An Intelligent Healthcare System for Automated Alzheimer's Disease Prediction and Personalized Care

Tawseef Ayoub Shaikh
Baba Ghulam Shah Badshah University, Rajouri, India
Rashid Ali, Waseem Ahmad Mir ,Izharuddin
Aligarh Muslim University,Aligarh India

Alzheimer's disease has posed the greatest threat among all the different types of neurodegenerative problems as it has assaulted humankind at quick pace than the others. Its manual revelation has become clinically insignificant because of the in expertise, high rate of false positives (FP) and false-negatives (FN). To reduce the false positive/ false negative rate, this paper frames a quick, affordable, and objective judgement of AD with a novel data mining method coalescing a global Maximum Relevance and Minimum Redundancy (MRMR) based filter heuristic with a globally optimised wrapper heuristic GANNIGMA with the intention of minimalising the consequence of an imbalanced healthcare dataset. The optimal feature subset yielding the best performance are utilised for model training of Decision Tree, k-NN, and SVM algorithms. The trial results on benchmark ADNI dataset using the proposed model displayed the Decision Tree attains TP rate of 0.778, and AUC of 0.798, k-NN acquires 0.764 TP Rate and 0.784 AUC, and SVM attains 0.997 TP Rate, and 0.996 as AUC. The results are far healthier than the separate results of these algorithms attained on the same dataset with fewer optimal feature subsets.

Keywords: Alzheimer's disease, Attribute Subset Selection, Automated diagnosis,Parkinson's disease, Quantile hyper-spheres, Smart healthcare system..

1. INTRODUCTION

Alzheimer's disease is the greatest public dementia among neurological disorders. Roughly 50 million individuals are experiencing dementia where AD represents 70–80% of the cases. Throughout the globe, 75.6 million people will be experiencing AD by 2030, and 131.5 million by 2050 . It has already exaggerated 26.6 million people globally in 2006, 44.4 million in 2013, and has been anticipated to influence 1 out of 85 individuals by 2050Adeli et al. [2008a]. The pace of commonness of AD worldwide is disturbing, and in each 3s, an individual falls prey to it. Promotion is additionally a huge factor in dismalness, as AD is positioned among the sixth powerful causes for deaths in the United States, and the fifth lashing means for deaths among those age 65 and more. The number of passings identified with coronary illness, stroke and prostate disease has diminished in between 2000 and 2015 anyway passings from AD expanded 123% Gaudiuso et al. [2020a]. The level of individuals with AD increments with age: 3% people of age 65–74, 17% people of age 75–84 and 32% people of age 85. The brain changes that arose from Alzheimer's may start 20 years or more before any indications Gaudiuso et al. [2020b].A large portion of the dementia cases (roughly 70%) will happen in low- and middle-income nations due to the expanded life expectancy around the world, as portrayed in Figure 1. Also, the financial burden of dementia presents an extraordinary cost for society. The worldwide yearly expense for dementia care was assessed at US$604 billion in 2010, which is double the Exxon Mobil income for the very year, and it is anticipated to amplify by 85% by reaching to 2 trillion dollars till 2030 Head et al. [2012a].The $818 billion costs for battling the disease has been anticipated only in 2015 Head et al. [2012b; Sado et al. [2018]. It may take 3 to 9 years for the AD to advanceAssociation et al. [2018]. The present-day manual means of AD detection generates accuracy values in between 85 to 95%, and depends on experienced clinicians, careful and comprehensive testing meetings, just as expensive and scant neuroimaging devices and intrusive techniquesShaikh and Ali [2019] . Tragically, clear AD finding is just affirmed through posthumous assessment of cerebrum tissue. These requirements limit the execution of early AD analysis in low-pay nations, distant and rural regions, just as in metropolitan zones where waiting occasions for non-crisis MRIs can be in the request for months. Mild cognitive impairment (MCI) is one of the promising phases of the AD,

which is especially perceived as its preclinical phase that may fill in as an objective speciality for early interventional techniques. Exploration contemplates showing that an estimation of 8–15% of MCI patients advances into AD stage yearly, while around 1–2% of people who were recently considered as healthy, obtained AD inside the same time period Amezquita-Sanchez et al. [2016]. Additionally, the hippocampus, a grey matter (GM) structure of the temporal lobe, is influenced during the beginning phases of AD. Thus, by analysing the hippocampus volume utilising magnetic resonance imaging (MRI), the illness level can be identified. It is very hard and tedious to estimate the hippocampus volume manually and programmed segmentation; thus, it becomes the ideal choice to sidestep these limitations and get the AD biomarkersKoh et al. [2020]. Figure 1 below gives the incidence rate of AD cases throughout the globe by 2030 Cassani et al. [2017].

Worldwide dementia prevalence in 2013 and predictable upsurge by 2030

The recent novel devices and biomarker tests have made it conceivable to analyse AD utilising progressed diagnostics. There exists a wide range of invasive and non-invasive neuroimaging advancements which are utilised for AD determinations, for example, Magnetic Resonance Imaging (MRI), functional MRI (fMRI), Positron Emission Tomography (PET), Computerized Tomography (CT), Electroencephalography (EEG), Magnetoencephalography (MEG) and Cerebrospinal fluid (CSF) biomarkers. A computer-aided diagnosis (CAD) system is the archetype from these neuroimaging data that help doctors and clinicians to improve medical services frameworks for ADAdeli et al. [2008b; Pich et al. [2014; Ossenkoppele et al. [2015] . MRI ends up being the highest quality level neuro-imaging methodology to assess the brain anatomic auxiliary and pathophysiological dimensions because of its delicate tissue natureNiazi et al. [2018; Suk et al. [2014].

1.1 Related Review

The computerised diagnosis of AD has engrossed prodigious devotion in recent years because of the effective neuroimaging innovations like structural magnetic resonance imaging (s-MRI) Papakostas et al. [2015; Aguilar et al. [2013; Westman et al. [2012], functional MRI (fMRI) Fan et al. [2008], and diffusion tensor imaging (DTI)Grana et al. [2011], along with positron emission tomography (PET), and single-photon emission computed tomography (SPECT) Hanyu et al. [2010]. The research has witnessed a huge trend in the automated AD diagnosis employing pattern recognition and AI strategies Huang and Aviyente [2008; Magnin et al. [2009; Kuncheva et al. [2010]. Different machine learning methods and algorithms have been executed for AD analysis by means of the influence of numerous numbers of classifiers together rather than one and thus making an ensemble of classifiersHinrichs et al. [2011; Liu et al. [2012; Seeley et al. [2014]. A fundamental challenge that remains unsolved in the neuroimaging field is the small sample feature size problem Maitra and Chatterjee [2006; Gonzlez-Castro et al. [2017; Ding and Peng [2005a]. The deep learning is the trending topic in the present area and has been widely used

in computerised AD detectionZhao and He [2014; McCrackin [2018; Jain et al. [2019]. Various different kinds of neuroimaging datasets like MRI, PET, SPECT Ghorbanian et al. [2012; Segovia et al. [2013], EEG Young et al. [2012; Fiscon et al. [2014; Ghorbanian et al. [2015], and multi-modal neuroimaging datasets Walhovd et al. [2010; Apostolova et al. [2010; Hinrichs et al. [2009] have been employed for computerised AD/MCI identification.

1.2 Organisation

The rest of the paper is organised out as: Section 2 discusses expansively the standard benchmark datasets that are contributing directly to accessing the performance of the proposed work. The section discusses the pre-processing techniques used for MRI image analysis and the performance indices used for evaluating the model. The proposed approach is thoroughly discussed in Section 3. The proposal is discussed in the form of flowcharts, equations, and the implementation of the scheme is also explained in the same section. The empirical and investigational results are discussed, interpreted, and tabulated in Section 4. At last, the conclusions drawn from the presented study are made in section 5.

2. MATERIALS AND METHODS

Both datasets, benchmark and local, employed in this work are the theme of explanation in this section. Starting pre-handling strategies with legitimate feature vector abstraction from input raw MRI brain images with a kind of the AI classifiers are likewise essential for this space.

2.1 Datasets

Alzheimer's disease Neuroimaging Initiative (ADNI) database is employed as a benchmark dataset in this work "www.adni.loni.usc.edu". It was launched in 2004 with the target of monitoring the progression of MCI and AD engaging several neuroimaging modalities like MRI, PET, other biological indicators, medical and neuropsychological inspections. As specified in Table I, the data of 303 contributors composing 158 as AD and 145 as HC is explored using high-resolution T1-weighted s-MRI, and the same is embodied as a mean ±standard deviation. While as gender presents p-value from a χ^2 test, for other sample characteristic, p-values are presented on right columns for two-sample t-tests. The second local dataset analysed in this study is conquered from 2 known medical imaging centres of Srinagar, India, specifically Medicare Diagnostic Center and DM diagnostic clinic over the last two years Banday and Mir [2017].

Table I: ADNI Dataset Used in this Study

Characteristics	AD	HC	p-Value
Sample Size	158	145	
Gender (Male/Female)	86/72	71/74	0.342
Age (Years)	75.21 ± 7.47	75.76 ± 4.42	0.432
CDR	4.68±1.740	.04± 0.14	less than0.000001
MMSE	23.30±2.05	29.14± 0.94	less than0.000001

Every image instance here consists of data taken from three sections, i.e., axial, coronal and sagittal. Depending upon the parameter assortment of axial resolution and slice thickness, the number of images differs across three segments.

2.2 MRI Data Pre-Processing

The eradication of unnecessary noise and artefacts from the raw data creates a healthier dataset with improved results using the pre-processing data approaches. To assure that each image voxel resembled with the identical functional location, all MRI data were initially spatially normalised by Statistical Parametric Mapping (SPM12), and Brain Voyager software Carballido-Gamio et al. [2017]. Nonbrain tissues for example neck and skull are removed from the image' scans using FSL-BET toolbox Smith [2002], which accomplishes brain withdrawal by approximating the intensity histogram-founded threshold, the centre of the gravity, and radius of the sphere of the brain's

surface. Then, FSL-MCFLIRT toolbox performs motion correctionJenkinson et al. [2002]. Also, FEAT Woolrich et al. [2001] module of FSL library, is used for Slice timing correction. Intensity normalisation is performed on data for guaranteeing that respective volume has an identical mean intensity. Also, Spatial smoothing is performed for plummeting the noise levels while conserving the original data. For registering images in MNI152 space, a linear conversion with 12 DOF (such as translation, scaling, shear, and rotation) is performed. The raw data is managed and transformed into 2D images by employing the aforesaid pre-processing approaches. Along these lines, we have made a dataset that was utilised for preparing the models.

2.3 Quality Assessment

The given investigation engages the associated quality evaluation approaches for demonstrating the competence of the proposed CAD framework. Accuracy: It is the proportion of accurate expectations divided by the total number of estimates. It is characterised as the potential of the model to hand-picked all instances that necessitate to be nominated and dispose of all instances that require to be disposed of Piramuthu [2004].

$$Accuracy = \frac{TP + TN}{TP + FP + TN + FN} \tag{1}$$

Sensitivity: It is essentially the accuracy of positive cases; i.e. how decent a test is in noticing positive disease. It is the competence of a prototypical to select all the cases that are vital to be selected.

$$Sensitivity = Recall = HitRate = TruePositiveRate = \frac{TP}{TP + FN} \tag{2}$$

False Positive Rate: It is the proportion of the number of negative instances mistakenly considered as positive (false positives) and the total number of real negative instances.

$$FalsePositiveRate = \frac{FP}{N} = \frac{FP}{FP + TN} \tag{3}$$

Precision: It is the percentage of true positive and the instances that are wrongly classified as positive, i.e. false positive.

$$Precision = \frac{TP}{TP + FP} \tag{4}$$

F-measure: Articulated basically by the measurements of precision and recall this measurement is utilised regularly in factual investigation and can be figured as the extent of the fluctuation among gatherings and the difference among gatherings.

$$F1 = 2 * \frac{Precision * Recall}{Precision + Recall} \tag{5}$$

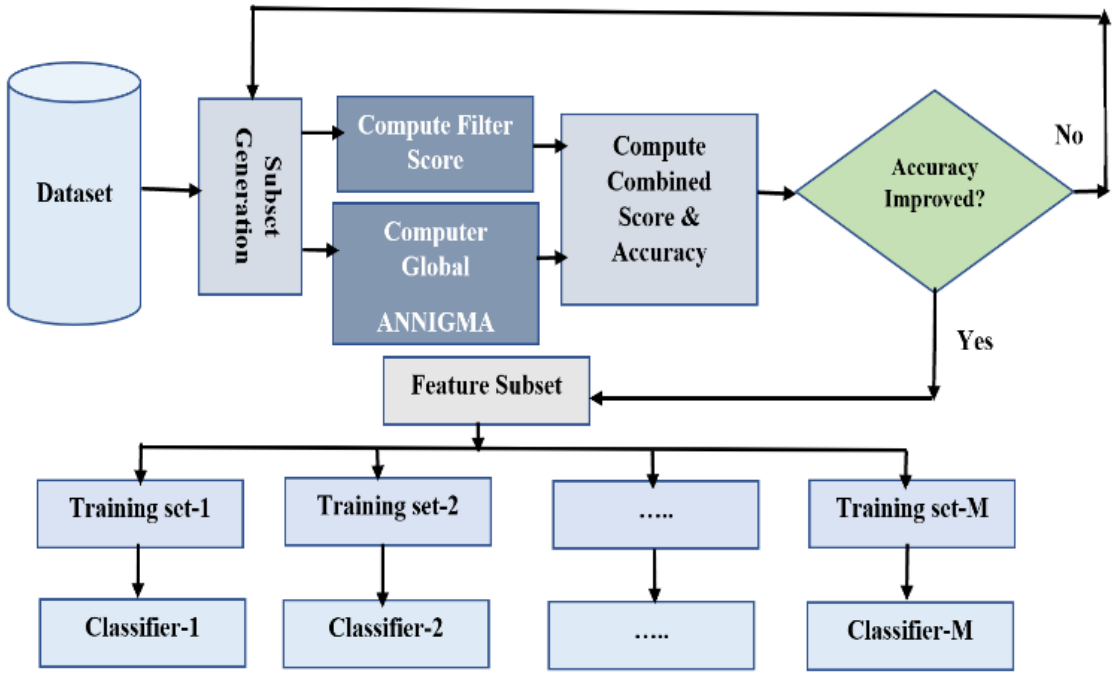
Area under receiver operating characteristics (AUROC): In a ROC curve, the of true positives (TP rate) is plotted against the false positives (FP rate). The estimations of the FP rate (1 – specificity or TN rate) is plotted on the horizontal axis and the TP rate estimations (sensitivity or recall) on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

3. PROPOSED METHODOLOGY

This section deals with the novel proposal and the prime motivation behind it.

3.1 Hybrid Feature Selection Using GANNIGMA

Here, the approach matures a hybrid feature selection to produce the indicative decision that is grounded on a globally optimised Artificial Neural Network Input Gain Measurement Approximation (GANNIGMA). The proposed GANNIGMA method unveils the most significant features,



Cataloguing architecture of the proposed novel feature selection method

which are then employed for various model training for improving the classification performances. The feature selection methodology is outlined in 2.

$$MRMR = \frac{1}{maxs|S|} \sum_{fp=s} I(F_p; C) - \frac{1}{|S|^2} \sum_{p,q=s} I(F_p; F_q) \tag{6}$$

3.1.1 *GANNIGMA Hybrid Feature Selection using Global Optimisation and Artificial Neural Network (ANN).* The inherent associations between the standalone analysis feature and AD class can be revealed by the Filter approach; however, the approach doesn't practice any presentation assessment standards grounded on accuracies. No doubt, the filter strategy is the cheapest, but it has the limitation in not choosing the significant AD features in terms of presentation. Wrapper approach might be computationally expensive but subset features picked from it can be guaranteed to be optimal as it employs accuracy founded performance evaluation during training. The proposed hybrid method receives compensations of the balancing possessions of both the methods and assimilates the acquaintance about the inherent association amid a specific feature with conforming class projected by the filter in the wrapper search process. So, mutual information (MI) grounded Maximum Relevance Minimum Redundancy (MRMR) strategy is imparted to the covering heuristic in our proposed AD highlight choice methodology for crossbreeding filter and wrapper approaches Hsu et al. [2002a]. An ANN is taken as a wrapper, and wrapper heuristic is computed employing Artificial Neural Network Input Gain Measurement Approximation (ANNIGMA). Figure 2 diagrammatically depicts the hybridisation of the wrapper and filter approaches. Also, a speed up search process is projected by including the ranking score from filter feature with the wrapper method. Then an optimal feature subset is explored by an induction algorithm in the wrapper training process. The maximum relevance (MR) [Huda et al., 2010] algorithm subsidised to severance while choosing the attributes that are extremely pertinent to class but enormously connected. Therefore, a Minimum Redundancy (MR) redundancy function is introduced into the MR algorithm as in Eq. (6). where $I(F_p, F_q)$ is the mutual information between the features F_p and F_q

3.1.2 *Global ANNIGMA (GANNIGMA) Score Calculation.* The wrapper training of ANN in which the input, hidden and output layer are signified as i, j, k, is used for the (ANNIGMA) value calculation Hsu et al. [2002b]. The logistic activation function is represented as "Q" as below:

$$Q(z) = (\frac{1}{1 + exp(-z)}) \tag{7}$$

The system yield is given by Eq. (8) after incorporating a linear function for first and second layer. Here is the input feature

$$O_k = \sum_j Q(\sum_i F_i * W_i j) * W_j k \tag{8}$$

The local gain for input feature is as:

$$LG_i k = \frac{\delta O_k}{\delta F_i} \tag{9}$$

The local gain is changed as in Eq. (10):

$$LG_i k = \sum_i (w_i j * W_j k) \tag{10}$$

becomes the local gain (LG) for feature- normalised based on a unity scale as (11) which is the ANNIGMA

$$ANNIGMA(F_i) = \frac{LG_i k}{maximum(i) LG_i k} \tag{11}$$

For an imbalanced dataset, the typical backpropagation training ANN technique delivers locally optimised parameter, which could have a negative impact on the performance. Therefore, a global optimisation method has been approved with the typical backpropagation training approach. An Algorithm for Global Optimisation Problem (AGOP) is plasticised for the optimal approximation of ANN limits in the training of ANN Hsu et al. [2002b; Huda et al. [2010]. N-fold cross-validation is adopted during the training of optimal wrapper heuristic for a Global ANNIGMA (GANNIGMA) scores. The GANNIGMA value over cross-validation can be found as Eq. (12):

$$GANNIGMA(F_p)_Average = (\frac{1}{n}) GANNIGMA(F_p)_1 + ... + GANNIGMA(F_p)_n \tag{12}$$

The MRMR score is derived from the most extreme significance score of a candidate AD highlight in candidate subset and repetition score between the candidate highlights from the remainder of the subset which has been uncovered in Eq. (13):

$$\max_{F_l \in F - F_{l-1}} \Big(\frac{1}{|S|} \sum_{J \in S} I(F; c) - \frac{1}{l-1} \sum_{F \in F_{l-1}} I(F; F_l) \Big) \tag{13}$$

At that point, an incremental search method technique is utilised for the figuring of the MRMR score for competitor AD highlight as the combination of the candidate highlights in the subset and rest of the highlights can be tremendous Ding and Peng [2005b]. Grounded on the corresponding weighted score as accessible in Eq. (14), the features are then ordered.

$$WeightedMRMR(F_p) = 1 - (\frac{RankFeature(F_p) inMRMR}{|F|}) \tag{14}$$

The final stage computes a collective value for the MRMR-GANNIGMA hybrid method as in Eq. (15):

$$CombinedScore(MRMRGANNIGMA; F_p) = WeightedMRMRScore(Fp) + GANNIGMA(F_P)_Average \tag{15}$$

4. RESULTS AND DISCUSSIONS

To assess the viability of our strategy, we do a few examinations utilising the pointers referenced previously. ADNI dataset utilised in this paper contains a total of 303 contributors with 158 (52.14 %) as AD and 145 (47.86%) as HC. The investigations are performed out in MATLAB R2019b, and on a computer system fortified with Intel Core i5-4590 with 3.30GHz, RAM of 8GB and 64-Bit Operating System. In the given study, an experimental examination is performed on the benchmark ADNI dataset using a two-stage experimental design. The first stage extracts the healthiest attribute subset among the immense subset space using the proposed GANNIGMA approach. The best combination of feature subset is employed for training various machine learning models. Following the attribute determination, classification of the test models is performed utilising Decision Trees, k-NN, and SVM calculations. The presentation of the model is likewise analysed and arranged. The given study diminishes the number of parameters and feature subsets rendering to their precision values. The proposed GANNIGMA guarantee to accomplish overpriced cataloguing performance and investigate the ideal blend of seven informative feature subsets from the gigantic feature subset space that can suggest imperative evidence to the clinicians for AD diagnosis. Originally, 34 subsets of attributes are accomplished by the GANNIGMA founded algorithm with the conveyance on the ADNI benchmark dataset. The decision of a couple of property subsets is alluring for the extra investigation as the number of reducts is excessively high. The calculation of the association of respective attribute subset with the decision (AD and HC) is completed using a "combination filtering" policy, i.e., choose the subsets which incorporate both the most noteworthy and least important features with the chosen component. The outcome exhibits the uppermost value is achieved by the attribute x9 (cortical thickness) and lowest value to attribute x17 (surface area), and thus, encouraging the toughest bending of the feature x9 with the class grouping, and x17 as the weakest one. For some indicator projects, feature significance (even solid importance) doesn't induce that an element must be in an ideal element subset, at times, a pitifully applicable element may likewise propel the forecast exactness. Hence, we embrace a "combination filtering" strategy to deal the ideal element subset, that is, we pick the subsets of traits which include both x9 with the most grounded importance and x17 with the most fragile significance as the ideal subset. As indicated by this technique, 27 characteristic subsets are disregarded, and all things being equal, the leftover 7 subsets are allotted to be the SVM Classifier contributions, as portrayed in II

Table II: The Seven Residual Attribute Sets

Attribute set number	Attribute sets	Classification accuracy on Benchmark ADNI dataset	Classification accuracy on Local MRI dataset
1	hippocampal atrophy, cortical thickness, surface area, grey matter volume	96.51	96.42
2	grey matter volume, White matter (WM), cortical thickness, surface area	96.87	95.98
3	grey matter volume, curvature, White matter (WM), cortical thickness, surface area	97.12	96.34
4	curvature, cortical thickness, hippocampal atrophy, Cerebral spinal fluid , surface area	97.35	95.23
5	hippocampal atrophy, cortical thickness, surface area, grey matter volume Cerebral spinal fluid , White matter (WM), curvature	98.23	96.87
6	Cerebral spinal fluid (CSF), White matter (WM), curvature, surface area, cortical thickness	97.79	94.24
7	White matter (WM), curvature, surface area, cortical thickness, hippocampal atrophy	96.21	94.23

The 7 abridged feature subsets conquered in the initial phase by the GANNIGMA based feature selection strategy is cross-validated using 10 folds to ensure the equal class spreading in the subset. Specifically, the subject dataset was arbitrarily partitioned into 10 sections, in which nine parts in each iteration is employed for training and the residual one part is used for as the testing set. Prior to developing the classifier, scaling of the data sets performed on each and every data

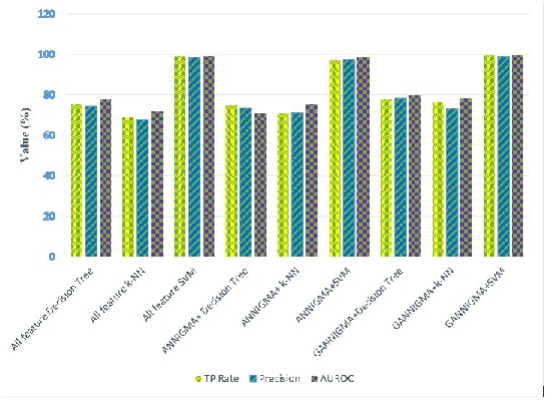
instance. The test collections are anticipated utilising the indicator model gained from the training of above advances appropriately. Each random training-test segment here is probably desired to gain dissimilar pair of best parameters as the subset fragmented randomly, and moulded by the 10- fold cross-validation. Along these lines, to mark the considerable examinations, the predetermined effort demeanours 1000 independent goes of the trials for each training-test segment, and the most noteworthy arrangement exactnesses are reckoned. In the fourth phase, optimal parameters train the predictor model as the output. The model trained with the ideal boundaries is yielded as the output in the fourth stage, which is used to envision the events in each testing set (both benchmark and local MRI datasets). The trained SVM model becomes ideal for fitting the actual data set classification. The SVM classifier visits the datasets and information highlights and requests them presenting to the commonality subsequent to its learning. The SVM is prepared with these confined improved highlights and gathers a more advantageous arrangement with limited computational expense and time. Table 2 shows the characterisation exactnesses on the testing information for the 7 quality subsets. Subset 5 accomplishes the maximum cataloguing accuracy among the 7 subsets in all the fold divisors, with 98.23% on the Benchmark ADNI dataset, and 96.87% on the local MRI dataset. The second-highest performance on the Benchmark ADNI dataset is acquired by the attribute subset number six with an accuracy of 97.79%, followed by attribute subset number four with the accuracy of 97.35%, followed by attribute subset number three with an accuracy of 97.12%, followed by attribute subset number two with an accuracy of 96.87%, followed by attribute subset number one with an accuracy of 96.51%, and finally the attribute subset number seven with an accuracy of 96.21%. In the case of local MRI dataset, attribute subset five once more surpassed the list by conquering an accuracy value of 96.87%, as depicted in Table 2. The outcomes are, without a doubt connoting the lessening in the error rates that arise from the misclassification issues immediately with the ideal choice of the training highlights. The SVM is proper to work on the genuine informational collection after learning about these 7 subsets of highlights. The SVM will give up on the previous information determined while preparing with the goal of ordering the new informational indexes and information highlights. At long last, the SVM offers improved order with limited computational expense and time after prepared with these obliged enlarged highlights. The subset number five is thus, selected for further analysis using various supplementary data mining algorithms besides SVM. The Decision Tree, k-NN, and SVM algorithms with various performance measures are tabulated in Table 3 and Figure 3. The SVM got the highest individual TP Rate of 0.992, lowest FP Rate of 0.028, Precision of 0.984, F-Measure of 0.991, and finally AUROC of 0.991. The Decision Tree individually got a TP rate of 0.754, and k-NN individually achieved a TP Rate of 0.689, which is the lowest among the group. All the three algorithms of Decision Tree, k-NN, and SVM are combined with the ANNIGMA scheme, and the new paradigm is evaluated again on the benchmark ADNI dataset. The TP Rate of Decision Tree and SVM algorithms got a slight decrease here from 0.754 to 0.749 in case of Decision Tree, from 0.992 to 0.971 in case of SVM. Only in the k-NN case, there occurred a lift in the TP Rate from 0.689 to 0.708. Similarly, all the three algorithms are merged with the proposed GANNIGMA scheme and once more evaluated on the ADNI benchmark dataset. In this case, the investigational results witnessed a hike in the TP Rate of all the three algorithms. The highest TP Rate in this combination is offered by the GANNIGMA+ SVM with a value of 0.997, followed by the TP Rate of GANNIGMA+ Decision Tree with the value of 0.778, and lastly the GANNIGMA+ k-NN with the TP Rate of 0.764

Table III: Assessment of cataloguing performance measures for various groupings of feature selection and cataloguing methods

Techniques	TP rate	FP rate	Precision	F Measure	AUROC
All feature Decision Tree	0.754	0.348	0.742	0.751	0.779
All feature k-NN	0.689	0.332	0.667	0.686	0.722
All feature SVM	0.992	0.028	0.984	0.991	0.991
ANNIGMA+Decision Tree	0.749	0.359	0.736	0.794	0.711
ANNIGMA+k-NN	0.708	0.321	0.712	0.704	0.753
ANNIGMA+SVM	0.971	0.034	0.975	0.971	0.984
GANNIGMA+Decision Tree	0.778	0.303	0.783	0.775	0.798
GANNIGMA+k-NN	0.764	0.308	0.732	0.744	0.784
GANNIGMA+SVM	0.997	0.004	0.991	0.994	0.996

The AUROC of the Decision Tree model got a minute reduction from 0.779 to 0.711 while

executing using ANNIGMA+ Decision Tree. But when AUROC got upliftment to a value of 0.798 using it in combination with GANNIGMA+ Decision Tree. The scenario reflected almost the same trend in case of SVM, where the AUROC values 0.991 is acquired using all features, and 0.984 while combining it with ANNIGMA+SVM. The AUROC value again here got amplified to 0.996 in case of GANNIGMA+SVM. Similarly, the k-NN algorithm achieved an AUROC of 0.722 while using all feature set, which got little diminished to 0.753 in combining it with ANNIGMA+ k-NN, but with GANNIGMA+ k-NN the AUROC value again upsurge to 0.784. The GANNIGMA heuristic in almost all the cases executed in the study performed well.



Performance evaluation of the proposal with standalone machine learning models using TP Rate, Precision, and AUROC parameters

The current findings exhibited that the ideal feature subset amalgam effectively coordinated different model properties, further improving classification execution. During the classification process, the mutual information (MI) investigation technique was performed to assess the commitments of chosen features. The mutual information (MI) weight got with different consolidated attributes was significantly more noteworthy than the weight acquired when just local attributes or subgraph attributes were embraced. With respect to the fundamental mechanism, this is probably going to be on the grounds that receiving various component subsets can coordinate integral organisation data, joining complementary network information and subgraph attributes, accordingly further improving classification exactness.

5. CONCLUSION

The works offer a novel data mining method coalescing a global optimisation-founded feature assortment strategy with the intention of minimalising the consequence of an imbalanced healthcare data. The proposal consolidates a Maximum Relevance and Minimum Redundancy (MRMR) filter heuristic with a globally optimised wrapper heuristic GANNIGMA. The execution of GANNIGMA optimisation process excerpts the best seven feature subsets out of the total 34 subset feature space. The optimally selected feature subset in the first step is employed for model training on various different data mining algorithms like Decision Tree, k-NN, and SVM. The trial results on benchmark ADNI dataset displayed the Decision Tree attains TP rate of 0.778, and AUC of 0.798, k-NN acquires 0.764 TP Rate and 0.784 AUC, and SVM attains 0.997 TP Rate, and 0.996 as AUC, using the proposed model. The results are far better than the individual results of these algorithms. In future, more optimisation techniques could be implemented for AD detection from neuroimaging data. Also, the union of multi-modal neuroimaging data can be an optimal choice for accurate and precise AD diagnosis.

References

Adeli, H., Ghosh-Dastidar, S., and Dadmehr, N. 2008a. A spatio-temporal wavelet-chaos methodology for eeg-based diagnosis of alzheimer's disease. *Neuroscience letters 444,* 2, 190–194.

Adeli, H., Ghosh-Dastidar, S., and Dadmehr, N. 2008b. A spatio-temporal wavelet-chaos methodology for eeg-based diagnosis of alzheimer's disease. *Neuroscience letters 444,* 2, 190–194.

Aguilar, C., Westman, E., Muehlboeck, J.-S., Mecocci, P., Vellas, B., Tsolaki, M., Kloszewska, I., Soininen, H., Lovestone, S., Spenger, C., et al. 2013. Different multivariate techniques for automated classification of mri data in alzheimer's disease and mild cognitive impairment. *Psychiatry Research: Neuroimaging 212,* 2, 89–98.

Amezquita-Sanchez, J. P., Adeli, A., and Adeli, H. 2016. A new methodology for automated diagnosis of mild cognitive impairment (mci) using magnetoencephalography (meg). *Behavioural brain research 305*, 174–180.

Apostolova, L. G., Hwang, K. S., Andrawis, J. P., Green, A. E., Babakchanian, S., Morra, J. H., Cummings, J. L., Toga, A. W., Trojanowski, J. Q., Shaw, L. M., et al. 2010. 3d pib and csf biomarker associations with hippocampal atrophy in adni subjects. *Neurobiology of aging 31,* 8, 1284–1303.

Association, A. et al. 2018. 2018 alzheimer's disease facts and figures. *Alzheimer's & Dementia 14,* 3, 367–429.

Banday, S. A. and Mir, A. H. 2017. Statistical textural feature and deformable model based brain tumor segmentation and volume estimation. *Multimedia Tools and Applications 76,* 3, 3809–3828.

Carballido-Gamio, J., Bonaretti, S., Kazakia, G. J., Khosla, S., Majumdar, S., Lang, T. F., and Burghardt, A. J. 2017. Statistical parametric mapping of hr-pqct images: a tool for population-based local comparisons of micro-scale bone features. *Annals of biomedical engineering 45,* 4, 949–962.

Cassani, R., Falk, T. H., Fraga, F. J., Cecchi, M., Moore, D. K., and Anghinah, R. 2017. Towards automated electroencephalography-based alzheimer's disease diagnosis using portable low-density devices. *Biomedical Signal Processing and Control 33*, 261–271.

Ding, C. and Peng, H. 2005a. Minimum redundancy feature selection from microarray gene expression data. *Journal of bioinformatics and computational biology 3,* 02, 185–205.

Ding, C. and Peng, H. 2005b. Minimum redundancy feature selection from microarray gene expression data. *Journal of bioinformatics and computational biology 3,* 02, 185–205.

Fan, Y., Resnick, S. M., Wu, X., and Davatzikos, C. 2008. Structural and functional biomarkers of prodromal alzheimer's disease: a high-dimensional pattern classification study. *Neuroimage 41,* 2, 277–285.

Fiscon, G., Weitschek, E., Felici, G., Bertolazzi, P., De Salvo, S., Bramanti, P., and De Cola, M. C. 2014. Alzheimer's disease patients classification through eeg signals processing. In *2014 IEEE Symposium on Computational Intelligence and Data Mining (CIDM)*. IEEE, 105–112.

Gaudiuso, R., Ewusi-Annan, E., Xia, W., and Melikechi, N. 2020a. Diagnosis of alzheimer's disease using laser-induced breakdown spectroscopy and machine learning. *Spectrochimica Acta Part B: Atomic Spectroscopy 171*, 105931.

Gaudiuso, R., Ewusi-Annan, E., Xia, W., and Melikechi, N. 2020b. Diagnosis of alzheimer's disease using laser-induced breakdown spectroscopy and machine learning. *Spectrochimica Acta Part B: Atomic Spectroscopy 171*, 105931.

Ghorbanian, P., Devilbiss, D., Simon, A., Bernstein, A., Hess, T., and Ashrafiuon, H. 2012. Continuous wavelet transform eeg features of alzheimer's disease. In *Dynamic Systems and Control Conference*. Vol. 45295. American Society of Mechanical Engineers, 567–572.

Ghorbanian, P., Devilbiss, D. M., Hess, T., Bernstein, A., Simon, A. J., and Ashrafiuon, H. 2015. Exploration of eeg features of alzheimer's disease using continuous wavelet transform. *Medical & biological engineering & computing 53,* 9, 843–855.

Gonzlez-Castro, V., Vals Hernandez, M. d. C., Chappell, F. M., Armitage, P. A., Makin, S., and Wardlaw, J. M. 2017. Reliability of an automatic classifier for brain enlarged perivascular spaces burden and comparison with human performance. *Clinical Science 131,* 13, 1465–1481.

Grana, M., Termenon, M., Savio, A., Gonzalez-Pinto, A., Echeveste, J., Pérez, J., and Besga, A. 2011. Computer aided diagnosis system for alzheimer disease using brain diffusion tensor imaging features selected by pearson's correlation. *Neuroscience letters 502,* 3, 225–229.

Hanyu, H., Sato, T., Hirao, K., Kanetaka, H., Iwamoto, T., and Koizumi, K. 2010. The progression of cognitive deterioration and regional cerebral blood flow patterns in alzheimer's disease: a longitudinal spect study. *Journal of the neurological sciences 290,* 1-2, 96–101.

Head, E., Powell, D., Gold, B. T., and Schmitt, F. A. 2012a. Alzheimer's disease in down syndrome. *European journal of neurodegenerative disease 1,* 3, 353.

Head, E., Powell, D., Gold, B. T., and Schmitt, F. A. 2012b. Alzheimer's disease in down syndrome. *European journal of neurodegenerative disease 1,* 3, 353.

Hinrichs, C., Singh, V., Mukherjee, L., Xu, G., Chung, M. K., Johnson, S. C., Initiative, A. D. N., et al. 2009. Spatially augmented lpboosting for ad classification with evaluations on the adni dataset. *Neuroimage 48,* 1, 138–149.

Hinrichs, C., Singh, V., Xu, G., Johnson, S. C., Initiative, A. D. N., et al. 2011. Predictive markers for ad in a multi-modality framework: an analysis of mci progression in the adni population. *Neuroimage 55,* 2, 574–589.

Hsu, C.-N., Huang, H.-J., and Dietrich, S. 2002a. The annigma-wrapper approach to fast feature selection for neural nets. *IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics) 32,* 2, 207–212.

Hsu, C.-N., Huang, H.-J., and Dietrich, S. 2002b. The annigma-wrapper approach to fast feature selection for neural nets. *IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics) 32,* 2, 207–212.

Huang, K. and Aviyente, S. 2008. Wavelet feature selection for image classification. *IEEE Transactions on Image Processing 17,* 9, 1709–1720.

Huda, S., Yearwood, J., and Strainieri, A. 2010. Hybrid wrapper-filter approaches for input feature selection using maximum relevance and artificial neural network input gain measurement approximation (annigma). In *2010 Fourth International Conference on Network and System Security*. IEEE, 442–449.

Jain, R., Jain, N., Aggarwal, A., and Hemanth, D. J. 2019. Convolutional neural network based alzheimer's disease classification from magnetic resonance brain images. *Cognitive Systems Research 57*, 147–159.

Jenkinson, M., Bannister, P., Brady, M., and Smith, S. 2002. Improved optimization for the robust and accurate linear registration and motion correction of brain images. *Neuroimage 17,* 2, 825–841.

Koh, J. E. W., Jahmunah, V., Pham, T.-H., Oh, S. L., Ciaccio, E. J., Acharya, U. R., Yeong, C. H., Fabell, M. K. M., Rahmat, K., Vijayananthan, A., et al. 2020. Automated detection of alzheimer's disease using bi-directional empirical model decomposition. *Pattern Recognition Letters 135*, 106–113.

Kuncheva, L. I., Rodríguez, J. J., Plumpton, C. O., Linden, D. E., and Johnston, S. J. 2010. Random subspace ensembles for fmri classification. *IEEE transactions on medical imaging 29,* 2, 531–542.

Liu, M., Zhang, D., Shen, D., Initiative, A. D. N., et al. 2012. Ensemble sparse classifi-

cation of alzheimer's disease. *NeuroImage 60,* 2, 1106–1116.

Magnin, B., Mesrob, L., Kinkingnéhun, S., Pélégrini-Issac, M., Colliot, O., Sarazin, M., Dubois, B., Lehéricy, S., and Benali, H. 2009. Support vector machine-based classification of alzheimer's disease from whole-brain anatomical mri. *Neuroradiology 51,* 2, 73–83.

Maitra, M. and Chatterjee, A. 2006. A slantlet transform based intelligent system for magnetic resonance brain image classification. *Biomedical Signal Processing and Control 1,* 4, 299–306.

McCrackin, L. 2018. Early detection of alzheimer's disease using deep learning. In *Canadian Conference on Artificial Intelligence*. Springer, 355–359.

Niazi, M., Karaman, M., Das, S., Zhou, X. J., Yushkevich, P., and Cai, K. 2018. Quantitative mri of perivascular spaces at 3t for early diagnosis of mild cognitive impairment. *American Journal of Neuroradiology 39,* 9, 1622–1628.

Ossenkoppele, R., Jansen, W. J., Rabinovici, G. D., Knol, D. L., van der Flier, W. M., van Berckel, B. N., Scheltens, P., Visser, P. J., Verfaillie, S. C., Zwan, M. D., et al. 2015. Prevalence of amyloid pet positivity in dementia syndromes: a meta-analysis. *Jama 313,* 19, 1939–1950.

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*, 37–42.

Pich, E. M., Jeromin, A., Frisoni, G. B., Hill, D., Lockhart, A., Schmidt, M. E., Turner, M. R., Mondello, S., and Potter, W. Z. 2014. Imaging as a biomarker in drug discovery for alzheimer's disease: is mri a suitable technology? *Alzheimer's research & therapy 6,* 4, 1–7.

Piramuthu, S. 2004. Evaluating feature selection methods for learning in data mining applications. *European journal of operational research 156,* 2, 483–494.

Sado, M., Ninomiya, A., Shikimoto, R., Ikeda, B., Baba, T., Yoshimura, K., and Mimura, M. 2018. The estimated cost of dementia in japan, the most aged society in the world. *PLoS One 13,* 11, e0206508.

Seeley, M., Clement, M., Giraud-Carrier, C., Snell, Q., Bodily, P., and Fujimoto, S. 2014. A structured approach to ensemble learning for alzheimer's disease prediction. In *Proceedings of the 5th ACM Conference on Bioinformatics, Computational Biology, and Health Informatics*. 605–606.

Segovia, F., Górriz, J., Ramírez, J., Salas-Gonzalez, D., and Álvarez, I. 2013. Early diagnosis of alzheimer's disease based on partial least squares and support vector machine. *Expert Systems with Applications 40,* 2, 677–683.

Shaikh, T. A. and Ali, R. 2019. Automated atrophy assessment for alzheimer's disease diagnosis from brain mri images. *Magnetic resonance imaging 62*, 167–173.

Smith, S. M. 2002. Fast robust automated brain extraction. *Human brain mapping 17,* 3, 143–155.

Suk, H.-I., Lee, S.-W., Shen, D., Initiative, A. D. N., et al. 2014. Hierarchical feature representation and multimodal fusion with deep learning for ad/mci diagnosis. *NeuroImage 101*, 569–582.

Walhovd, K., Fjell, A., Brewer, J., McEvoy, L., Fennema-Notestine, C., Hagler, D., Jennings, R., Karow, D., Dale, A., Initiative, A. D. N., et al. 2010. Combining mr imaging, positron-emission tomography, and csf biomarkers in the diagnosis and prognosis of alzheimer disease. *American Journal of Neuroradiology 31,* 2, 347–354.

Westman, E., Muehlboeck, J.-S., and Simmons, A. 2012. Combining mri and csf measures for classification of alzheimer's disease and prediction of mild cognitive impairment conversion. *Neuroimage 62,* 1, 229–238.

Woolrich, M. W., Ripley, B. D., Brady, M., and Smith, S. M. 2001. Temporal autocor-

relation in univariate linear modeling of fmri data. *Neuroimage 14,* 6, 1370–1386.

Young, J., Modat, M., Cardoso, M. J., Ashburner, J., and Ourselin, S. 2012. Classification of alzheimer's disease patients and controls with gaussian processes. In *2012 9th IEEE International Symposium on Biomedical Imaging (ISBI)*. IEEE, 1523–1526.

Zhao, Y. and He, L. 2014. Deep learning in the eeg diagnosis of alzheimer's disease. In *Asian conference on computer vision*. Springer, 340–353.

Tawseef Ayoub Shaikh is Assistant Professor at Baba Ghulam Shah Badshah University, Rajouri, India, besides he is Doctorate from Zakir Hussain College of Engineering and Technology, Aligarh Muslim University , Aligarh, India. Prior to this he has done M-Tech from Guru Nanak Dev University Amritsar India and B-tech in Computer Engineering from Islamic University of Science and Technology Jammu and Kashmir , India. He has a wide range of research experience in Medical Data Analysis and Artificial Intelligence. His research interests include Medical Image processing, Medical signal processing, cloud computing, Deep Learning, Soft Computing, particularly how search engines, advertising platforms can actively learn from interaction using Machine Learning, Data Mining methods, and Soft Computing.

Rashid Ali obtained his B.Tech. and M.Tech. from A.M.U. Aligarh, India in 1999 and 2001 respectively. He obtained his PhD in Computer Engineering in February 2010 from A.M.U. Aligarh. His PhD work was on performance evaluation of Web Search Engines. He has authored about 125 papers in various International Journals and International conference proceedings. He has presented papers in many International conferences and has also chaired sessions in few International conferences. He has reviewed articles for some of the reputed International Journals and International conference proceedings. His research interests include Web-Searching, Web-Mining, soft computing Techniques (Rough-Set, Artificial Neural Networks, fuzzy logic etc), Recommender Systems and Online Social Network Analysis

Waseem Ahmad Mir is currently a Research Scholar in Department of Computer Engineering, Aligarh Muslim University, Aligarh, India,working in the domain of medical image and signal processing using Artificial Intelligence. Prior to this he completed his M-Tech From Department of CS and IT, Maulana Azad National Urdu University Hyderabad, India. He completed his B-Tech from Department of Electronics and Communication Engineering, Islamic University of Science and Technology, Jammu and Kashmir, India. His research interests are Energy Efficient Routing in Wireless Sensor Networks, Medical Data Analysis, Artificial Intelligence, and Brain Disorder Detection using Machine Learning and signal processing techniques.

Izharuddin did his B.Sc. Engg. and M.Sc.Engg. and Ph.D from Z. H. College of Engineeringand Technology AMU, Aligarh, India,in 1988, 1991 and 2014, respectively. He worked for two years as a Hardware Engineerin Morsad Computer, in Oman,till 1997. In 1997 he joined the Department of Electronics Engineering at AMU, Aligarh, as an Assistant Professor. Currently he is Professor in the Department of Computer Engineering, AMU Aligarh 202002. His main area of research interest is low-power VLSI design (ASIC and FPGA based Design) for signal processing specifically security and surveillance applications and security of Medical signals, other area of interest is Device Modelling. His worked are published in various reputed journals. He is a member of Institution of Electrical and Electronics Engineers (IEEE) (U.S.A.) and Institute of Electronics and Telecommunication Engineers, IETE, India.