science, engineering and health studies

Research Article

Science, Engineering and Health Studies
https://lio1.tci-thaijo.org/index.php/sehs
ISSN (Online): 2630-0087

Automated Alzheimer's disease detection from brain magnetic resonance imaging using a smart classifier fusion approach

Tawseef Ayoub Shaikh[1] and Rashid Ali[2]*

[1] Department of Computer Science and Engineering, Pandit Deendayal Energy University, Gandhinagar, Gujrat 382007, India
[2] Department of Computer Engineering, Aligarh Muslim University, Uttar Pradesh 202001, India

[*]***Corresponding author***:
Tawseef Ayoub Shaikh
tawseef.shaikh@pdpu.ac.in

Received: *13 November 2020*
Revised: *17 May 2021*
Accepted: *25 May 2021*
Published: *21 June 2022*

Citation:
Shaikh, T. A., and Ali, R. (2022). Automated Alzheimer's disease detection from brain magnetic resonance imaging using a smart classifier fusion approach. Science, Engineering and Health Studies, 16, 22040002.

ABSTRACT

This aim of this research was to use a novel method for discriminating between Alzheimer's disease from normal controls from brain magnetic resonance imaging. Here, the six diverse connecting rules (mean, product, maximum, minimum, and voting) related to the consolidation of classifiers. The collection of the relevant benchmark data was taken from the Alzheimer's disease neuroimaging initiative (ADNI) data set for the proposal evaluation. The empirical investigations uncovered the four individual classifiers out of thirteen classifiers, viz BayesNet, linear discriminant classifier, quadratic Bayes normal classifier, and kernel support vector machine from numerous machine learning groups, gained the highest recognition percentages of 74.77%, 71.62%, 77.76, and 76.13%, respectively. These four-best performing classifiers were employed for prototyping the classifier fusion model, which displayed a much healthier performance with a shared mean error rate of 0.2123, in contrast to the mean error rate of 0.2493 before ensemble. Our analyses have shown that a smart classifier-based fusion method outperforms the base-classifier method.

Keywords: classifier fusion strategy; pattern recognition; performance evaluation indices; error rates

1. INTRODUCTION

The increase in the prevalence of Alzheimer's disease (AD) from the previous decade has increased the burden to rank it seventh globally and costs society approximately $818 billion in lost output. It is made up of 90% of the population over the age of 65, who are among the 50 million individuals in the world suffering from dementia (Chien et al., 2019). All dementias vary in terms of symptoms and disease, and there are more than 600 types of dementia that contrast with others, and AD is the most among them (Chien et al., 2019). Estimation in 2001 estimated that the prevalence of dementia would rise to 42.3 million by 2020 and hit 81.1 million in 2040 (Tierney et al., 2005). About 3.7 million people with dementia will cost India around 14.7 billion US dollars. According to the report, such increases will triple the expenditure.

Currently, AD is administered using a group of methods like the Montreal cognitive assessment (MoCA) (Orimaye et al., 2017) and mini-mental state examination (MMSE). Out of a total score of 30, a case here is negated to have any dementia with an MMSE score of 27 or above (Pozueta et al., 2011). The manual approach to these tests makes them more difficult because they have a low rate of accessibility for AD. As with the intensity of the analysis, the tests, the clinician's involvement and ability to analyze the many subsets of the illness varies as well. Studies of Alzheimer's disease and mild cognitive impairment have discovered, such modalities as structural magnetic resonance imaging (sMRI), functional MRI (fMRI), flourodeoxy glucose positron emission tomography (FDG-PET) and amyloid PET, for example, Pittsburgh compound B (PiB-PET),

florbetapir, and flutemetamol (Thal et al., 2015). s-MRI has the capability to detect more subtle morphological alterations in the brain's microstructure than any other imaging techniques currently available (Tessitore et al., 2016). Consequently, this would result in negative variations in imaging attributes, and help to contain sickness transfer. Also known as diffuse neuronal atrophy, the phenomenon of progressive cerebral atrophy can be noted by modern MRI technology. In addition, measures of cortical volume and cortical thickness have been used to gain a better understanding of the overall picture of the underlying pathology of AD.

sMRI is predominately employed for stokes, blood clots, brain tumors or additional aberrations that might account for multiple sclerosis (MS) or Alzheimer's because of its high-resolution imaging and high brain tissue contrast capabilities. Three MR tissue parameters specify the MRI scan, e.g., spin-lattice (T1), spin-spin (T2) relaxation times, and proton density (PD), which are labelled as T1 weighted, T2 weighted, and PD weighted. Under a magnetic field, the various subparts of brain-like gray matter (GM), white matter (WM), glial matter, cerebrospinal fluid (CSF), fat, muscle/skin, display exclusive characteristics. The T1, T2 and PD occupy the distribution of the noteworthy tissues of brain-like CSF, GM and WM. Subsequently, spatial, along with tissue characteristic-based features, can be mined from these MRI scans. The manual cataloguing for these bulky tissue volumes is not only tedious but non-reproducible. In this manner, the advancement of entirely programmed and exact mind tissue characterization from MRI in the various disease manifestations like tumors, MS, AD and other WM lesions is of incredible intrigue. AD tends to be recognized at the beginning phase with the assistance of MRI to stay away from irreversible harm of the brain with a legitimate treatment plan. MRI can represent 'atrophy' – a reduction in the size of various different regions of the brain triggered due to the degenerative process of brain tissues in rejoinder to a disease process like AD (Sadek, 2013).

Once thought to be of divine origin, neuroimaging disorders have been deciphered by MRI (Papakostas et al., 2015). Also, electroencephalography has been used as a marker for the earlier and early detection of Alzheimer's disease in many studies (Sankari and Adeli, 2011). Time and effort have been dedicated for earlier diagnosis of Alzheimer's disease that has (Varatharajan et al., 2018). Numerous studies have focused on the overall power of ensemble for early detection (Kumar et al., 2018). The features' quality has a direct impact on the classifier's predictive capability. Preclinical researchers have made a significant contribution to extracting the most relevant feature vectors that differentiate between AD subjects and healthy individuals (Lahmiri and Shmuel, 2018). Deep learning (DL) is nowadays the most trending area and has been widely employed in automated neuroimaging analysis, particularly the AD diagnosis (He et al., 2019).

While these methodologies yield excellent results, current application for characterizing clinical information has not yet uncovered consistent data. This particular methodology has problems because it uses only about a few classifiers. Furthermore, they often leave unacknowledged factors, such as age and sex, both of which have a significant impact on results. To sum up, good performance depends on one particularity, while a good presentation may need a mix of specifics. It utilized a high-scale dataset and a complex methodology in this paper. This research was laying the groundwork for numerous new measurements to be taken, so it was conceivable that additional ones would follow. Neuroimaging data determined how precisely AD software needed to be made. These rules (mean, median, maximum, minimum, and voting, and aggregation) were discussed in the report.

2. MATERIALS AND METHODS

This work was primarily focused on the Alzheimer's disease neuroimaging initiative (ADNI) dataset and feature extraction approaches. Statistical analysis in the form of mathematical formulae did also occupy the space here, which were used for comparison estimations of classifiers, like classification accuracy, sensitivity, specificity, area under curve (AUC), and error rate.

2.1 Data acquisition

The proposal was validated using the ADNI database, which involved high-resolution T1-weighted s-MRI of 668 candidates (www.adni.loni.usc.edu). The main goal of its launching in 2004 was to gauge the advancement of mild cognitive impairment (MCI) and AD, employing various neuroimaging modalities, like MRI, PET, other biological markers clinical and neuropsychological examinations. For this work, a total of 668 subjects in the ADNI2 baseline dataset was employed, which consisted of 146 AD cases, 336 MCI cases, and 186 CN (Table 1).

Table 1. Demographic data of 668 subjects in the ADNI2 baseline dataset

	Male	Female	Age	Min/ Max Age	APOE1	APOE2	FAQ
AD	85	61	74.73±8.15	56/ 90	3.11	3.63	13.39
MCI	186	150	71.30±7.50	55/ 91	2.94	3.42	02.08
CN	89	97	73.50±6.25	57/ 89	2.86	3.24	00.16

Note: APOE is apolipoprotein E and is the chief genetic risk feature for Alzheimer's disease, FAQ stands for functional activities questionnaire

2.2 Neuroimage processing and feature extraction

Initial pre-processing of the raw MRI images was carried out with the FreeSurfer image analysis suite, which was open-source brain suite software (http://surfer.nmr.mgh.harvard.edu). It performed approximately 31 pre-processing steps like averaging of multiple volumetric T1 weighted images, removal of non-brain tissue using a hybrid watershed/surface deformation procedure, motion correction, etc. Cortical thickness (CT), which were the average values for each region of grey matter probability (GMP), were dug out in the s-MRI case. It was calculated as the closest distance from the grey/white boundary to the grey/CSF boundary at each vertex on the tessellated surface. Individual brain atlases using statistical parametric mapping (IBASPM) were used for volumetric analysis of brain MRIs utilized, an extension of SPM-5. The calculation

of the brain structure was performed as follows: 1) MRIs were segmented into grey matter, white matter and CSF, using IBASPM segmentation, 2) MRI scans were spatially transformed into Montreal Neurological Institute (MNI) space, using affine transformation for approximate registration and non-linear transformation for fine registration to obtain the transformation parameters (Shaikh and Ali, 2019). Then two different feature reduction phases were used in the proposed scheme by inheriting the best of both filter and wrapper feature extracting approaches. Given that some features were uninformative, irrelevant or redundant for classification, reducing the number of features not only speed up computation but also improve classification performance. Therefore, an initial feature selection step was adopted. The feature ranking approach has been widely used in feature selection. In this study, the F-score method was employed for feature ranking followed by feature reducing techniques for extraction of optimal feature vector space for model training.

2.3 Quality assessment

The given examination employed the accompanying quality evaluation strategies to show the adequacy of the proposed CAD framework.

- Accuracy: accuracy is the ratio of right expectations partitioned by the absolute number of forecasts. It is characterized as the capacity of the classifier to choose all cases that require to be chosen and reject all cases that require to be dismissed (Nisbet et al., 2009; Shaikh et al., 2020).

$$Accuracy = \frac{TP+TN}{TP+FP+TN+FN} \quad (1)$$

where TP is true positive, TN is true negative, FP is false positive, and FN is false negative.

- Sensitivity: sensitivity is actually the accuracy of positive cases, i.e., how good a test is in detecting positive disease. It is the capability of a model to pick all the cases that are essential to be picked.

$$Sensitivity = Recall = Hit\ rate = True\ positive\ rate\ (TPR) = \frac{TP}{TP+FN} \quad (2)$$

- Specificity: specificity is actually the accuracy of negative cases, i.e., how likely patients without the disease can be ruled out correctly. It is the capability of a model to throw away all the cases that need to be rejected.

$$Specificity = \frac{TN}{TN+FP} \quad (3)$$

- Area under receiver operating characteristics (AUROC): In a ROC curve, the TP rate is plotted against the FP rate. The values of the FP rate (1 – specificity or TN rate) is plotted on the horizontal axis and the TP rate values (sensitivity or recall) on the vertical axis.

$$Error = 1 = Accuracy = \frac{FP+FN}{TP+FP+TN+FN} \quad (4)$$

2.4 Proposed methodology

The present segment of the projected classification organization designated the classification integration. Their foremost concern was for automatic computer-aided detection tools and techniques that enabled clinical specialists to discover signs of AD before it occurs. The new methodology would make it possible for clinical experts to view data continuously collected for the patients at risk or who have already been diagnosed with AD.

2.4.1 Ensemble

From a conceptual perspective, rather than a technical standpoint, the computerized diagnosis must focus on the fruitful arrangement of characteristics with the lowest possible error rates. A classifier that is fit for use in general cataloguing may not be useful for other types of feature representation. The machine learning models have a wide range of functionality, but their behaviours are quite different depending on what type of job they're applied to (Dietterich, 2000). Settling on an incorrect classification takes less time if there are fewer classifiers, proficiently making good decisions allows the cataloguing of data to be faster. Similarly, a classifier fusion probes from various starting states, helping to avoid exploratory local classifiers that may run into local minima. While the best classifier cannot be present at all times, a selection of classifiers can be constructed for ideal performance (Dietterich, 2000).

The model learning capability and generalization amplified by joining the consensus of multiple classifiers. When merging classifiers in parallel, the posterior class estimates become more accurate (Dai et al., 2015). Tree-like classifier fusion analysis offers more accurate results using weighted averaging for real-time datasets (Dai et al., 2015). This process has proved useful in different application, for, for example, documentation images. Juggling these results together reveals notable improvements in the classifiers' performance (Zheng et al., 2017).

The goal of ensemble learning was to produce diverse base learners (weak classifiers) and integrate their performance on all datasets in order to create a more accurate yield on the whole. While there are many ensemble techniques in the literature, they all need to grapple with the issues of training the base learners and merging the results of the numerous base learners. Ensembles with values of 50% or more excellent must have precision (Li et al., 2012). When classifying results were computed, voting or averaging were employed, respectively.

A module known as "combiner" is expected to amass the classifiers together in the second stage after the choice of specific base-level classifiers gets matured in the initial stage. The combination rule decides the compatibility in the working of the base-level classifiers. The presented merging strategy is implemented by a means of a parallel architecture because of its straightforwardness, less computational time, and furthermore higher certainty level (Eom et al., 2019). The literature represents two standard rules for consolidating: 1) fixed combining rules, 2) trained combining rules (Lebedev et al., 2014). Table 2 shows a list of fundamental fixed joining rules. Trained combining rules, on the other hand, train a discretionary classifier utilising all the training data in the moderate space. Here, the classifiers were typically trained as a yield classifier, using a similar training data set. The posterior probabilities were straightforwardly utilized for the structure of the moderate space (Lebedev et al., 2014)

Table 2. List of the fixed combining rules

Number	Combining rules	Description
1	Median and mean	Both averages the posterior probability estimates, thereby reducing the estimation error
2	Product	Work by taking the product of posterior probabilities of each classifier
3	Maximum	Selects the outcome of the classifier producing the highest estimated confidence, which seems to noise-sensitive
4	Minimum	Selects the outcome of the classifier that has the least objection against a certain class
5	Majority voting	Counts the vote for each class over the input classifiers and selects the majority class

2.4.2 A classifier merger approach to progress the early uncovering of AD

Principle steps that create an ensemble learning procedure for ADNI benchmark data cataloguing for AD diagnosis are as follows (Talia et al., 2016):

1. Initially, the benchmark ADNI dataset is recognized as a biomarker to perceive the incidence of AD, which is then fragmented into respective training and test sets.
2. The n classification algorithms that synchronously run on diverse computation units to shape n autonomous models are fed with the training set.
3. A voter tool v entrees the n models so as to achieve an ensemble cataloguing by conveying every test set instance the label projected by the majority of the n models.

Figure 1 displays the proposed classifier merging approach. Our proposed strategy gave a potential plan that clarified how the plan objectives have been consolidated inside the plan and feature the curiosity of our methodology.

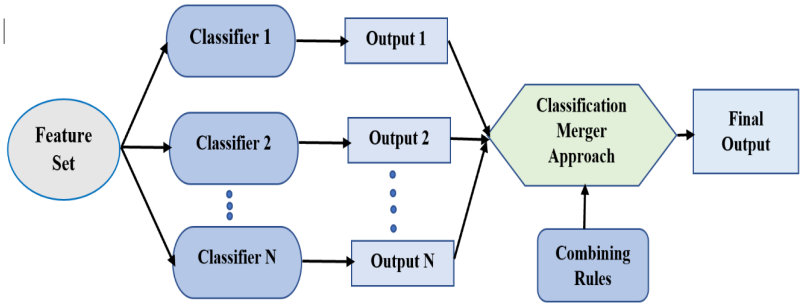

Figure 1. A classifier merger approach

2.4.3 Feature classification

The compulsory feature sets for correct AD identification were offered by the benchmark ADNI dataset designated in the above section. This dataset was utilized to choose a classifier, train it, test it, assess the outcome to decide whether the right arrangement was achieved. The calculation was legitimately corresponding to the quantity and quality of the features considered in the dataset. Various dissimilar classifiers have been explored using a specific feature vector so as to assess the best among the lot for the final classifier merging process. Among the linear classifiers, BayesNet (probabilistic algorithm), stochastic gradient descent (SGD), Fisher's (fishrerc), J48 (decision tree algorithm), RepTree (decision tree algorithm), and nearest means (nmc) were considered, while as a quadratic discriminant classifier (qdc), linear discriminant classifier (ldc), and the quadratic Bayes normal classifier (udc) were employed for density-based classification, and finally, kernel support vector machine (KSVM), simple logistic (SL), multilayer perceptron (MLP), and iBK (instance-based learning) were chosen from the set of non-linear classifiers.

A linear classifier forecasts the class tags grounded on a weighted linear amalgamation of features or the pre-defined variables. Figure 2 shows the classification accuracy of classifiers on the ADNI dataset. Here, the confusion matrix procedure was employed to govern the spreading of errors across all classes. Among the linear classifiers, the BayesNet acquired the best results followed by SGD, Fisherc, nmc, RepTree. Similarly, in the density-based classification group, udc achieved the best accuracy value of 76.13%. Finally, among non-linear classifiers, KSVM topped the list with an accuracy value of 76.13%.

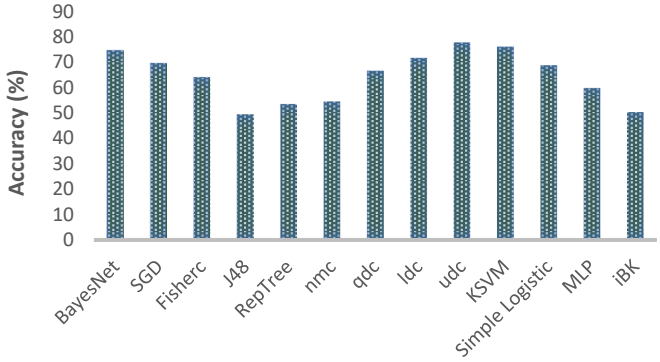

Figure 2. Classification accuracy of classifiers tested

2.4.4 Fusion classifier selection

The four best accomplishing classifiers from the above mentioned were encompassed in the final classifier merging plan. The final classifier fusion was grounded on the amalgamation of these four top-performing classifiers. The complete scenario with simulated results gained throughout the assessment of the ADNI dataset using the thirteen base-level classifiers is depicted in Figure 3. A single array held the results attained from the four topmost performing classifiers, and their error rates were calculated.

2.4.5 Implementation

The individual base-level classifiers and the classifier merging policy was realized via Weka (Version 3.7) and MATLAB R2019b. Initially, the computation commenced by distributing the ADNI dataset into three classes, viz AD, MCI, and healthy control (HC) (Figure 4). A specific label was allocated to each and every class (1 for AD, 2 for MCI, and 3 for HC), and the dataset was randomly fragmented into two equal parts; 70% for training the classifier and the rest 30% for testing. An array (w) = [w1, w2, w3, w4] was created which comprises combination of the four untrained classifiers (w1 = BayesNet, w2 = ldc, w3=udc, and w4 = KSVM) without rules. A cell array (v) encompassing the trained classifiers was formed by the concurrent training of a set of untrained classifiers consuming the training dataset.

The amalgamation of the optimal performing base-level classifiers into a cell array consuming a set of fixed combining rules and their performance outcomes on the unseen test datasets is depicted in Figure 5. The specified work accomplished the examination of the ensemble on six different rules were; mean selection (meanc), median selection (median), product combiner (prodc), maximum selection (maxc), minimum selection (minc), and voting selection (votec). The assessment of a cell array of trained classifiers (v) and the untrained classifiers united with rules (vc) was completed by testing the (testset) set. The performance outcomes displayed by the top four classifiers are deposited in a single array, and their error rates were calculated, which came to be 0.2523 in BayesNet, 0.2838 in ldc, 0.2224 in udc, and finally 0.2387 in the KSVM case. The mean error rate for these four classifiers was 0.2493. The same four classifiers are then merged into a cell array exhausting six different rules, and their error rates were calculated which came to be 0.2537 in mean, 0.2612 in the median, 0.2131 in the product, 0.1809 in maximum, 0.2021 in case of minimum, and finally 0.1631 in case of voting. The mean error rate for the combined classifiers was 0.2123, which was less than the mean error rate of the base-level classifiers.

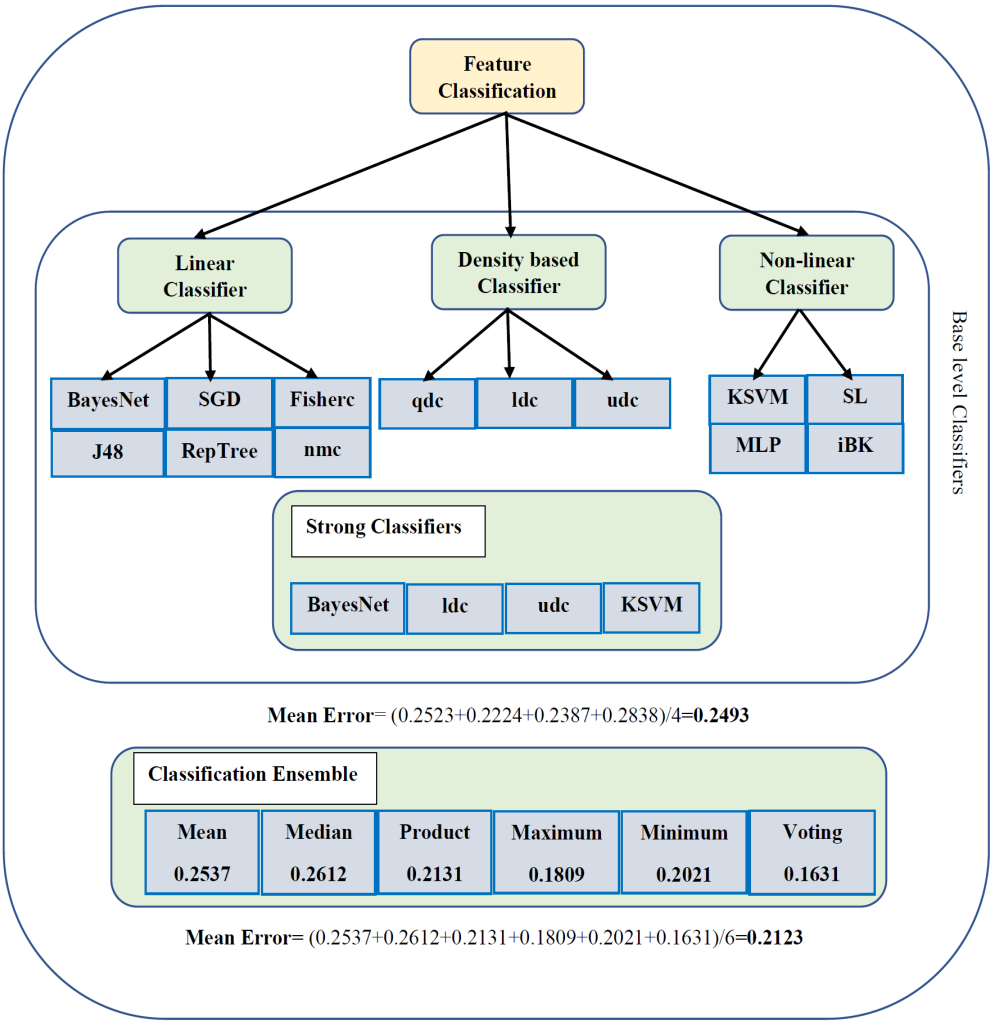

Figure 3. Performance evaluation of feature organisation by base-level classifiers and fusion approach

```
1:  a = dataset((data),genlab([20 20 20],[1;2;3;]));
2:  [trainset,testset] = gendat(a,0.70); %Divide dataset into two equal halves
3:  w1 = BayesNet; % Untrained classifiers w1, w2, w3 and w4
4:  w2 = ldc;
5:  w3 = udc;
6:  w4=KSVM:
7:  w = {w1,w2,w3,w4}; % Combining Classifiers
8:  v = trainset * w; %  % Training of Classifiers
9:  disp([newline 'Errors for individual classifiers']
10: testc(testset, v); % Display errors

        Errors for individual classifiers

        Test results result for:
                    clsf_1 : BayesNet
                    clsf_2 : ldc
                    clsf_3 : udc
                    clsf_4 : KSVM


                     clsf_1    clsf_2     clsf_3    clsf_4

                    0.2523   0.22240    0.2387    0.2838
```

Figure 4. Concurrent training of a set of four base-level classifiers

```
11: comb_base = [v{:}]; %Combining Classifiers into Cell array
12: wc = {meanc,medianc, prodc,maxc,minc,votec}; % Combining Rules into Cell array
13: vc = comb_base * wc; % Base Level classifiers combined by rules
14: testc(testset,vc); % Testing on 30% test data

                    Test results result for:

                     clsf_1 : Mean combiner
                     clsf_2 : Median combiner
                     clsf_3 : Product combiner
                     clsf_4 : Maximum combiner
                     clsf_5 : Minimum combiner
                     clsf_6 : Voting combiner

                   clsf_1     clsf_2     clsf_3     clsf_4     clsf_5     clsf_6

                  0.2537     0.2612     0.2131     0.1809     0.2021     0.1631
```

Figure 5. Combing base-level classifiers

3. RESULTS AND DISCUSSION

To assess the viability of our strategy, a few examinations utilizing the pointers referenced previously. ADNI dataset utilized in this paper contained a total of 668 subjects with 146 AD (21.86%) cases, 336 MCI (50.30%) cases, and 186 CN (27.85%). The investigations were performed in MATLAB R2019b, and WEKA toolkit on a computer system fortified with Intel Core i5-4590 with 3.30GHz, RAM of 8GB and 64-Bit operating system.

The proposal was validated on the benchmark ADNI dataset using various different evaluation parameters. Initially, the "testc" routine of the MATLAB was employed for carrying out the validation estimations for the trained classifier on a test dataset. Figure 6 depicts the performance of the top-performing standalone classifiers and classier merging strategy using accuracy, error rate, and AUC as the evaluating metrics. The four best accomplishing classifiers BayesNet, ldc, udc, and KSVM delivered the greatest results from the thirteen classifiers tested with their accuracy 74.77%, 71.62%, 77.76%, and 76.13%, respectively. The ensemble method decorated from the same four algorithms yielded an accuracy of 74.63% in mean, 73.88% in the median, 78.69% in the product, 81.91% in maximum, 79.79% in minimum, and lastly 83.69% in voting case (Figure 6). Clearly, the results in the classifier fusion methodology got enhanced than the standalone cases. The mean error rate of 0.2493 was generated by the standalone trained classifiers on the concealed test dataset. In contrast, the classifier merging strategy generated a mean error value of 0.2123 using different combining rules. The results clearly indicated the classifier merging approach (using combining rules) accomplishes improved performance values than the individual classifiers by 0.037 reductions in mean error rate. Additionally, the voting combination rule achieved the best overall results than other combining rules used (Figure 6).

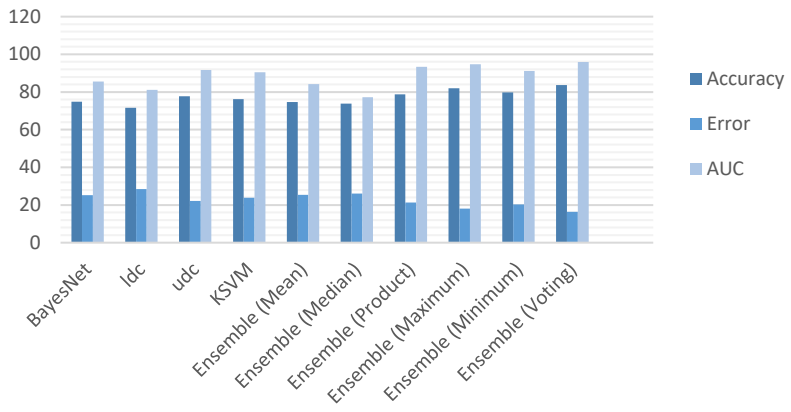

Figure 6. Performance analysis of classifier merging strategy using various evaluation parameters

Since the accuracy metric could be deceptive when the data were imbalanced, which was mainly the case in the medical domain, credit card fraud, etc. The outcomes proposed that model assessment measurements might uncover more about the appropriation of classes than they did about the genuine exhibition of models when the information was imbalanced because the model converged towards the majority class. So, to enhance the subjectivity of the proposal, the TP rate and FP rate and validated the same using AUC from receiver operating characteristics (ROC) were computed (Figure 7).

Figure 7 demonstrates the performance values of the classifier merging strategy utilizing several combining rule algorithms that embrace mean, median, product, maximum, minimum and voting. The finest outcomes were returned by the "voting combiner" with a value of 16.31% error rate. This was meticulously trailed by the "maximum combiner," with a value of 18.09%, followed by "minimum combiner" with a value of 20.21%, followed by "product combiner" with a value of 21.31%, followed by "mean combiner" with a value of 25.37%, and finally by "median combiner" with an error rate of 26.12%. The TP rate and FP rate values were calculated on the individual top four algorithms and then on various classification merging strategies. BayesNet algorithm got an accuracy of 74.77% and AUC of 0.856, ldc accomplished an accuracy of 71.62 and AUC of 0.811, udc with an accuracy of 77.76 and AUC of 0.916, and KSVM with an accuracy of 76.13 and AUC of 0.904. Similarly, the parameter values of other ensemble algorithms using various merging strategies were calculated.

The ensemble of the top four algorithms (BayesNet, ldc, udc, KSVM) with mean as the merging approach acquired an accuracy of 74.63 and AUC of 0.841. In the same way, the median merging approach acquired an accuracy of 73.88 but with the lowest AUC of 0.772. The product merging strategy obtained an accuracy of 78.69 and AUC of 0.933, the minimum combining strategy with an accuracy of 79.79 and AUC of 0.919. The second highest AUROC value of 0.946 was yielded by the ensemble of four algorithms with a maximum of probabilities as parameter setting and represented at number seven in ROC curve, and finally, the voting merging strategy topped the list with an accuracy value of 83.69 and AUC of 0.959 and made it the highest yielded AUC among the lot. Also, the ROC curve of the said ensemble was nearer to the upper left corner (Number 8), which again justified it being the best performer (Figure 7).

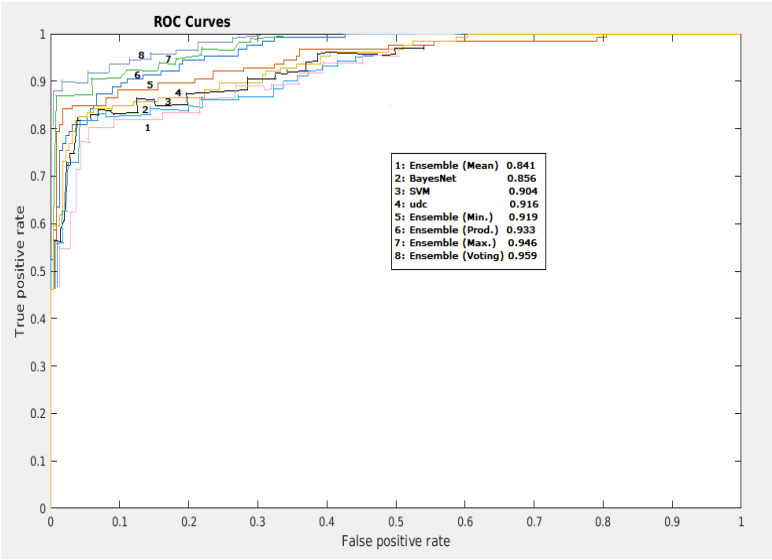

Figure 7. ROC analysis of the classifier merging strategy

The values of TP rate, FP rate and AUC definitely revealed the worthiness of the proposed method as compared to their individual counterparts. Similarly, the combinations for various methods and algorithms with diverse other parameter settings favor the clinical significance of ensemble methods for AD detection upon validating with various parameters.

4. CONCLUSION

Neuroimaging results realized an original classifier fusion idea for the use of which was to help in diagnosing AD. Initially, the base classifiers from diverse areas of machine learning on the ADNI dataset were examined. The four best-performing models had all experienced a drop in performance. The four best performing models were caught viz. BayesNet, ldc, udc, and KSVM, and ensembled to deliver a brilliant classifier combining plan. The six different combiner selection techniques were utilized to plan classifier groups. In the study it is found that the accuracy rate of fusing classifiers with consolidating rule was more noteworthy than joining the base-level classifiers with no standard normal classifier accuracy. The outcomes were better than one-by-machine learning. Optimality, combined with greater classificatory freedom, offers enormous potential for future applications. Future directions for the multidimensional analysis for disease diagnosis are nearly limitless.

REFERENCES

Chien, Y. W., Hong, S. Y., Cheah, W., Yao, L. H., Chang, Y. L., and Fu, L. (2019). An automatic assessment system for Alzheimer's disease based on speech using feature sequence generator and recurrent neural network. *Scientific Reports*, 9, 19597.

Dai, P., Sridhar, F. G., Bauer, M., and Borrie, M. (2015). Bagging ensembles for the diagnosis and prognostication of Alzheimer's disease. In *Proceedings of Thirtieth AAAI Conference on Artificial Intelligence*, pp. 3944-3951. Arizona, USA.

Dietterich, T. G. (2000). Ensemble methods in machine learning. In *Proceedings of the First International Workshop on Multiple Classifier Systems*, pp. 1-15. Berlin, Germany.

Eom, J., Jang, H., Kim, S., Jang, J., and Hwang, D. (2019). Study on discrimination of Alzheimer's disease states using an ensemble neural network's model. In *Proceedings of SPIE Medical Imaging 2019: Computer-Aided Diagnosis,* pp. 16-21. California, USA.

He, X., Chen, L., Li, X., and Fu, H. (2019). Brain image feature recognition method for Alzheimer's disease. *Cluster Computing*, 22(4), 8109-8117.

Kumar, P. R., Arunprasath, T., Rajasekaran, M. P., and Vishnuvarthanan, G. (2018). Computer-aided automated discrimination of Alzheimer's disease and its clinical progression in magnetic resonance images using hybrid clustering and game theory-based classification strategies. *Computers and Electrical Engineering*, 72, 283-295.

Lahmiri, S., and Shmuel, A. (2018). Performance of machine learning methods applied to structural MRI and ADAS cognitive scores in diagnosing Alzheimer's disease. *Biomedical Signal Processing and Control*, 52, 414-419.

Lebedev, A. V., Westman, E., Van Westen, G. J. P., Kramberger, M. G., Lundervold, A., Aarsland, D., Soininen, H., Kłoszewska, I., Mecocci, P., Tsolaki, M., Vellas, B., Lovestone, S., and Simmons, A. (2014). Random forest ensembles for detection and prediction of Alzheimer's disease with a good between-cohort robustness. *NeuroImage: Clinical*, 6, 115-125.

Li, K., Liu, Z., and Han, Y. (2012). Study of selective ensemble learning methods based on support vector machine. In *Proceedings of the International Conference on Medical Physics and Biomedical Engineering*, pp. 1518-1525. Paris, France.

Nisbet, R., Elder, J., and Miner, G. (2009). Model evaluation and enhancement. In *Handbook of Statistical Analysis and Data Mining Applications*, 2nd ed., pp. 285-312. Cambridge, Massachusetts: Academic Press.

Orimaye, S. O., Wong, J. S. M., Golden, K. J., Wong, C. P., and Soyiri, I. N. (2017). Predicting probable Alzheimer's disease using linguistic deficits and biomarkers. *BMC Bioinformatics*, 18, 34.

Papakostas, G. A., Savio, A., Graña, M., and Kaburlasos, V. G. (2015). A lattice computing approach to Alzheimer's disease computer-assisted diagnosis based on MRI data. *Neurocomputing*, 150(A), 37-42.

Pozueta, A., Rodríguez, E. R., Higuera, J. L. V., Mateo, I., Juan, P. S., Perez, S. G., Berciano, J., and Combarros, O. (2011). Detection of early Alzheimer's disease in MCI patients by the combination of MMSE and an episodic memory test. *BMC Neurology*, 11, 78.

Sadek, R. A. (2013). Regional atrophy analysis of MRI for early detection of Alzheimer's disease. *International Journal of Signal Processing, Image Processing and Pattern Recognition*, 6(1), 49-58.

Sankari, Z., and Adeli, H. (2011). Probabilistic neural networks for EEG-based diagnosis of Alzheimer's disease using conventional and wavelet coherence. *Journal of Neuroscience Methods*, 197(1), 165-170.

Shaikh, T. A., Ali, R., and Beg, M. M. S. (2020). Transfer learning privileged information fuels CAD diagnosis of breast cancer. *Machine Vision and Applications*, 31, 9.

Shaikh, T. A., and Ali, R. (2019). Automated atrophy assessment for Alzheimer's disease diagnosis from brain MRI images. *Magnetic Resonance Imaging*, 62(2), 167-173.

Talia, D., Trunfio, P., and Marozzo, F. (2016). Introduction to data mining. In *Data Analysis in the Cloud: Computer Science Reviews and Trends,* 3rd ed., pp. 1-25. Amsterdam: Elsevier.

Tessitore, A., Santangelo, G., Micco, R. D., Vitale, C., Giordano, A., Raimo, S., Corbo, D., Amboni, M., Barone, P., and Tedeschi, G. (2016). Cortical thickness changes in patients with Parkinson's disease and impulse control disorders. *Parkinsonism Related Disorders*, 24, 119-125.

Thal, D. R., Beach, T. G., Zanette, M., Heurling, K., Chakrabarty, A., Ismail, A., Smith, A. P. L., and Buckley, C. (2015). [18F]flutemetamol amyloid positron emission tomography in preclinical and symptomatic Alzheimer's disease: specific detection of advanced phases of amyloid-β pathology. *Alzheimer's & Dementia*, 11(8), 975-985.

Tierney, M. C., Yao, C., Kiss, A., and McDowell, I. (2005). Neuropsychological tests accurately predict incident

Alzheimer disease after 5 and 10 years. *Neurology*, 64(11), 1853-1859.

Varatharajan, R., Manogaran, G., Priyan, M. K., and Sundarasekar, R. (2018). Wearable sensor devices for early detection of Alzheimer disease using dynamic time warping algorithm. *Cluster Computing*, 21(6), 681-690.

Zheng, X., Shi, J., Zhang, Q., Ying, S., and Li, Y. (2017). Improving MRI-based diagnosis of Alzheimer's disease via an ensemble privileged information learning algorithm. In *Proceedings 14th IEEE International Symposium on Biomedical Imaging*, pp. 456-459. Melbourne, Australia.