10.6977/IJoSI.202209_7(3).0003
G. A. Ansari, Sivakani. R, S. Srisakthi / Int. J. Systematic Innovation, 7(3), 28-38 (2022)

Precise diagnosis of alzheimer's disease using recursive feature elimination method

Gufran Ahmad Ansari[1], R. Sivakani[2*], S. Srisakthi[2]
[1] Department of Computer Application, B.S. Abdur Rahman Crescent Institute of Science and Technology, Vandalur, Chennai, Tamilnadu, India

[2] Department of Computer Science and Engineering, B.S. Abdur Rahman Crescent Institute of Science and Technology, Vandalur, Chennai, Tamilnadu, India

[*] Corresponding author E-mail: sivakani13@gmail.com
(Received 10 December 2021; Final version received 15 June 2022; Accepted 5 July 2022)

Abstract

One of the prevalent diseases that the elderly tend to have patients has been Alzheimer's disease (AD). It is a neurological disease where the brain cells start to deteriorate. As the disease progresses it eventually leads to the death of the brain cells. Death in brain cells results in various problems like memory loss change in behavior patterns and many more. The most challenging problem has been in predicting an early diagnosis of AD in patients. The importance of the disease is that it is detected early. If early detection is done, the death of brain cells can be reduced. The disease is predicted based on the various features of the patient. Feature selection has been one of the important steps in predicting the disease. This paper takes the OASIS data set and implements the different algorithms and proposes a model. The proposed model identifies the salient feature by recursively considering smaller and smaller sets of the features. The classification has been done for evaluating the feature selection. The result has been compared before the feature selection method and after the feature selection method. The performance metrics show improved scores after applying the feature section concept.

Keywords: Alzheimer, early diagnosis, feature selection, recursive feature elimination (RFE).

1. Introduction

Alzheimer's has been a neurodegenerative disease that mostly occurs in elderly people. Neurodegenerative, is a continuous worsening of the neurons, this affects the competence of the central nervous in a very intense and progressive manner (Scatena et al, 2007). AD is one of the prevalent neurodegenerative disorders (Small and D. H., 2005) and it is a cureless disease and the only treatment is to slow down its progression (Unay et al, 2010). The disease mainly affects the aged 60 to 65. A person at this age is at a high risk of being vulnerable to AD (Cummings et al, 2014). The disease grows progressively from a Cognitively Normal (CN) person to AD through Mild Cognitive Impairment (MCI) (Matsuda H, 2007), (Mosconi L et al. 2008). An Early Diagnosis (ED) of the disease would be helpful so that its progression can be reduced.

Feature selection is an important task in machine learning. If the dataset is processed without feature selection means then the accuracy of the prediction will be reduced, and the processing time will be increased so to avoid these issues the feature selection will be the best.

The diagnosis of the disease is done on studying the brain images or the Magnetic Resonance Imaging (MRI) of the brain. Along with MRI, some doctors use the s-MRI and resting-state functional magnetic resonance imaging (rs-fMRI) also. Both these have been used as the common method to analyze the changes, activities in the brain (Jun Jie Ng et al, 2016). Thus, the MRI serves as input using which AD is being diagnosed. The MRI of a patient has many features like gray matter densities, cortical thickness. The features related to diagnose and AD, needs to be retrieved from the MRI. When the correct features are selected, it helps in the correct diagnosis. The feature selection greatly affects or influences the classification performance and hence acts as the first step towards the prediction.

To have higher prediction accuracy the feature selection process needs to be done. The importance of feature selection process can be stated as to

- It decreases the training time of the machine learning algorithm.
- It increases the model accuracy as it uses the right subset
- It minimizes over fitting

Hence, the feature selection step becomes crucial. To perform the feature selection process a number of algorithms and techniques have been used. Feature selection is the process of selecting a subset from the original input data; also use these features as input to next step. The feature selection models have been classified to be of three methods:

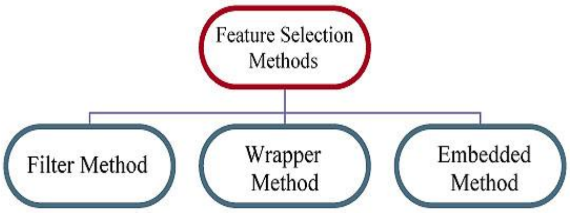

Fig. 1. Feature selection methods.

- Filter based
- Wrapper based
- Embedded

In the filter-based methods some metrics are specified as filters and based on those filters the features are selected. Some examples for this type of feature selection are Chi-squared test, Information gain. In the wrapper-based methods the feature selection is taken as a problem where a search technique is deployed. To perform the searching process algorithms like a best-first search are utilized. An example of these algorithms that use the wrapper method is the recursive feature elimination (RFE). Embedded methods work on the logic of learning about the features that contribute to the accuracy of the model that is being created. An example of an algorithm that uses the embedded methods is the regularization method.

Each of the above methods has its own advantage and disadvantages. Selecting and using the feature selection method depends on the problem and it also depends upon the existing data that the model has. ML(Machine Learning) algorithms in particular have been used in all the stages of AD prediction. Researchers have used these methods in the feature selection process to diagnose AD.

In this paper, the importance of the feature selection for an optimal prediction has been studied and analyzed. In the medical field the prediction of the disease is very important, which is done based on the symptoms. By constructing a machine learning model to predict the disease is done based on the features presented. Thus, the features play a vital role in the prediction of the disease. Appropriate selection of the features needs to be done for obtaining an optimal result. The detection of the Alzheimer disease is done based on the features available. For the study as carried out in this paper, OASIS dataset has been considered for the AD prediction. The different features of the disease are listed out and the feature selection is being done. Identifying the best feature has been one of the major tasks in predicting the disease. From the literature it shows that many of the previous works does not involve feature selection. Using the machine learning techniques, the feature selection is being done and the best feature has been identified and implemented.

The feature selection is necessary for selecting the required attributes for the prediction. It will remove the unwanted and the less priority attributes from the dataset using various feature selection techniques. It will use the attributes and the features effectively and efficiently for good prediction. If the attributes are not selected properly, the prediction of the disease will not be accurate. The feature selection will train the machine learning model very quickly also the over-fitting problem is reduced.

Outlines of the contributions are:

- Selecting the feature selection method.
- Constructing the feature selection model using machine learning.
- Select the attributes using the various feature selection methods for the prediction.
- Train the model and test the model.
- Evaluate the results.
- Comparing the results to other models.

- Analyzing the performance of the proposed model.

The proposed model in the paper uses the wrapper method. The paper is structured as Section 2 explains the related work, Section 3 describes the dataset, the design and working of the proposed model, the analysis is presented in section 4 and Section 5 explains the discussion with the future work and section 6 describes the conclusion.

2. Related works

Feature selection has been a vital step in building a machine learning model. To improve the accuracy of the predictive model, it is required to reduce the count of input variables or the features. This also reduces the computational cost of building a model. Researchers working on the prediction of AD, have used feature selection on two inputs, That are used in building the prediction model. Feature selection has been used on MRI and on the dataset. The proposed model presented in this paper takes the dataset and performs the feature selection on it.

There are many data sets that are available online, where all the details pertaining to an AD are stored. Some of the most used data sets are:

- The Alzheimer's Disease Neuroimaging Initiative (ADNI).

- Open Access Series of Imaging Studies (OASIS)

The ADNI data set is a longitudinal multicenter study that was created to develop clinical, imaging, genetic, and biochemical biomarkers. These were used for the early identification and diagnosis of Alzheimer's disease (AD). The data set includes participants who were recruited across North America during the stage of the study, and these participants had decided to complete diversity of imaging and medical assessments. Once the participants registered with ADNI, they were followed and reassessed over time. This is being done so as to track the pathology of the disease in the course of its progress. OASIS data set aims to provide neuroimaging datasets to the scientific and the research community. This helps in future discoveries. Many researchers have taken these two data sets and used them for building feature selection models that can be used in predicting AD. In detecting AD, the uses of computer technology have been used in vast numbers.

(Hinrichs et al, 2011) have made the ADNI data set, they used 48 AD patients and 66 Normal Controls (NC) for the diagnosis of AD, and they got an accuracy of 87.60% with the help of two image inputs of Positron emission tomography (PET) and MRI. The authors also achieved the result of 92.40% using the modalities of PET, MRI, Cerebro Spinal Fluid CSF, APOE, and cognitive values. In the model proposed by (Gray et al, 2013), they had used 37 AD patients, 75 MCI patients, and 35 NC for classify the AD and MCI patients. They used four modalities of PET, CSF, MRI, and genetics. With their model, they achieved an accuracy of 89.00% for AD classification and had an accuracy of 74.60% for MCI classification.

(Zhang et al, 2011) have proposed a model that used the same ADNI dataset and got an accuracy of 90.60% for the classification of AD, they have used the MRI and PET. They were able to achieve an accuracy of 93.20% for the classification of AD by using three inputs the MRI, the PET, and the CSF. (Feng Liu et al, 2014) have proposed a model that gave a result of 94.37% and the ROC curve (AUC) of 0.9724 in detecting AD. They also achieved an accuracy of 78.80% and an AUC of 0.8284 in identifying.

(Trambaiolli et al, 2017) have developed a model that used the Filtered Subset Evaluator technique and were able to achieve the best performance improvement for the patient is 91.18% of accuracy and on an epoch basis is 85.29±21.62%. They first removed 8.76±1.12% of the original features. Riedel et al., have proposed a model where AD features were selected from the different neuroimaging modalities. These were used to create more useful measures, and these features included mean gray matter densities, subcortical, cortical thickness, and cerebral amyloid-b accumulation in regions of interest (ROIs). (Lecun et al, 2015) have proposed a model and used the Deep learning (DL) method of convolutional neural networks (CNN), to build models. The authors have proved that this method of using DL has been shown to outperform the other existing machine learning methods. Thus, many feature selection models have been developed using different approaches. The model proposed in this paper uses the Recursive Feature Elimination method. (Sivakani et al, 2020) have generated the missing values using the algorithms EM, KNN, and RF algorithms. (Sivakani et al, 2020) have done the feature selection using the best-first search algorithm and cfssubsetevalu evaluator.

(Wiharto et al, 2022) have done a study for the diagnosis of heart disease using a feature selection method developed with the genetic algorithm and support vector machine. Out of 54 features, only 5 features are selected and produced 87% accuracy. (Z. D. Akşehir et al, 2022) have introduced a new rule-based labeling algorithm and a feature selection method for the prediction of CNN model performance. (Chen et al, 2020) have specified four reasons for showing the importance of feature selection. The feature selection reduces the parameters, decreases the model building time enhances the generalization, and reduces the dimensionality. The evaluation has been done using the random forest, support vector machines, K-Nearest neighbors, and Linear Discriminant Analysis. (Jianting Chen et al, 2021) have proposed a novel self-learning feature selection method using the wrapper method the improvement the accuracy of the machine learning model. The evaluation has been done using sixteen UCI repository datasets. (Lianxi Wang et al, 2021) explained the uses of the feature selection algorithm to increase the accuracy of the classification. The evaluation has been done using the UCI datasets. (Chu Y.M et al, 2020) studied hybrid nanoparticles for various mixtures and their applications. (F. Heydarpour et al, 2020) introduced a system of ordinary differential equations (ODE) to predict tumor growth. An artificial neural network has been applied to solve the problem in the ODE. (He Z. Y et al, 2022) presented a new fractional-order discrete-time susceptible-infected-recovered (SIR) epidemic model with vaccination to find the system's dynamics using the numeric value. The complexity of the system has been analyzed and verified. (Rahiminasab A et. al, 2020) introduced a model for choosing a cluster head for the energy prediction, using the attributes energy, mobility, distance and the length of data queues. (Fang Jin et. al, 2022) proposed a new system to prove the uniqueness of the result using the fixed point theory and the Picard technique.

The literature clearly shows the importance of feature selection in medical and other fields. So a model has been proposed to predict the disease with the feature selection. In this paper, the disease prediction has been made with various feature selection techniques and without applying feature selection techniques; then, the result has been compared and proved that the effect produced with the feature selection technique is the best. The proposed model has been compared with the model (Trambaiolli, L. R., et al, 2017) and shows the better accuracy.

3. Proposed model

AD has been one of the challenging diseases that have been very difficult to diagnose at an early stage. When a patient is being tested for the symptoms of the disease, a number of features are taken into account. Evaluating these features and diagnosing the disease based on these features has been a critical method. Using ML, a number of approaches are available to do the feature selection. The model proposed here uses the wrapper approach which is found to have certain advantages when compared to the other approaches. This section first presents the dataset used in the paper and then discusses the design of the proposed model. The dataset has been preprocessed and is subjected to feature selection. Various feature selection methods have been used and the result is compared to predict the best feature selection method. The RFE method will remove the least important attribute for the prediction. The features for the processing will be selected and a subset has been generated using the machine learning techniques. Based on the less importance of the attributes, it will eliminate and with the other attributes, the processing will be done for the best prediction.

3.1 Dataset

The dataset taken for the proposed model is the OASIS dataset. The recent dataset of OASIS has been the OASIS-3 which has displayed data for above than 1000 participants. All these participants were across several ongoing projects for 30 years. The partakers included 609 CN adults and 489 patients at several phases of cognitive decline, all the participants were aged from 42-95years. This dataset has the details such as patient ID, Gender, Dominant Hand, Age, Education detail, etc. These features are taken and the processing is done.

3.2 Design of the model

The feature that is used to train any machine learning model has a great impact on the performance of the model. A feature that is irrelevant or partially relevant can have a negative impact. The proposed model makes use of the RFE method that follows the wrapper approach. RFE is based on the scheme to repeatedly construct a model and then the best or worst performing feature is chosen. The chosen feature is set aside and the process is repeated again taking the rest of the features. This process is functionally applied until all the features in the dataset are taken out. All the features are then ranked according to when they were eliminated. This technique can be called a greedy optimization problem that can be used for finding the best performing subset of features. The process of RFE is given in Fig. 2.

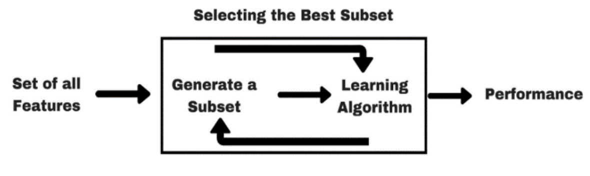

Fig. 2. Recursive feature elimination

The various features present in the dataset are tabulated in the Table 1.

Table 1. Data description

S.No	Attribute_ Fields	Description
1	Subject ID	Subject Identification number
2	MRI ID	MRI Identification number
3	Group	Group of the patient
4	Visit	Visit of the patient
5	MR Delay	MR delay of the visit
6	M/F	Gender
7	Hand	Mental State test
8	Age	Patient's age
9	EDUC	Education status
10	SES	Social economic Status
11	MMSE	Mini Mental State Examination
12	CDR	Clinical Dementia Rating
13	eTIV	Estimated Total Intracranial Volume
14	nWBV	Normalize Whole Brain Volume
15	ASF	Atlas Scaling Factor

4. Result and discussion

The data set has been used in various algorithms and the result has been evaluated. The OASIS dataset has been used for the evaluation of the feature selection algorithms. The result has been analyzed in two-step processes; first, the dataset has been preprocessed and the classification has been done without applying feature selection methods and in the second step the classification of the dataset has been done with the feature selection method. Finally, the result has been compared for finding the best feature selection method. In this regard, the CFS subset attributes selection, classifier attribute selection, correlation attribute, and relief attribute selection methods have been used for the feature selection and for the evaluation these methods Naïve Bayes, Logistic Regression, SVM, Bagging, Logitboost, Multiclass, JRip, J48, Random Forest, REP Tree classifiers has been used to find the best feature selection method.

4.1 CFS subset attributes selection

CFS is the Correlation-based feature selection; this algorithm selects the attributes which have a high correlation with the classification task. The equation used in the subset evaluation is given below:

$$r_{zc} = \frac{k\overline{r_{zi}}}{\sqrt{k + k(k-1)\overline{r_{ii}}}}, \tag{1}$$

Where, r_{zc} is the correlation of the attributes, k is the number of attributes, $\overline{r_{zi}}$ is the average of the correlations $\overline{r_{ii}}$ is the average inter-correlation among the attributes.

We have applied CfsSubset algorithm to evaluate the subsets. The search methods applied for finding the best attributes are best first search method, greedy search method. The direction of the attribute search taken place in forward direction. Total we have 373 instances and 15 attributes.

When we are applying the CfsSubset algorithm along with the best first search method the evaluated subsets are 104 and the accuracy of the best attribute search is 0.98 and the selected attributes are 1,4,12,13,14. Totally the selected attributes are Subject ID, Visit, CDR, eTIV, nWBV

When we are applying the CfsSubset algorithm along with the greedy search method the evaluated subsets are 104 and the accuracy of the best attribute search is 0.98 and the selected attributes are 1,4,12,13,14. Totally the selected attributes are Subject ID, Visit, CDR, eTIV, nWBV. Both the methods selected the same attributes as the best.

4.2 Classifier attribute selection

The classifier attributes algorithm works based on the prediction concept for selecting the best subset. The search method applied with the evaluator is the ranking method. Rank will be generated for each attribute; the rank will be generated between -1 and +1. The ranking method is applied along with this algorithm to choose the best attribute. For each attribute, the rank is generated, and based on the rank the best attribute is selected. The wrapper and RMSE methods are used for subset evaluation. The accuracy is 5 and the Selected attributes are 14,4,5,3,13,2,6,7,8,9,12,11,10,1; total 14 out of the 15 attributes. The rank generated for each attribute is given below:

Table 2. Rank generated by the classifier attribute selection algorithm

Sl. No.	Rank	Attribute number	Attribute Name
1	0	14	nWBV
2	0	4	Visit
3	0	5	MR Delay
4	0	3	Group
5	0	13	eTIV
6	0	2	MRI ID
7	0	6	M/F
8	0	7	Hand
9	0	8	Age
10	0	9	EDUC
11	0	12	CDR
12	0	11	MMSE
13	0	10	SES
14	0	1	Subject ID

4.3 Correlation attributes selection

The correlation attribute is a prediction of the best attribute based on the linear relationship between the attributes. Along with the correlation attribute, the ranking search method is applied to find the best attributes based on the rank generated by this algorithm. The selected attributes are 6,10,14,1,2,11,3,7,12,8,4,5,9,13 and totally 14 attributes are selected. The accuracy is 0.138. The rank generated for each attribute is given below:

Table 3. Rank generated by the correlation attributes selection algorithm

Sl. No.	Rank	Attribute number	Attribute Name
1	0.5616	6	M/F
2	0.2467	10	SES
3	0.2135	14	nWBV
4	0.0663	1	Subject ID
5	0.041	2	MRI ID
6	0.0395	11	MMSE
7	0.0239	3	Group
8	0	7	Hand
9	-0.0293	12	CDR
10	-0.0351	8	Age
11	-0.1204	4	Visit
12	-0.1235	5	MR Delay
13	-0.2418	9	EDUC
14	-0.9889	13	eTIV

4.4 Relief attributes selection

The relief attribute algorithm predicts the best attributes based on the weight allocated for each attribute. Along with the Relief algorithm the ranking search method is applied to select the best attributes. The selected attributes are 1,13,9,10,14,8,11,7,2,6,3,12,5,4; totally 14 attributes are selected as the best attributes. The accuracy is 0.138. The ranks generated for each attribute are given below:

Table 4. Rank generated by the relief attribute selection algorithm

Sl. No.	Rank	Attribute number	Attribute Name
1	0.15901	1	Subject ID
2	0.10424	13	eTIV
3	0.01446	9	EDUC
4	0.01144	10	SES
5	0.00874	14	nWBV
6	0.00453	8	Age
7	0.00392	11	MMSE
8	0	7	Hand
9	0.0000	2	MRI ID
10	-0.00106	6	M/F
11	-0.00354	3	Group
12	-0.0052	12	CDR
13	-0.0510	5	MR Delay
14	-0.0549	4	Visit

Representation of number of attributes selected is given in the below graph.

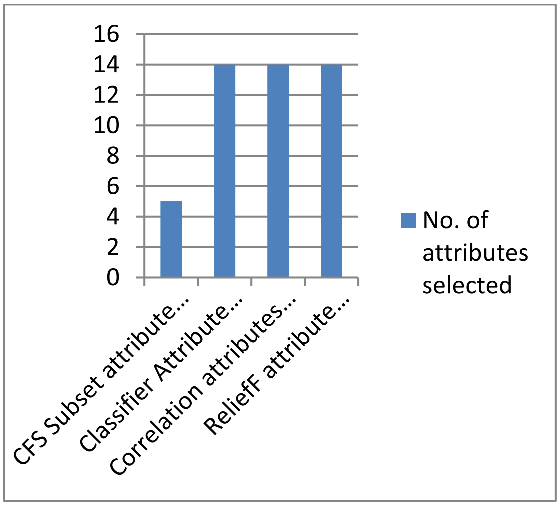

Fig. 3 Representation of the attributes selected

This Fig. 3 shows the representation of the attributes selected by the algorithms. From the above graph we can identify that the CFS subset attribute selection algorithm has been selected 5 attributes and the other algorithms has selected 14 attributes. So, for a good classification or a prediction we need these 14 attributes from this dataset.

4.5. Classification and comparison of the classifier before and after applying the feature selection concept

The classification has been done for evaluating the performance of the classifiers and comparison has been done before applying the feature selection and after applying the feature selection. The classifiers considered for the evaluation are Naïve Bayes, Logistic Regression, SVM, Bagging, Logitboost, Multiclass, JRip, J48, Random Forest, REP Tree classifiers.

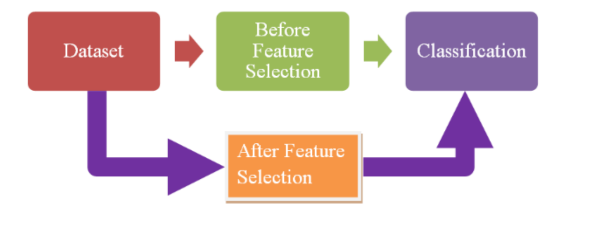

Fig. 4. Representation of the attributes selected

Fig. 4, describes that the dataset has been classified before and after feature selection. Before applying feature selection the score is less and after the feature selection the result has been improved.

Table 5. Performance of Naïve Bayes Classifier

Sl. No.	Parameter Metrics	Result before feature selection	Result after Feature selection
1	Correctly Classified	92.7	93.29
2	Precision	0.923	0.939
3	Recall	0.928	0.933
4	F-Measure	0.924	0.930
5	ROC	0.981	0.991

Table 5, shows the metrics comparison for the Naïve Bayes classifier before and after the feature selection.

Table 6. Performance of logistic regression classifier

Sl. No.	Parameter Metrics	Result before feature selection	Result after Feature selection
1	Correctly Classified	98.3	98.41
2	Precision	0.984	0.995
3	Recall	0.984	0.994
4	F-Measure	0.984	0.994
5	ROC	0.955	0.995

Table 6, shows the metrics comparison for the Logistic Regression classifier before and after the feature selection.

Table 7. Performance of SVM classifier

Sl. No.	Parameter Metrics	Result before feature selection	Result after Feature selection
1	Correctly Classified	98.6	98.75
2	Precision	0.987	0.987
3	Recall	0.987	0.997
4	F-Measure	0.986	0.986
5	ROC	0.982	0.991

Table 7, shows the metrics comparison for the SVM classifier before and after the feature selection.

Table 8. Performance of bagging classifier

Sl. No.	Parameter Metrics	Result before feature selection	Result after Feature selection
1	Correctly Classified	97.58	97.68
2	Precision	0.968	0.977
3	Recall	0.976	0.976
4	F-Measure	0.976	0.98
5	ROC	0.989	0.999

Table 8, shows the metrics comparison for the Bagging classifier before and after the feature selection.

Table 9. Performance of logitboost classifier

Sl. No.	Parameter Metrics	Result before feature selection	Result after Feature selection
1	Correctly Classified	91.68	91.78
2	Precision	0.916	0.917
3	Recall	0.917	0.927
4	F-Measure	0.896	0.896
5	ROC	0.943	0.953

Table 9, shows the metrics comparison for the Logitboost classifier before and after the feature selection.

Table 10. Performance of multiclass classifier

Sl. No.	Parameter Metrics	Result before feature selection	Result after Feature selection
1	Correctly Classified	98.39	98.65
2	Precision	0.984	0.987
3	Recall	0.984	0.997
4	F-Measure	0.984	0.987
5	ROC	0.989	0.999

Table 10, shows the metrics comparison for the Multi classifier before and after the feature selection.

Table 11. Performance of JRip classifier

Sl. No.	Parameter Metrics	Result before feature selection	Result after Feature selection
1	Correctly Classified	91.15	93.29
2	Precision	0.901	0.928
3	Recall	0.912	0.933
4	F-Measure	0.904	0.926
5	ROC	0.921	0.943

Table 11, shows the metrics comparison for the JRip classifier before and after the feature selection.

Table 12. Performance of J48 classifier

Sl. No.	Parameter Metrics	Result before feature selection	Result after Feature selection
1	Correctly Classified	87.39	87.45
2	Precision	0.805	0.815
3	Recall	0.874	0.894
4	F-Measure	0.838	0.868
5	ROC	0.868	0.878

Table 12, shows the metrics comparison for the J48 classifier before and after the feature selection.

Table 13. Performance of random forest classifier

Sl. No.	Parameter Metrics	Result before feature selection	Result after Feature selection
1	Correctly Classified	95.97	96.24
2	Precision	0.962	0.974
3	Recall	0.960	0.962
4	F-Measure	0.955	0.958
5	ROC	1.00	1.00

Table 13, shows the metrics comparison for the Random Forest classifier before and after the feature selection.

Table 14. Performance of REP classifier

Sl. No.	Parameter Metrics	Result before feature selection	Result after Feature selection
1	Correctly Classified	97.85	97.95
2	Precision	0.979	0.989
3	Recall	0.979	0.98
4	F-Measure	0.978	0.988
5	ROC	0.989	0.999

Table 14, shows the metrics comparison for the REP classifier before and after the feature selection.

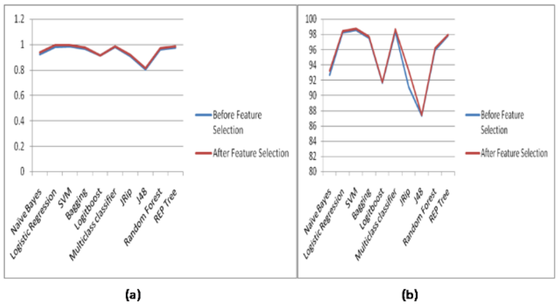

Fig. 4(a) & 4(b). Metrics comparison for precision and accuracy of classifiers

Figure 4(a) shows the comparison of precision values before and after feature selection for the classifiers. 4(b) shows the comparison of accuracy values before and after feature selection for the classifiers.

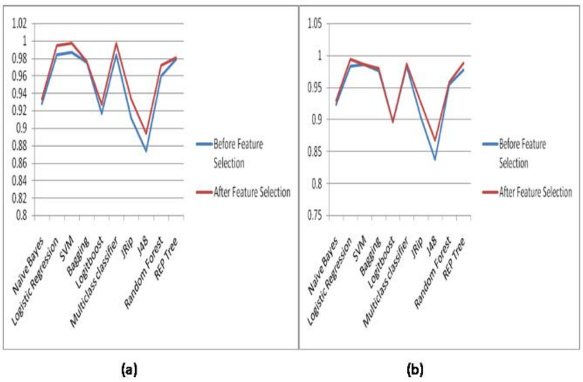

Fig. 5(a) & 5(b). Metrics comparison for recall and f-measure of classifiers

Fig. 5(a) shows the comparison of Recall values before and after feature selection for the classifiers.

5(b) shows the comparison of F-Measure values before and after feature selection for the classifiers.

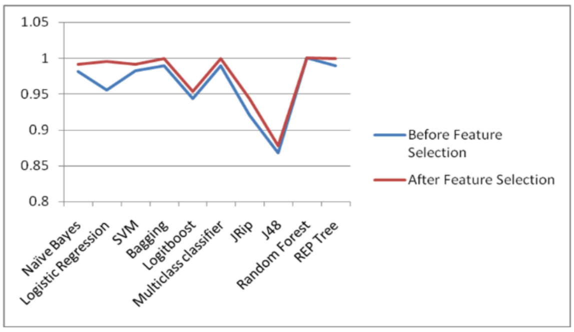

Fig. 6. Metrics comparison for ROC of classifiers

Fig.6, shows the comparison of ROC values before and after feature selection for the classifiers.

From the comparison, it has been analyzed that the performance for all the classifiers given improved scores after applying the feature selection.

5. Conclusion and Future Enhancement

This paper discusses various feature selection algorithms for selecting the best attributes for the classification or prediction. The best attributes can be selected by using the feature selection method. CFS Subset attributes selection, Classifier Attribute Selection, Correlation attributes selection, and Relief attribute selection are the algorithms applied to the oasis dataset; in the dataset, there are 15 attributes, and among those attributes, the best attributes selected are 14 attributes. CFS Subset attributes selection algorithm selected only 5 attributes, and the other algorithms selected 14 attributes so it concluded that for the classification of this dataset 14 attributes are best. Also, the classification has been done to evaluate the performance of the feature selection process. The evaluation metrics considered are accuracy, Recall, F-Measure, and ROC. All the performance metrics show better results after applying the feature selection. The results have been compared without feature selection and with the feature selection process; all the evaluation metrics show the better result with the feature selection process. Also, the proposed model has been compared with the result of the previous model and shows a better result. The previous model has 91.18%, and this proposed model gives the best result of 98.7%. So it is concluded for the best prediction, the feature selection should be done. In future work, we decided to focus on other feature selection algorithms.

References

Chen, RC., Dewi, C., Huang, SW. & et al. (2020). Selecting critical features for data classification based on machine learning methods. J Big Data. https://doi.org/10.1186/s40537-020-00327-4

Chu Y.M., Nazir U., Sohail M., Selim M.M., & Lee J.R. (2020). Enhancement in Thermal Energy and Solute Particles Using Hybrid Nano particles by Engaging Activation Energy and Chemical Reaction over a Parabolic Surface via Finite Element Approach. Fractal Fract. No. 5, 119. https://doi.org/10.3390/fractalfract5030119.

Cummings, J. L., Morstorf, T., & Zhong, K. (2014). Alzheimer's disease drug-development pipeline: few candidates, frequent failures, Alzheimer's research & therapy. 6(4):37, https://doi.org/10.1186/alzrt269.

Jianting Chen, Shuhan Yuan, Dongdong Lv, & Yang Xiang. (2021). A novel self-learning feature selection approach based on feature attributions, Expert Systems with Applications, https://doi.org/10.1016/j.eswa.2021.115219.

Lianxi Wang, Shengyi Jiang, & Siyu Jiang. (2021). A feature selection method via analysis of relevance, redundancy, and interaction. Expert Systems with Applications. https://doi.org/10.1016/j.eswa.2021.115365.

F. Heydarpour, E. Abbasi, M. J. Ebadi, & S. M. Karbassi. (2020). Solving an Optimal Control Problem of Cancer Treatment by Artificial Neural Networks. International Journal of Interactive Multimedia and Artificial Intelligence. Vol. 6. No.4.

Fang Jin, Zi-Shan Qian, Yu-Ming Chu, & Mati ur Rahman. (2022). On Nonlinear Evolution Model for Drinking Behavior Under Caputo-Fabrizio Derivative [J]. Journal of Applied Analysis & Computation, 12(2): 790-806. doi: 10.11948/2021035.

Gray, K.R., Aljabar, P., Heckemann, R.A., Hammers, A., & Rueckert, D. (2013). Random forestbased similarity measures for multi-modal classification of Alzheimer's disease. NeuroImage. 65, 167–175, https://doi.org/10.1016/j.neuroimage.2012.09.065.

He Z-Y, Abbes A, Jahanshahi H, Alotaibi ND, & Wang Y. (2022). Fractional-Order Discrete-Time SIR Epidemic Model with Vaccination: Chaos and Complexity. Mathematics.10 (2):165. https://doi.org/10.3390/math10020165.

Hinrichs, C., Singh, V., Xu, G., & Johnson, S.C. (2011). Predictive markers for AD in a multimodality framework: an analysis of MCI progression in the ADNI population. NeuroImage. 55, 574–589, https://doi.org/10.1016/j.neuroimage.2010.10.081.

Jun Jie Ng, & et al. (2016). Machine Learning, Wearable Computing and Alzheimer's Disease. Electrical Engineering and computer Sciences University of California at Berkeley.

Lecun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. Nature 521:436, https://doi.org/10.1038/nature14539.

Liu, Feng, & et al. (2014). Inter-modality relationship constrained multi-modality multi-task feature selection for Alzheimer's Disease and mild cognitive impairment identification. NeuroImage. 84: 466-475, https://doi.org/10.1016/j.neuroimage.2013.09.015.

Matsuda H. (2007). The role of neuroimaging in mild cognitive impairment. Neuropathology. 27(6):570, https://doi.org/10.1111/j.1440-1789.2007.00794.x.

Mosconi L, Tsui WH, Herholz K., & et al. (2008). Multicenter standardized 18F-FDG PET diagnosis of mild cognitive impairment, Alzheimer's disease, and other dementias. Journal of Nuclear Medicine. 49(3):390–398, https://doi.org/10.2967/jnumed.107.045385

Rahiminasab A, Tirandazi P, Ebadi MJ, Ahmadian A, & Salimi M. (2020). An Energy-Aware Method for Selecting Cluster Heads in Wireless Sensor Networks. Applied Sciences. 10(21):7886. https://doi.org/10.3390/app10217886.

Riedel, B. C., Daianu, M., Ver Steeg, G., Mezher, A., Salminen, L. E., Galstyan, A., & et al. (2018). Uncovering biologically coherent peripheral signatures of health and risk for Alzheimer's disease in the aging brain. Front. Aging Neurosci.10:390. https://doi.org/10.3389/fnagi.2018.00390.

Scatena, R., Martorana, G. E., Bottoni, P., Botta, G., Pastore, P., & Giardina, B. (2007). An update on pharmacological approaches to neurodegenerative diseases. Expert opinion on investigational drugs. 16(1):59–72., https://doi.org/10.1517/13543784.16.1.59.

Small & D. H. (2005). Acetylcholinesterase inhibitors for the treatment of dementia in alzheimer's disease: do we need new inhibitors? Expert opinion on emerging drugs. 10(4):817–825, https://doi.org/10.1517/14728214.10.4.817.

Sivakani, R & Ansari G.A. (2020). Imputation Using Machine Learning Technique. IEEE International Conference on Computer Communication and Signal Processing (ICCSSP). 1-6, https://doi: 10.1109/ICCCSP49186.2020.9315205.

Sivakani, R & Ansari G.A. (2020). Machine Learning Framework for Implementing Alzheimer's Disease. IEEE Advancing Technology for Humanity International Conference on Communication and Signal Processing (ICCSP), 0588-0592, https://doi. 10.1109/ICCSP48568.2020.9182220.

Trambaiolli, L. R., et al. (2017). Feature selection before EEG classification supports the diagnosis of Alzheimer's disease. Clinical Neurophysiology, 128: 2058-2067, https://doi.org/10.1016/j.clinph.2017.06.251.

Unay, D., Ekin, A., & Jasinschi, R. S. (2010). Local structure-based region-of-interest retrieval in brain mri images. IEEE Transactions on Information technology in Biomedicine. 14(4):897–903, https://doi.org/10.1109/TITB.2009.2038152.

Wiharto, E. Suryani, S. Setyawan and B. P. Putra. (2022). The Cost-Based Feature Selection Model for Coronary Heart Disease Diagnosis System Using Deep Neural Network. IEEE Access. 29687-29697. doi: 0.1109/ACCESS.2022.3158752.

Z. D. Akşehir & E. Kiliç. (2022). How to Handle Data Imbalance and Feature Selection Problems in CNN-Based Stock Price Forecasting. IEEE Access. 31297-31305. doi: 0.1109/ACCESS.2022.3160797.

Zhang, D., Wang, Y., Zhou, L., Yuan, H., & Shen, D. (2011). Multimodal classification of Alzheimer's disease and mild cognitive impairment. NeuroImage. 55, 856–867, https://doi.org/10.1016/j.neuroimage.2011.01.008.

AUTHOR BIOGRAPHIES

Prof. Gufran Ahmad Ansari is a professor in Department of Computer Application and Director for Smart Campus in B.S. Abdur Rahman Crescent Institute of Science and Technology, Chennai, Tamilnadu, India. He received his B.Sc. degree from Lucknow University in Computer science. He also received his M.C.A degree from Dr. B.R. Ambedkar University. He received his Ph.D degree from Lucknow University. He has more than 16 years' experience in teaching. He has contributed more than 50 research articles and conferences in the reputed journals. He has published a book and 2 patents. He has completed 7 funded projects and has membership in more than twenty professional bodies. His research area are Software Engineering, Information System, Management Information System, Decision Support System, Cyber security, UML modeling, Artificial Intelligence, Machine Learning, Data Mining, Data Security etc. Email: gufran@crescent.education.

Sivakani. R received her B.Tech. degree from Anna university in Information Technology. She also received her M.E. degree from Sathyabama University in Computer Science and Engineering. She is currently pursuing her Ph.D degree in B.S. Abdur Rahman Crescent Institute of Science and Technology, Chennai, Tamilnadu, India. Email:sivakani13@gmail.com.

Dr. S. Srisakthi is working assistant professor in Department of Computer Application in B.S. Abdur Rahman Crescent Institute of Science and Technology, Chennai, Tamilnadu, India. She received her B.Sc. degree from Madras University in Computer science. She also received her M.C.A degree from Bharathidasan University. She received her Ph.D degree from Anna University. She has more than 7 years' experience in teaching. She has contributed many research articles and conferences in the reputed journals. Her research area are Cyber security, Computer Networks, Information System, Management Information System, Decision Support System, UML modeling, Artificial Intelligence, Machine Learning, Data Mining, Data Security etc. Email: srisakthi@crescent.education.