Tianjin DaxueXuebao (ZiranKexueyuGongchengJishu Ban)/
Journal of Tianjin University Science and Technology
ISSN (Online): 0493-2137
E-Publication: Online Open Access
Vol:54 Issue:04:2021
DOI 10.17605/OSF.IO/V9TX7

FEATURE SELECTION USING HYBRID KRILL HERD – GENETIC ALGORITHM AND STRUCTURE OPTIMIZED NEURAL NETWORK OPTIMIZATION FOR ALZHEIMER'S DISEASE

S. SUMANTH
Assistant Professor, Dept. of Computer Science, Smt. V.H.D. Central Institute of Home Science (A), Bengaluru

Dr. A .SURESH
Head, Dept. of Computer Science, Sona College of Arts and Science, Salem, India.

ABSTRACT

Alzheimer's disease (AD) is an incurable neurodegenerative disease that mainly affects the aged population. For early AD diagnosis, there is a requirement for effective automated techniques. Researchers have propounded numerous novel approaches for AD classification. Nevertheless, for more comprehensive knowledge about AD research, more effective learning techniques are essential. Application of feature selection has a major impact on the classification procedure's speed due to the removal of unnecessary features. The feature selects an optimized feature subset from a larger feature set. This work has proposed the feature selection utilizing a hybrid Krill Herd-Genetic Algorithm (KH-GA) algorithm and optimizes the Neural Network (NN). The computation of the NN's weights or optimization minimizes the function cost or error for attaining feasible outcome. The stochastic nature-inspired optimization algorithm known as the Krill Herd (KH) has been successfully applied to resolve numerous complex optimization problems. The KH algorithm's performance can sporadically get influenced by its poor exploration (diversification) ability. To attain the best global optima, the KH-GA algorithm boosts the KH algorithm's global (diversification) search ability. This proposed algorithm has been operated via the addition of the Genetic Algorithm's (GA) global search operator. This will boost the KH'sexploration around the optimal solution and kill individuals that shift towards the global best solution. In comparison to the other approaches, the experimental outcomes have demonstrated the proposed algorithm's superior performance.

Keywords: Alzheimer's disease (AD), Feature Selection, Genetic Algorithm (GA), Krill Herd (KH), Artificial Neural Network (ANN), and Structure Optimized Neural Network.

I. INTRODUCTION

A common type of dementia is Alzheimer's disease (AD). It is a progressive neurodegenerative disease that causes gradual yet continued memory deterioration and impairments intellectual abilities and other mental functions. AD can result in changes to the brain's structure. Despite the symptoms' gradual development, they will become worse as time passes by. At first, the patient experiences Mild Cognitive Impairment (MCI) that will gradually progress into AD. The Ad’s intermediate stage is the MCI. However, not all MCI patients become AD patients. Even though AD is presently incurable, early-stage detection can prevent this disease's progression. While 26.6 million human beings were suffering from AD in 2006, AD is predicted to affect 1 in 85 human beings worldwide by 2050, out of which, a high level of care [1] will be required for roughly 43% of the universal cases.

The highly advanced Magnetic Resonance Imaging (MRI) is utilized for clinical medicine and medical imaging. MRI is an efficient tool for studying the human brain's various states. The images from the MRIs are examined, diagnosed, and experimented upon for the brain's clinical analysis so as to obtain a lot of valuable information. With this information, it is possible to verify whether the examined brain is either abnormal or normal. The images' extracted data is huge. Thus, the determination of a conclusive diagnosis that is dependent on this raw data is quite difficult. For these types of cases, several image analysis tools are required for the MRI images' analysis and the conclusive information's extraction in order to classify into normal or abnormalities of the brain. With the availability of 3-D and 2-D images of the body's [2] various internal organs, there has been a swift growth in the MRI images' detail levels.

Definition for features is given as the objects of interest's characteristics. When features are chosen cautiously, they can represent the image's maximum relevant information offered for a lesion's complete characterization. Images and objects are analyzed with feature extraction methodologies for extraction of the most critical representative features of objects' diverse classes. Features are utilized as inputs to classifiers that allocate them to their representative as Feature space reduction for minimization of the computation time as well as enhancement of the prediction accuracy is done with feature selection's support. This is accomplished through the removal of features that are noisy, redundant, and irrelevant feature sub set that is capable of offering the best performance with regards to computation time [3] and accuracy will be chosen.

Being under the artificial intelligence's umbrella, machine learning has numerous tools for making statistical and probabilistic decisions dependent on previous learning. Classification of new events and prediction of new patterns in machine learning is done with past learning (training). In comparison to commonly-used statistical tools, machine learning is extremely powerful. For the attainment of effective outcomes in machine learning, there must be excellent comprehension of the problem and the algorithms' constraints. Hence, an even chance for success is possible if there has been proper execution of an experiment, careful and correct utilization of the training, and vigorous validation of the results. Moreover, every machine learning techniques and algorithms are quite distinct. For example, some techniques designed based on certain assumptions or certain data types would become inapplicable for other data types[4].

In the earlier decade, machine learning approaches are proved to be quite beneficial for AD diagnosis. Support Vector Machine (SVM), Artificial Neural Network (ANN), and deep learning are among the most extensively utilized classification approaches. The optimization problem's nature is the key difference between the SVM and the ANN. While a global optimal solution is given by the SVM, the ANN gives a local optimal solution. A critical step in [5] in SVM as well as in ANN is the feature extraction.

The accuracy of the AD data's classification is reliant on the problem type. For example, the accuracy is the least for MCI vs. AD, lesser for Control normal (CN) vs. MCI, and the highest for CN vs. AD. It is also quite difficult to classify MCI converters (MCIc) **vs**. non-converters (MCInc) and amnestic MCI (aMCI) vs. non-amnestic MCI (naMCI. Furthermore, the MRI scanners' generated 3-D data will result in huge-sized datasets. Therefore, it is necessary for this data analysis to use effective feature extraction, selection, and classification methods. These meta-heuristic methods are Non-deterministic Polynomial (NP)-hard, and therefore, more efficient in feature selection [6] when compared with statistical methods such as Correlation Based Features (CFS) and Mutual Information (MI). An optimized feature subset can be attained through effective use of the Particle Swarm Optimization (PSO) and the Genetic Algorithm (GA).

This work proposes feature selection that utilizes KH-GA for AD classification of MRI images. The rest of the investigation has been arranged in the following sections. The

Tianjin DaxueXuebao (ZiranKexueyuGongchengJishu Ban)/
Journal of Tianjin University Science and Technology
ISSN (Online): 0493-2137
E-Publication: Online Open Access
Vol:54 Issue:04:2021
DOI 10.17605/OSF.IO/V9TX7

associated literary works are outlined in Section Two. The various techniques employed in this work are detailed in Section Three. The experimental outcomes are discussed in Section Four, while the work's conclusions are given in Section Five.

II. RELATED WORKS

A new classifier of brain MRI was propounded by Neffati et al., [7],which was dependent on the novel Downsized Kernel Principal Component Analysis (DKPCA) as well as the multi-class SVM. AD MRIs' classification could be done with this proposed technique. At first, a multi-objective optimization technique would be employed for determining the kernel function's optimal parameter such that, simultaneously, there is the assurance of good classification outcomes as well as minimization of the number of retained principle components. The optimized DKPCA model is constructed by utilizing the optimal parameter. Secondly, there is application of DKPCA to the normalized features. Afterwards, the downsized features are given as the classifier's input to get the prediction as output. For validation of the proposed technique's effectiveness, DKPCA was gauged with synthetic data to verify its efficiency on dimensionality reduction, and then, the DKPCA-based technique was assessed on the Open Access Series of Imaging Studies (OASIS) MRI database, and the outcomes were found to be satisfactory in comparison to traditional techniques.

A model that was based on 12-layer Convolution Neural Network (CNN)was presented by Hossain et al., [8] for binary classification and for AD detection with Brain MRI data. The proposed model's performance was compared with certain currently-used CNN models with regards to the OASIS dataset's Receiver Operating Characteristic (ROC) curve, F1 score, recall, precision, and accuracy. This work's major contribution was a 12-layer CNN model which had a high accuracy of 97.75% when compared with this dataset's any other published currently-used CNN models. Side by side comparisons of the proposed model and the pre-trained CNN models (MobilenetV2, Xception, InceptioV3, and so on) were also offered in this work. The proposed model's superiority over the currently-used models was indicated from the experiments' outcomes.

Proposal for a novel model for AD with Brain Image Analysis (BIA) was given by Shankar et al., [9]. In the beginning, there was consideration of the image database for the unwanted region's removal. Later, an intermediate output is emitted for the extraction of considerable features like scale-invariant transform, histogram, and texture from the brain's magnetic resonance images. The decision tree, the K-Nearest Neighbour (KNN), and the CNN classifiers were utilized to increase the Group Grey Wolf Optimization (GGWO) approach's detection performance. These can be utilized for identification of the reduced set of useful features without performance degradation. When compared with existing literature's more competitive schemes, the proposed technique is found to have 96.23% accuracy for AD detection.

For brain MRI images' classification into two distinct classes, namely, Dementia class or Non-Dementia class, Bharanidharan & Harikumar[10]had propounded a novel Modified Grey Wolf Randomized Optimization technique. Grey Wolf Optimization (GWO) technique and other such swarm intelligence algorithms are primarily utilized for the resolution of problems related to feature selection and optimization. However, it is intelligent to utilize this technique so as to categorize brain MRI images for classification of dementia. This work has used a collection of cross-sectional MRI from 52 Dementia and 65 Non-Dementia subjects from OASIS. The original GWO's randomness makes its performance quite poor. The modified GWO will discard this randomness and also add control parameters. Despite that, the

Tianjin DaxueXuebao (ZiranKexueyuGongchengJishu Ban)/
Journal of Tianjin University Science and Technology
ISSN (Online): 0493-2137
E-Publication: Online Open Access
Vol:54 Issue:04:2021
DOI 10.17605/OSF.IO/V9TX7

problem of local optima makes this technique's accuracy quite inadequate. To this end, there was proposal of the novel modified grey wolf randomized optimization with controlled randomness at the original algorithm's appropriate position. The modified GWO variants are then added with Principal Component Analysis(PCA), Detrended fluctuation analysis, Hilbert Transform, and K-Means clustering for boosting the accuracy as well as for testing without or with statistical features. The Hilbert Transform-Modified Grey Wolf Randomized Optimization technique without statistical features has offered the highest accuracy of 93.16%, whereas, the original GWO with statistical features has an accuracy of 51.28%, and the original GWO without statistical features has an accuracy of 52.99%.

The selection of a subset containing optimal features may turn out to be inflexible and every feature selection-related issue will be displayed as the NP-hard. For an optimal feature set's election with feature selection, Valarmathy & Ramani[11]had further examined with the utilized optimization approaches. In large-scale signal processing, the PCAis quite applicable. There are possibilities for source separations well as noise estimation. This is accomplished by optimizing the Radial Basis Function (RBF) and its classifier to this structure through the GA-Artificial Immune System (AIS) algorithm's utilization. This optimized classifier of the RBF is capable of classifying a feature set which has been given by the feature selection's GA, AIS, and GA-AIS algorithms. Assessment of a classifier is dependent on its performance metrics. All classifiers are assessed by retaining the sensitivity, specificity, and accuracy in utilizing an optimized feature set. Experimental outcomes have demonstrated the feature selection as well as its effectivity in enhancing the classification accuracy of all images.

For the diagnostic procedure's time complexity reduction as well as for enhancement of its computational efficiency, Thiruvasagam & Palanisamy [12] had proposed automated brain tissue segmentation for magnetic resonance brain images. The brain tissues' pre-processing and its segmentation with Histogram-based Swarm Optimization techniques are the two processes incorporated in this approach. Investigation of this proposed approach was done with images acquired from twenty volumes and eighteen volumes of T1-Weighted images acquired from the Internet Brain Segmentation Repository (IBSR), AD images acquired from Minimum Interval Resonance Imaging in Alzheimer's disease (MIRIAD), and T2-Weighted real-time images acquired from SBC Scan Center Dindigul.

III. METHODOLOGY

Nevertheless, for further enhancement of the diagnostic accuracy and acquisition of better knowledge about disease-related brain atrophies, data-driven feature selection is essential. By removing uninformative features, the feature selection can overcome the high dimensionality issue and the small sample size issue. Discussions about the GA feature selection with NN, the KH feature selection with NN, and the GA-KH feature selection-structure optimized NN techniques are provided in this section.

Feature Selection using Genetic Algorithm (GA)-NN

Introduced by Holland, GA is a computational optimization paradigm modelled upon the biological evolution concept. This optimization approach can operate within binary search spaces and manipulate a population of potential solutions. A chromosome is a finite sequence of 1s and 0s that is a representation of a point in the search space. Evaluation of the fitness function will give potential solutions ‘quality. The survival probability is proportionate to the fitness value of the chromosome. In GA, three operators, i.e., selection, crossover, and mutation, cause the initial population's random generation. Choosing elites for direct transfer into the next generation is done by the selection operator. A portion of

Tianjin DaxueXuebao (ZiranKexueyuGongchengJishu Ban)/
Journal of Tianjin University Science and Technology
ISSN (Online): 0493-2137
E-Publication: Online Open Access
Vol:54 Issue:04:2021
DOI 10.17605/OSF.IO/V9TX7

chromosomes between the two selected parents is randomly swapped by the cross over operator to produce off spring chromosomes. A bit in chromosomes [13] is randomly alerted by the mutation.

Elimination of the insignificant features in this work is done with the GA. To this end, as a mask for features, it has defined chromosomes. That is, every chromosome will be a feature subset. The chromosome's size will be equivalent to the number of features which represent an AD's specification. A binary string that is either 1 or 0 is the chromosome's representation. 1 will indicate that the associated feature's selection, whereas 0 will indicate the associated feature's de-selection.

Feature Selection using Krill Herd (KH) Algorithm-NN

ANN efficiency can be improved through the KH algorithm's utilization since this swarm intelligence optimization algorithm is dependent on the krill herd's behaviour. In the KH algorithm, the krill population will seek krill food sources within a multidimensional search space and later will propose diverse decisions. Never the less, the target will be the distance between the krill individuals and the excess food associated with the costs. Subsequently, a krill individual's time-dependent position can be assessed by the three operational procedures: (i) motion, which was induced by the presence of other individuals, (ii) foraging motion, and (iii) random physical diffusion. The following are the KH algorithm's key benefits: all agents have a role in the procedure; it is not essential to have derivative information; both the crossover operator and the mutation operator [14] are exploited. However, the KH algorithm necessitates an optimal approach for determining the primary krill distribution as well as parameters and more comprehensive basic motions within the algorithm.

Proposed Feature Selection using GA-KH- Classification using 4 Hidden Layer NN

With a huge search space, the GA is a stochastic meta-heuristic search scheme for detecting the global solution. The standard Evolutionary Algorithms (EA) were this algorithm's inspirations. The genetic operators, encoded within a genome, are executed exceptionally for offspring creation through asexual reproduction. In sexual reproduction, the chromosomes' swapping and reordering for birthing an offspring that has the cross-breeding of genetic information from both parents. The genetic information's swapping is commonly known as a crossover operation. The solution diversity is increased through a mutation operator's utilization for the avoidance of premature convergence. For its performance [15] improvement, there is the incorporation of the genetic operators into the KH algorithm.

KH Algorithm's Crossover Operator: An efficient process for global solutions. A probability Cr will control this procedure through the generation of a random value with a uniform distribution between [0, 1]. Equation (1) and Equation (2) provide the expression of the m_{th} component of $x_{i,m}$ as below:

$$x_{i,m} = \begin{cases} x_{p,m}, if\ rand < Cr \\ x_{q,m}\ else. \end{cases} \qquad (1)$$

$$Cr = 0.2\widehat{K}_{i,best} \qquad (2)$$

Wherein, Equation (1) will determine the crossover probability, p and q will indicate two solutions that are picked for the crossover operator, p, q ∈ {1, 2, . . . ,i − 1, i +1, . . . , n}, the Cr will increase with the decreasing fitness function,$\widehat{K}_{i,best}$ = $K_i - K^{best}$,in which K_i will indicate the i_{th}KI's objective function value, and K^{best}will indicate is the i_{th}KI's best objective function value.

Tianjin DaxueXuebao (ZiranKexueyuGongchengJishu Ban)/
Journal of Tianjin University Science and Technology
ISSN (Online): 0493-2137
E-Publication: Online Open Access
Vol:54 Issue:04:2021
DOI 10.17605/OSF.IO/V9TX7

KH Algorithm's Mutation Operator: This operator is an efficient global solution strategy. A probability Mu will control this strategy. Equation (3) and Equation (4) will express the mutation operator as below:

$$x_{i,m} = \begin{cases} x_{gbest,m} + \mu(x_{p,m} - x_{q,m}), if\ rand < Mu \\ x_{i,m}\ else. \end{cases} \quad (3)$$

$$Mu = 0.05/\widehat{K}_{i,best} \quad (4)$$

Wherein, the mutation probability will be determined by Equation (4). p, q ∈ {1, 2, . . . ,i −1, i + 1, . . . , S}, Mu is a value between [0, 1] which will increase with the fitness function's decrease.

The GA-KH was employed for NN [16] training. This algorithm was demonstrated to be superior in comparison to the utilization of a simple GA and KH algorithm for the detection of a back propagations' initial values. Rather than a binary number, since a real number was utilized for each weight's encoding, the binary encodings inherent deficiency in accuracy was avoided. Instead of performing crossover on every feature, it was only done on a random number of features. In addition to that, the mutation was done one random digit that was within the real number of weight. As these were not binary features, a "reverse significance of 9"operation (for instance, 3 would mutate to 6, 4 would mutate to 5, etc.) was done by the mutation.

While GA-KH algorithms have been employed for the optimization of one-layered networks, these were too small for even resolving problems that were moderately complicated. For optimization of NNs, which had a set number of layers, numerous other GA-KH algorithms have also been utilized. However, this approach's problem involved the GA-KH's one-time execution for each of the hidden layers 'different numbers. Images were classified by applying a variable string GA-KH algorithm to determine the feed forward NN's initial weights and the hidden layer's number of neurons. There were fixed (input, hidden, and output) a number of layers; however, the GA-KH could search through differently-sized networks by adjusting the number of neurons.

Proposed Feature Selection using KH-GA - Classification using Structures Optimized NN

Selection of the weights and the structure or architecture, namely, the graph that describes the number of neurons and the manner in which the neurons are connected, form the basis of the NNs 'performance. It is essential to have a feasible architecture for fast learning or learning with small data amounts. Evolutionary (KH and GA)structure optimization of the NNs has been verified to be quite effective for selecting architecture and weights. Embedding of the approximate model's structure optimization into the design optimization algorithm [17] is, in essence, possible.

Akin to the KH algorithm [18], the GA is a stochastic search algorithm based on the population. The algorithm has maintained a solution population in which every solution will represent a candidate solution for the problem's resolution. Therefore, a hybrid of the KH algorithm and GA (KH-GA) has been proposed as a novel technique for AD. There is random initialization of a population of sizes * n in the KH-GA. While these individuals are viewed as ahead for the KH algorithm, they are viewed as genetic operators for GA. For the creation of a memory of size S * n by KHacts, krill individuals are provided to the KH memory. Afterward, there is an organization of the krill individuals depending on the fitness function. These are then provided to the GA to boost the GA-KH algorithm through a new

Tianjin DaxueXuebao (ZiranKexueyuGongchengJishu Ban)/
Journal of Tianjin University Science and Technology
ISSN (Online): 0493-2137
E-Publication: Online Open Access
Vol:54 Issue:04:2021
DOI 10.17605/OSF.IO/V9TX7

solution's generation and its addition to the population if it is better in comparison to the KH memory's worst solution. Hence, better individuals are utilized by the KH-GA in every iteration for searching the optimum solution and the solutions' enhancement. Moreover, with GA control, poorly performing krill individuals can be retained for the avoidance of premature convergence.

For the given fitness function's [19] maximization, a hybrid KH -GA was utilized in the proposed system to detect feature combinations. The feature space with each feature is represented in an individual dimension, and each dimension has a very large span that extends from 0 to 1.Therefore, an intelligent searching technique is required for the detection of an optimal point within the search space, which is able to maximize a provided fitness function (Equation (5)). For a given training data, the KH-GS's fitness function will maximize the classification performance over the validation set whilst retaining a minimum number of selected features.

$$f_\theta = \omega * E + (1-\omega)\frac{\sum_i \theta_i}{N} \tag{5}$$

Wherein, $f_\theta$will indicate the fitness function for a provided vector $\theta$with N size and having 0/1 elements which represent the unselected/selected features; N will indicate the dataset's overall number of features; E will indicate the classifier error rate, and ωwill indicate a constant that controls the significance of the classification performance to the number of selected features.

KH-GA structure optimized NN optimization's flowchart has been illustrated in Figure 1.

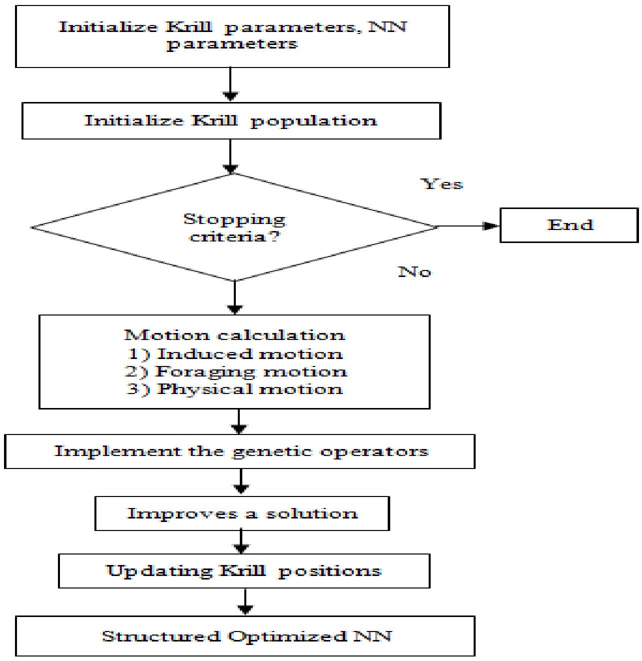

Figure 1 Flowchart for KH-GA Structure Optimized Neural Network Optimization

Tianjin DaxueXuebao (ZiranKexueyuGongchengJishu Ban)/
Journal of Tianjin University Science and Technology
ISSN (Online): 0493-2137
E-Publication: Online Open Access
Vol:54 Issue:04:2021
DOI 10.17605/OSF.IO/V9TX7

IV. RESULTS AND DISCUSSION

In this section, the GA feature selection –NN, KH feature selection – NN, feature selection using GA-KH- classification using 4 hidden layer NN and feature selection using KH-GA-classification using structure optimized NN methods are used. Table 1 shows the summary of the results. The classification accuracy, recall, and precision, as shown in figures 2 to 4.

Table 1 Summary of Results

	GA Feature Selection - NN	KH Feature Selection - NN	Feature Selection using GA-KH-Classification using 4 hidden layer NN	Feature Selection using KH-GA-Classification using structure optimized NN
Classification Accuracy	88.27	90.23	92.51	94.79
Recall for normal	0.898	0.9143	0.951	0.9633
Recall for AD	0.8226	0.8548	0.8226	0.8871
Precision for normal	0.8226	0.8548	0.9549	0.9712
Precision for AD	0.898	0.9143	0.8095	0.8594

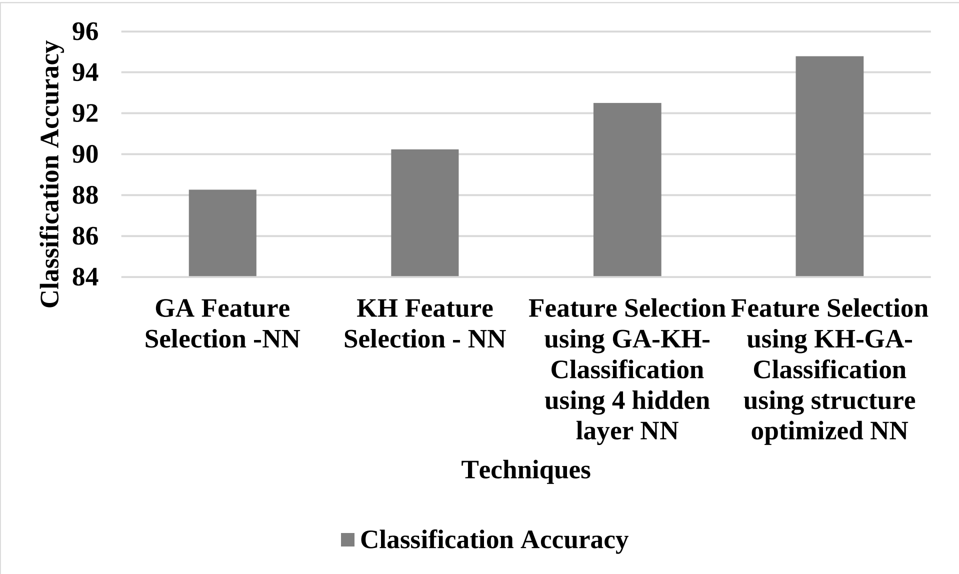

Figure 2 Classification Accuracy for Feature Selection using KH-GA- Classification using structure optimized NN

Tianjin DaxueXuebao (ZiranKexueyuGongchengJishu Ban)/
Journal of Tianjin University Science and Technology
ISSN (Online): 0493-2137
E-Publication: Online Open Access
Vol:54 Issue:04:2021
DOI 10.17605/OSF.IO/V9TX7

From figure 2, it can be observed that the feature selection using KH-GA- classification using structure optimized NN has higher classification accuracy by 7.12% for GA feature selection -NN, by 4.93% for KH feature selection - NN, and by 2.43% for feature selection using GA-KH- classification using 4 hidden layer NN respectively.

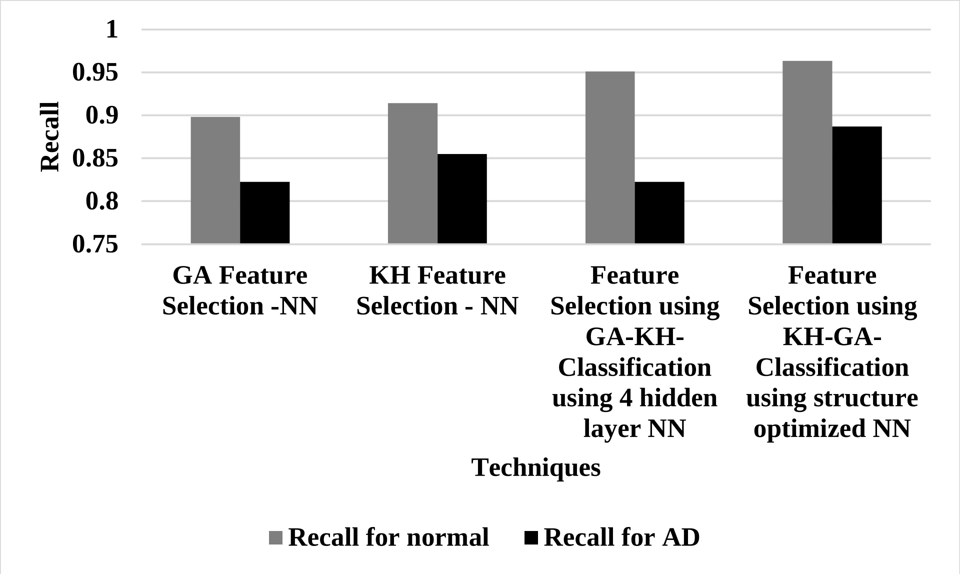

Figure 3 Recall for Feature Selection using KH-GA- Classification using structure optimized NN

From figure 3, it can be observed that the feature selection using KH-GA- classification using structure optimized NN has a higher recall for normal by 7.02% for GA feature selection -NN, by 5.23% for KH feature selection - NN, and by 1.28% for feature selection using GA-KH- classification using 4 hidden layer NN respectively. The feature selection using KH-GA- classification using structure optimized NN has a higher recall for AD by 7.54% for GA feature selection -NN, by 3.71% for KH feature selection - NN and by 7.54% for feature selection using GA-KH- classification using 4 hidden layer NN respectively.

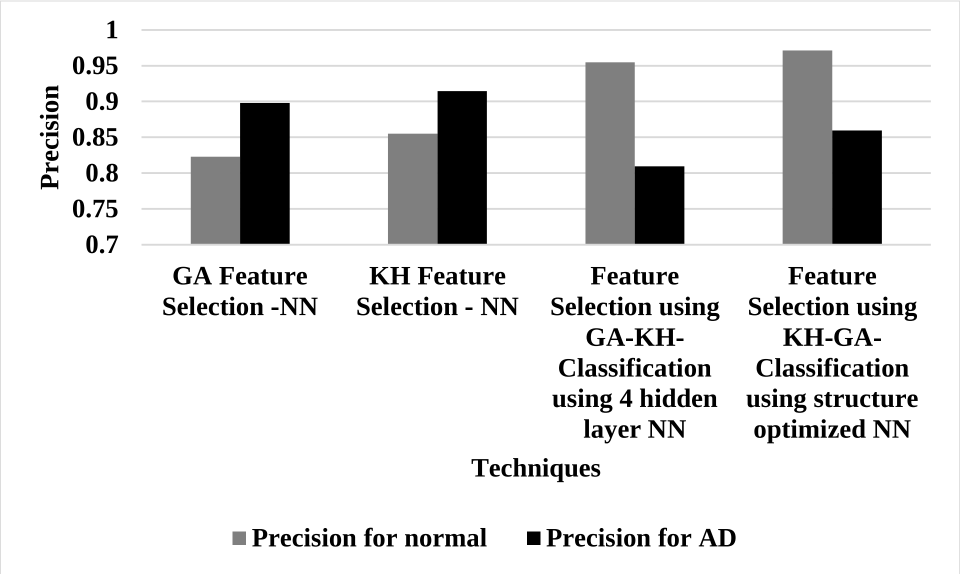

Figure 4 Precision for Feature Selection using KH-GA- Classification using structure optimized NN

Tianjin DaxueXuebao (ZiranKexueyuGongchengJishu Ban)/
Journal of Tianjin University Science and Technology
ISSN (Online): 0493-2137
E-Publication: Online Open Access
Vol:54 Issue:04:2021
DOI 10.17605/OSF.IO/V9TX7

From figure 4, it can be observed that the feature selection using KH-GA- classification using structure optimized NN has higher precision for normal by 16.56% for GA feature selection - NN, by 12.75% for KH feature selection - NN, and by 1.69% for feature selection using GA-KH- classification using 4 hidden layer NN respectively. The feature selection using KH-GA-classification using structure optimized NN has higher precision for AD by 4.39% for GA feature selection -NN, by 6.19% for KH feature selection - NN,and by 5.98% for feature selection using GA-KH- classification using 4 hidden layer NN respectively.

V. CONCLUSION

For the progression of treatments, it is necessary to have an early AD diagnosis. A framework is capable of gaining from extensive and complicated datasets through the utilization of a diverse range of probabilistic and optimization techniques from an artificial intelligence branch known as machine learning. Due to this, for the early stages of AD, researchers usually concentrate on machine learning. The KH algorithm's concepts can be described within the context of sensing the food and producing a high-density KI herd. Crossover and mutation are the two genetic operators applied by the KH algorithm. In this algorithm, physical random diffusion, as well as genetic operators, constitutes the exploration steps. Meanwhile, foraging motion as well as the movements induced by other krill individuals, constitute the algorithm's exploitation steps. The NNs' structural optimization has been done in this work prior to their utilization in evolutionary design optimizations 'fitness evaluations. The purpose of this work is to propose hybrid KH-GA for feature selection to pick a minimal number of features (attributes) and attain comparable or even better classification accuracy from the utilization of all these attributes. It is demonstrated in this work that, for feature selection problems, KH-GA is an efficient search algorithm. While the applied fitness function's primary target is classification accuracy, its secondary target is the reduction size. As a result, selected features that have maximum accuracy and the minimum size can be obtained. It is evident from the outcomes that the feature selection which utilized KH-GA-classification with structure optimized NN had higher classification accuracy by 7.12% for GA feature selection-NN, by 4.93% for KH feature selection–NN, and by 2.43% for feature selection that utilized GA-KH-classification with 4 hidden layer NN.

REFERENCES

1. Mahajan, S., Bangar, G., &Kulkarni, N. (2020). Machine Learning Algorithms for Classification of Various Stages of Alzheimer's Disease: A review. Machine Learning, 7(08).
2. Singh, L., &Chetty, G. (2012, December). A comparative study of MRI data using various machine learning and pattern recognition algorithms to detect brain abnormalities. In Proceedings of the Tenth Australasian Data Mining Conference-Volume 134 (pp. 157-165). Australian Computer Society, Inc.
3. Vasantha, M., Bharathi, V. S., &Dhamodharan, R. (2010). Medical image feature, extraction, selection and classification. International Journal of Engineering Science and Technology, 2(6), 2071-2076.
4. Khan, A., &Usman, M. (2015, November). Early diagnosis of Alzheimer's disease using machine learning techniques: A review paper. In 2015 7th International Joint Conference on Knowledge Discovery, Knowledge Engineering and Knowledge Management (IC3K) (Vol. 1, pp. 380-387). IEEE.
5. Tanveer, M., Richhariya, B., Khan, R. U., Rashid, A. H., Khanna, P., Prasad, M., & Lin, C. T. (2020). Machine learning techniques for the diagnosis of Alzheimer's disease: A review. ACM Transactions on Multimedia Computing, Communications, and Applications (TOMM), 16(1s), 1-35.

Tianjin DaxueXuebao (ZiranKexueyuGongchengJishu Ban)/
Journal of Tianjin University Science and Technology
ISSN (Online): 0493-2137
E-Publication: Online Open Access
Vol:54 Issue:04:2021
DOI 10.17605/OSF.IO/V9TX7

6. Sountharrajan, S., &Thangaraj, P. (2016). Optimized feature selection technique for automatic classification of MRI images for Alzheimer's disease. Journal of Medical Imaging and Health Informatics, 6(8), 2057-2062.
7. Neffati, S., Ben Abdellafou, K., Jaffel, I., Taouali, O., &Bouzrara, K. (2019). An improved machine learning technique based on downsized KPCA for Alzheimer's disease classification. International Journal of Imaging Systems and Technology, 29(2), 121-131.
8. Hossain, E., Hasan, M., Hassan, S. Z., Azmi, T. H., Rahman, M. A., &Parvez, M. Z. (2020, November). Deep learning based binary classification for Alzheimer's disease detection using brain MRI images. In 15th IEEE Conference on Industrial Electronics and Applications: ICIEA 2020. IEEE Xplore.
9. Shankar, K., Lakshmanaprabu, S. K., Khanna, A., Tanwar, S., Rodrigues, J. J., & Roy, N. R. (2019). Alzheimer detection using Group Grey Wolf Optimization based features with convolutional classifier. Computers & Electrical Engineering, 77, 230-243.
10. Bharanidharan, N., &Harikumar, R. (2020). Modified Grey Wolf Randomized Optimization in Dementia Classification Using MRI Images. IETE Journal of Research, 1-10.
11. Valarmathy, S., &Ramani, R. (2019). Evaluating the Efficiency of Radial Basis Function Classifier with Different Feature Selection for Identifying Dementia. Journal of Computational and Theoretical Nanoscience, 16(2), 627-632.
12. Thiruvasagam, P., &Palanisamy, K. (2020). Brain Tissue Segmentation from Magnetic Resonance Brain Images Using Histogram Based Swarm Optimization Techniques. Current Medical Imaging, 16(6), 752-765.
13. Aalaei, S., Shahraki, H., Rowhanimanesh, A., &Eslami, S. (2016). Feature selection using genetic algorithm for breast cancer diagnosis: experiment on three different datasets. Iranian journal of basic medical sciences, 19(5), 476.
14. Asteris, P. G., Nozhati, S., Nikoo, M., Cavaleri, L., &Nikoo, M. (2019). Krill herd algorithm-based neural network in structural seismic reliability evaluation. Mechanics of Advanced Materials and Structures, 26(13), 1146-1153.
15. Abualigah, L. M. Q. (2019). Feature selection and enhanced krill herd algorithm for text document clustering (pp. 1-165). Berlin: Springer.
16. Nowakowski, G., Dorogyy, Y., &Doroga-Ivaniuk, O. (2018). Neural network structure optimization algorithm. Journal of Automation Mobile Robotics and Intelligent Systems, 12.
17. Hüsken, M., Jin, Y., &Sendhoff, B. (2005). Structure optimization of neural networks for evolutionary design optimization. Soft Computing, 9(1), 21-28.
18. Abualigah, L. M., Khader, A. T., &Hanandeh, E. S. (2018). A hybrid strategy for krill herd algorithm with harmony search algorithm to improve the data clustering**1**. Intelligent Decision Technologies, 12(1), 3-14.
19. Hafez, A. I., Hassanien, A. E., Zawbaa, H. M., &Emary, E. (2015, December). Hybrid monkey algorithm with krill herd algorithm optimization for feature selection. In 2015 11th International Computer Engineering Conference (ICENCO) (pp. 273-277). IEEE.