalz.2019.01.010.
2. R. Perneczky (Ed.), Biomarkers for Preclinical Alzheimer’s Disease, Vol. 137 of Neuromethods, Humana Press, New York, NY, 2018. doi:https://doi.org/10.1007/978-1-4939-7674-4.
3. P. Vemuri, T. G. Lesnick, S. A. Przybelski, D. S. Knopman, V. J. Lowe, J. Graff-Radford, R. O. Roberts, M. M. Mielke, M. M. Machulda, R. C. Petersen, et al., Age, vascular health, and Alzheimer disease biomarkers in an elderly sample, Annals of neurology 82 (2017) 706–718. doi:https://doi.org/10.1002/ana.25071.

4. Geerts H and Dacks PA and Devanarayan V and Haas M and Khachaturian ZS and Gordon MF and Maudsley S and Romero K and Stephenson D and Brain Health Modeling Initiative (BHMI), Big data to smart data in alzheimer's disease: The brain health modeling initiative to foster actionable knowledge, Alzheimer's & Dementia 12 (9) (2016) 1014–1021. doi:10.1016/j.jalz.2016.04.008.
5. George A. Edwards III, Nazaret Gamez, Gabriel Escobedo Jr, Olivia Calderon, Ines Moreno-Gonzalez, Modifiable risk factors for alzheimer's disease, Frontiers in aging neuroscience 11 (164). doi:10.3389/fnagi.2019.00146.
6. M. Tanveer, B. Richhariya, R. U. Khan, A. H. Rashid, P. Khanna, M. Prasad, C. T. Lin, Machine learning techniques for the diagnosis of alzheimer's disease: A review, ACM Transactions on Multimedia Computing, Communications, and Applications 16 (15). doi:https://doi.org/10.1145/3344998.
7. V. Bol´on-Canedo, N. S´anchez-Maro˜no, A. Alonso-Betanzos, Feature Selection for High-Dimensional Data, Artificial Intelligence: Foundations, Theory, and Algorithms, Springer, 2015.
8. X. HONG, R. LIN, C. YANG, C. CAI, K. C. . C. of Computer Science, X. . C. Technology, Huaqiao University, Adpm: An alzheimer's disease prediction model for time series neuroimage analysis, IEEE Access 8 (2020) 62601–62609. doi:10.1109/ACCESS.2020.2979969.
9. N. Zeng, H. Qiu, Z. Wang, W. Liu, H. Zhang, Y. Lid, A new switching-delayed-pso-based optimized svm algorithm for diagnosis of alzheimer's disease, Neurocomputing 320 (2018) 195–202. doi:https://doi.org/10.1016/j.neucom.2018.09.001.
10. Y. Farouk, S. Rady, Early diagnosis of alzheimer's disease using unsupervised clustering, International Journal of Intelligent Computing and Information Sciences 20 (2) (2020) 112–124. doi:10.21608/ ijicis.2021.51180.1044.
11. T. Jo, K. Nho, A. J. Saykin, Deep learning in alzheimer's disease: Diagnostic classification and prognostic prediction using neuroimaging data, Frontiers in Aging Neuroscience 11. doi:10.3389/fnagi.2019.00220.
12. X. Hong, R. Lin, C. Yang, N. Zeng, C. Cai, J. Gou, J. Yang, Predicting alzheimer's disease using lstm, IEEE Access 7 (2019) 80893–80901. doi:10.1109/ACCESS.2019.2919385.
13. F. Ramzan, M. U. G. Khan, A. Rehmat, S. Iqbal, T. Saba, A. Rehman, Z. Mehmood, A deep learning approach for automated diagnosis and multi-class classification of alzheimer's disease stages using resting-state fmri and residual neural networks, Journal of Medical Systems 44 (37). doi:https://doi.org/10.1007/s10916-019-1475-2.
14. S. Basaia, F. Agosta, L. Wagner, E. Canu, G. Magnani, R. Santangelo, M. F. for the Alzheimer's Disease Neuroimaging Initiative, Automated classification of alzheimer's disease and mild cognitive impairment using a single mri and deep neural networks, NeuroImage: Clinicaldoi:10.1016/j.nicl.2018.101645.
15. M. A. Ebrahimighahnavieh, S. Luo, R. Chiong, Deep learning to detect alzheimer's disease from neuroimaging: A systematic literature review, Computer Methods and Programs in Biomedicine 187 (2020) 105242. doi:https://doi.org/10.1016/j.cmpb.2019.105242.
16. Heung-Il Suk, Dinggang Shen, Deep learning-based feature representation for ad/mci classification, in: K. Mori, I. Sakuma, Y. Sato, C. Barillot, N. Navab (Eds.), Medical Image Computing and Computer-Assisted Intervention – MICCAI 2013, Vol. 16, Springer Berlin Heidelberg, Berlin Heidelberg, 2013, p. 583–590.
17. W. Lin, T. Tong, Q. Gao, D. Guo, X. Du, Y. Yang, G. Guo, M. Xiao, M. Du, X. Qu, T. ADNI., Convolutional neural networks-based mri image analysis for the alzheimer's disease prediction from

mild cognitive impairment, Frontiers in Neuroscience 12 (2018) 777. doi:10.3389/fnins.2018.00777. URL https://www.frontiersin.org/article/10.3389/fnins.2018.00777
18. J. Wen, E. Thibeau-Sutre, M. Diaz-Melo, J. Samper-Gonz´alez, A. Routier, S. Bottani, D. Dormont, S. Durrleman, N. Burgos, O. Colliot, Convolutional neural networks for classification of alzheimer's disease: Overview and reproducible evaluation, Medical Image Analysis 63 (2020) 101694.
19. N. Amoroso, D. Diacono, A. Fanizzi, M. La Rocca, A. Monaco, A. Lombardi, C. Guaragnella, R. Bellotti, S. Tangaro, Deep learning reveals alzheimer's disease onset in mci subjects: Results from an international challenge, Journal of Neuroscience Methods 302 (2018) 3–9, a machine learning neuroimaging challenge for automated diagnosis of Alzheimer's disease.
20. R. R. Janghel, Deep-Learning-Based Classification and Diagnosis of Alzheimer's Disease, IGI Global, 2020, Ch. 76, p. 25. doi:10.4018/978-1-7998-0414-7.ch076.
21. M. Geaur Rahman, M. Zahidul Islam, Discretization of continuous attributes through low frequency numerical values and attribute interdependency, Expert Systems with Applications 45 (2016) 410 – 423. doi:https://doi.org/10.1016/j.eswa.2015.10.005.
22. H. Kang, The prevention and handling of the missing data, Korean journal of anesthesiology 65 (5) (2013) 402–406. doi:https://doi.org/10.4097/kjae.2013.64.5.402.
23. G. Chandrashekar, F. Sahin, A survey on feature selection methods, Computers and Electrical Engineering 40 (1) (2014) 16 – 28. doi:10.1016/j.compeleceng.2013.11.024.
24. S. Gu, R. Cheng, Y. Jin, Feature selection for high-dimensional classification using a competitive swarm optimizer, Soft Computing 22 (2018) 811– 822. doi:doi.org/10.1007/s00500-016-2385-6.
25. D. Wang, D. Tan, L. Liu, Particle swarm optimization algorithm: an overview, Soft Computing 22 (2018) 387–408. doi:https://doi.org/10.1007/s00500-016-2474-6.
26. A. Darwish, A. E. Hassanien, S. Das, A survey of swarm and evolutionary computing approaches for deep learning, Artificial Intelligence Reviewdoi:https://doi.org/10.1007/s10462-019-09719-2.
27. Dimitriadis SI, Liparas D, Tsolaki MN, Alzheimer's Disease Neuroimaging Initiative, Random forest feature selection, fusion and ensemble strategy: Combining multiple morphological mri measures to discriminate among healhy elderly, mci, cmci and alzheimer's disease patients: From the alzheimer's disease neuroimaging initiative (adni) database, Journal of Neuroscience Methods 15 (302) (2018) 14–23. doi:10.1016/j.jneumeth.2017.12.010.
28. A. Sarica, A. Cerasa, A. Quattrone, Random forest algorithm for the classification of neuroimaging data in alzheimer's disease: A systematic review, Frontiers in Aging Neuroscience 9 (2017) 329. doi:10.3389/fnagi.2017.00329.
29. J. Li, K. Cheng, S. Wang, F. Morstatter, R. P. Trevino, J. Tang, H. Liu, Feature selection: A data perspective, ACM Computing Surveys (CSUR) 50 (6). doi:10.1145/3136625.
30. A. P. Bradley, The use of the area under the roc curve in the evaluation of machine learning algorithms, Pattern Recognition 30 (7) (1997) 1145–1159. doi:https://doi.org/10.1016/S0031-3203(96) 00142-2.
31. S. Qiu, G. H. Chang, M. Panagia, D. M. Gopal, R. Au, V. B. Kolachalama, Fusion of deep learning models of mri scans, mini-mental state examination, and logical memory test enhances diagnosis of mild cognitive impairment, Alzheimers Dement (Amst). 28 (10) (2018) 737–749. doi:10.1016/j.dadm.2018.08.013.
32. Grassi M, Rouleaux N, Caldirola D, Loewenstein D, Schruers K, Perna G, Dumontier M, Alzheimer's Disease Neuroimaging Initiative, A novel ensemble-based machine learning algorithm to predict the conversion from mild cognitive impairment to alzheimer's disease using

sociodemographic characteristics, clinical information, and neuropsychological measures, Front Neurology 10 (756). doi:10.3389/fneur.2019.00756.

33. Haaksma ML, Calderón-Larrañaga, Olde Rikkert MG, Melis RJ, Leoutsakos JS, Cognitive and functional progression in Alzheimer disease: A prediction model of latent classes, International journal of geriatric psychiatry 33 (8). doi:10.1002/gps.4893.
34. A. Shikalgar, S. Sonavane, Hybrid deep learning approach for classifying alzheimer disease based on multimodal data, in: B. Iyer, P. S. Deshpande, S. C. Sharma, U. Shiurkar (Eds.), Computing in Engineering and Technology, Springer Singapore, Singapore, 2020, pp. 511–520. doi:https://doi.org/10.1007/978-981-32-9515-5_4.
35. X. Bi, X. Hu, H. Wu, Y. Wang, Multimodal data analysis of alzheimer's disease based on clustering evolutionary random forest, IEEE Journal of Biomedical and Health Informaticsdoi:10.1109/JBHI.2020.2973324.