FEATURE SELECTION AND CLASSIFICATION OF CLINICAL DATASETS USING BIO-INSPIRED ALGORITHMS AND BPNN

Dr S Murugesan[1*], K Pathmapriya[2], A Nithya[3], Dr R Manikandan[4], S D Lalitha[5], S Karthick Murugan[6]

1. Associate Professor, Department of Computer Science and Engineering, R.M.D Engineering College, Kavaraipettai, Tiruvallur – 601206.
2. Alumna, College of Engineering Guindy, Chennai – 600025.
3. Assistant Professor, Department of Computer Science and Engineering, R.M.K College of Engineering and Technology, Tiruvallur – 601206.
4. Professor, Department of Electronics and Communication Engineering, Panimalar Engineering College, Chennai-600123.
5. Assistant Professor, Department of Computer Science and Engineering, R.M.K Engineering College, Tiruvallur – 601206.
6. Assistant Professor, Department of Computer Science and Engineering, R.M.D Engineering College, Kavaraipettai, Tiruvallur – 601206.

*Corresponding Author: Dr S Murugesan

bsmurugesan@gmail.com

ABSTRACT

Clinical decision Support System (CDSS) aid in analyzing clinical data and using them to predict disease and supply necessary care. In this research work, a paradigm for predict the accuracy of majority classifier. Clinical data classification was performed using two datasets from the University of California, Irvine (UCI) machine learning repository, namely the Hypothyroid and Autistic Screening datasets, which were aided by feature selection using bio-inspired computational methods and Back Propagation Neural Network (BPNN). Bat Algorithm (BA) and Crow Search Algorithm (CSA) are used for feature selection in Wrapper approach method and Stochastic Gradient Descent with Back Propagation is used to evaluate as the fitness function. It entails the selection and categorization of clinical data features in order to improve classification accuracy while minimizing the size of the feature subset and hence the learning duration. The selected optimal features are used to train Ensemble Learning (EL), which consists of classifiers such as Support Vector Machine (SVM), k-Nearest Neighbour (k-NN), and Naive Bayes (NB). Experimentation is carried out using the Hypothyroid and Autistic Screening datasets from the

University of California Irvine machine learning repository. The accuracy of the selected subset was classified by 62 % with BA and 61 % with CSA for Hypothyroid dataset and 71 % with BA and 72 % with CSA for Autistic Screening dataset. Other performance indicators, such as specificity, recall, and accuracy are used to evaluate the classifier's performance.

Keywords: Feature Selection, Classification, Bat Algorithm, Crow Search Algorithm, Back Propagation Neural Network

1. INTRODUCTION

In the era of value-based health care, digital innovation, and huge data, clinical decision support systems tools became vital for health care organizations seeking to reinforce provision. Clinical decision support (CDS) tools have the ability to analysis large volumes of knowledge and counsel next steps for treatment, tired potential problems and enhancing care team efficiency. Machine learning (ML) is also a growing field in medication. The presence of irrelevant, outlier and noisy values can affect the accuracy of the classifiers. The Redundant features and unprocessed medical data are processed for better performance of the system.

Feature selection helps to avoid the problems by reducing the number of features in models, and trying to optimize the model performance. Feature selection algorithms are classified into three varieties as wrapper, filter and embedded algorithms. In wrapper algorithms, the analysis of the feature subsets involves the classifier to be used in the next step, classifier accuracy is employed because the criterion function in most of the cases, Wrapper approaches take account of the classifier that the features are to be selected, yielding higher performance accuracy however incur a high computational cost In filter algorithms, measure the features by their intrinsic discriminatory property irrespective of the classifier and select features irrespective of the classifier, and so they are computationally light with lower accuracy. In embedded algorithms, perform feature selection combine both the feature selection algorithm and learning algorithm are integrated with each other.

One of the most significant aspects of engineering models is optimization. Optimization is a method for determining the best solution to a problem by calculating the solution's quality. The heuristic is an important method for addressing optimization issues. It employs a practical step that is not guaranteed to be the ideal or optimal solution to the problem, but is enough for the current objective. The meta-heuristic algorithms represent population-based global improvement strategies namely Bat Algorithm [4], [7] and Crow Search Algorithm [2-3], [10]. They typically use some sort of variation operators on population members (candidate solutions) to make new solutions. In most activities, we typically have a huge number of options that might be employed. Feature selection is a process for picking a subset of features from a large set of knowledge. The most effective subset of features has the fewest dimensions, which add to the model's accuracy by reducing unnecessary and redundant alternatives.

In this work, a classification of clinical data is aided by feature selection using to bio-inspired algorithm namely Bat algorithm and Crow search algorithm. It is implemented using a wrapper approach and a back Propagation Neural Network (BPNN) with a stochastic Gradient Descent is used to evaluate the feature subsets. Finally, a classification was done using Ensemble Classifier with SVM, k-Nearest Neighbour (k-NN), and Naive Bayes (NB), and the performance of a classifier was assessed.

The following is the paper's structure: A related work is discussed in the second chapter. The third chapter is a dataset description. Chapter 4 presents the proposed classifier. Chapter 5 discusses the findings and evaluation metrics. Finally, Chapter 6 contains a conclusion and a work scope for the future.

2. Literature Survey

Yang, B., et al. (2017) [6], in their study, introduced a feature selection technique dubbed MBAFS that uses a modified Bat Algorithm to avoid premature convergence. These MBAFS variations were used for wrapper-based feature selection, and the classifier was Support Vector Machine (SVM). Experimentation was carried out utilizing twelve datasets from the machine learning repository at the University of California, Irvine (UCI), including WBCD1, Wine, Australian, Zoo, Vehicle, German, WBCD2, Ionosphere, Lung, Sonar, Hill valley, and Musk1. It was deduced that MBAFS achieves higher outcomes than NBBA, BBA, IBPSO, and BPSO in a shorter amount of time.

Jeyasingh, S., et al. (2017) [5] suggested a feature selection strategy based on a modified Bat Algorithm in their study. The bat method uses the simple random sampling approach to select a random instance from the dataset. The most effective solutions are rated, and their characteristics are utilised to train the Random Forest (RF) classifier. The Wisconsin Diagnosis Breast Cancer (WDBC) dataset from the University of California, Irvine (UCI) machine learning repository was used for the experiments. The WDBC has a claimed accuracy of 96.85%. The experimental outcome is compared to Correlation-based Feature Selection (CFS) at 89.51% and Gain Ratio (GR) at 90.21%.

Anter, A. M., et al. (2020) [3] suggested a feature selection mechanism based on a crow search algorithm (CSA), which uses the global optimization approach to prevent local optimization. The chaotic crow search optimization technique employs the fuzzy c-means (FCM) objective function as a cost function. Diabetes, heart, breast cancer, lung cancer, liver disorders, Radiopaedia CT liver, cardiotocography, ILPD (Indian Liver Patient Data), hepatitis, and arrhythmia were among the ten clinical datasets evaluated from the University of California, Irvine (UCI) machine learning repository. It was deduced that the computing time of CFSA is smaller than that of other comparable methods such as Bat Algorithm, BCSA, CALO and BALO.

In their work, Yaghoubzadeh et al. (2021) [2] established a feature selection procedure, applying a Binary Bat Algorithm (BBA) to choose a valuable subset of characteristics, which were then evaluated for classification and assessment criteria. The evaluation criteria were classified using the naïve Bayes (NB), support vector machine (SVM), and J48 algorithms. The performance was evaluated using a breast cancer dataset from the University of California, Irvine (UCI) machine learning repository. The BBA obtained 99.28%, 96.43%, and 92.86% accuracy in the SVM, NB, and J48 algorithms, respectively. According to the findings, feature selection combined with the BBA and SVM algorithms yields good results in sickness diagnosis.

In their paper, Sreejith, S., et al. (2020) [12] proposed a feature selection strategy based on Chaotic Multi-Verse Optimization (CMVO). The Synthetic Minority Over-sampling approach (SMOTE) was used to balance the datasets, which was enhanced utilizing Orchard's approach. A wrapper-based feature selection was performed with these CMVO variants, and the Random Forest (RF) classifier was used as the fitness function. Experimentation was carried out using three datasets from the University of California, Irvine (UCI) machine learning repository, including the Indian Liver Patient Dataset (ILPD), the Thoracic Surgery Dataset (TSD), and the Pima Indian Diabetes (PID). The ILPD is 82.46% accurate, the TSD is 86.88% accurate, and the PID dataset is 89.04% accurate.

In their study, Isaac, A., et al. (2021) [11] suggested a feature selection technique employing an Artificial Bee Colony optimization (I-ABCO) that is enhanced for Rough Dependency Measure (RDM) and mutual information (MI), to encourage better search space exploitation. Support Vector Machine (SVM) was employed as the classifier in a wrapper-based feature selection. The assessment criteria were categorized using a Radial Basis Function Neural Network (RBFNN) classifier and their performance was tested. The Tuberculosis dataset and the Lung Image Database Consortium-Image Database Resource Initiative (LIDC-IDRI) dataset were used in the experiments. For both datasets, the accuracy of CAD systems utilizing I-ABCO with MI and I-

ABCO with RDM is (88.34%, 92.63%) and (87.32%, 90.17%), respectively.

3. Outline of the Datasets Used

In this research work, for binary classification two clinical datasets from the UCI ML repository, Autistic Screening Disorder and Hypothyroid, were employed. Table 1 provides an overview of each dataset used.

Name of the dataset	Number of Instances	Number of Attributes	Number of Class Label	Number of Missing Values
Autistic Screening Disorder	704	22	2	2
Hypothyroid	3163	25	2	2426

Table 1: Outline of dataset Characteristics

4. System Framework

The framework of feature selection and classification for clinical dataset is presented in **Figure 1**. The main components of the CDSS are preprocessing subsystem, Feature Selection subsystem, and Classification subsystem.

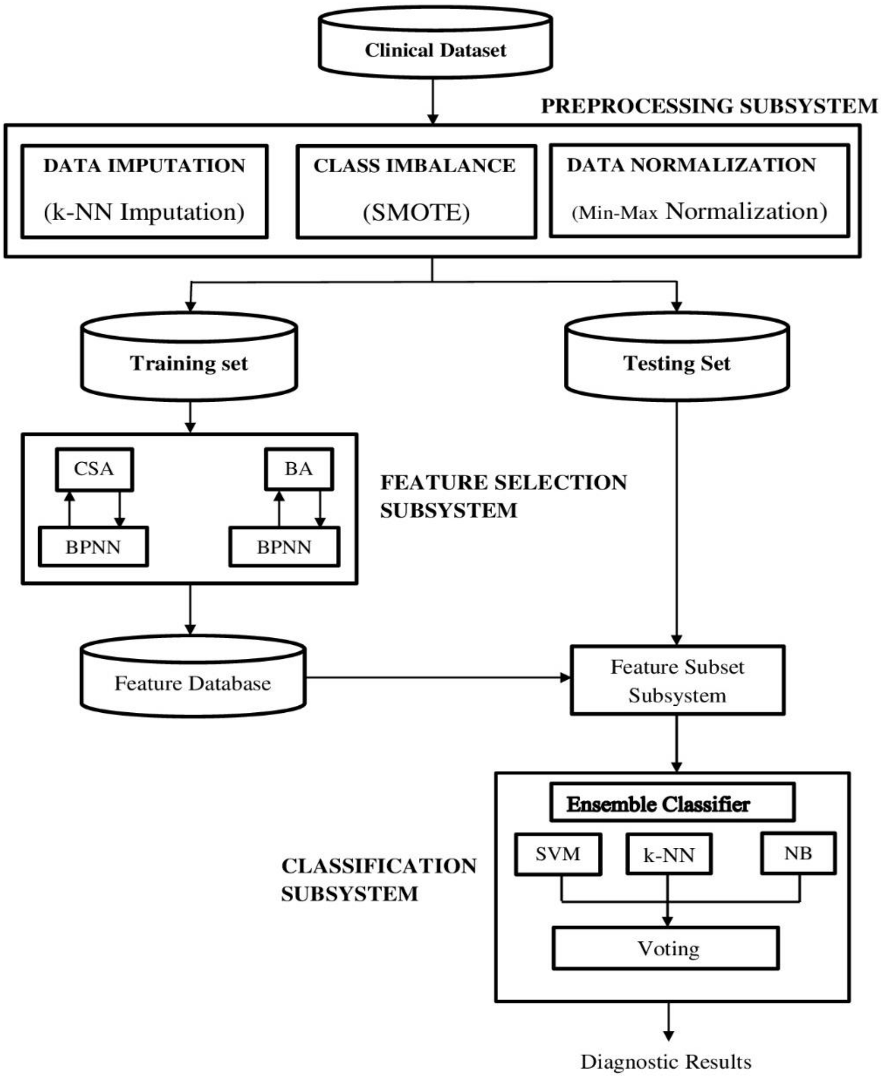

Figure 1: System Framework of Proposed Model

4.1 Pre-processing subsystem

The problem of missing values and class imbalance are affecting the accuracy of the classifiers. The missing values have been imputed by k-NN imputation. The dataset now has a class imbalance, which is balanced by producing more instances from the minority class using the Synthetic Minority Oversampling Technique (SMOTE). A min-max normalization approach is employed on the dataset to scale the values in the range of 0 to 1. The equation represents the min-max normalization.

$$x_{new} = \frac{x - \min(x)}{\max(x) - \min(x)} \quad (1)$$

Where x_{new} represents the normalized value, x represents the attribute's actual value, and min (x) and $\max(x)$ represent the minimum and maximum values of the attribute's range. Because min-max normalization is used to normalize variables from 0 to 1, the value of $\max(x)$ is 1 and min (x) is 0. One strategy for dealing with uneven data in clinical datasets is the minority class. Even if the repeated examples bring no new information to the model, the easiest technique is to duplicate instances from the minority class. Table 1 shows the number of instances in the training set and the number of instances in the testing set.

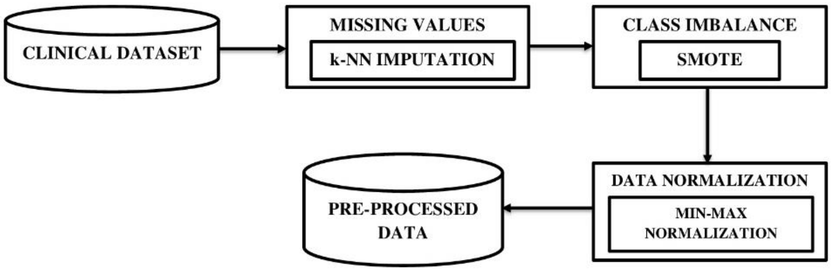

Figure 2: Architecture of Data Pre-Processing

There will be two types of data sets used in supervised learning: training data sets and test data sets. The labelled data instances are included in the training data set. The test data set comprises fresh data that was not present in the training data set. The dataset is divided into two parts: training (80%) and testing (20%).

4.2 Feature selection

Feature selection is a process of reducing the input for processing and analysis; find the most important feature from a dataset. Wrapper approach can be applied for feature selection, it generates relevant feature subsets by using Nature-inspired algorithms including Bat Algorithm (BA) and Crow Search Algorithm have been successfully used for the purpose of feature selection. The selected features can often provide approximately as better results. The selected features will be passed to BPNN with Stochastic Gradient Descent (SGD). This SGD classifier can in turn be used as a fitness function for evaluation.

4.2.1 Outline for CSA Algorithm based feature selection

Crows (crow family or corvids) are considered the most intelligent birds, a population-based meta-heuristic algorithm. Crows have been spotted monitoring other birds, observing where opposing birds hide their food, and then snatching it after the owner has departed. If a crow has stolen anything, it will take extra effort to avoid being a victim again in the local phase, such as moving hiding sites. In fact, as a global phase, they use their own skill as thieves to predict a pilferer's actions and may choose the safest strategy to protect their caches from being pilfered.

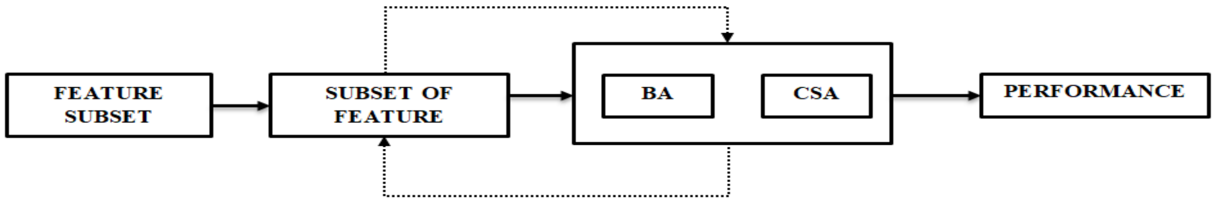

Figure 3: Feature Selection of Wrapper Approach

Table 4 summarizes the factors that are crucial in CSA. Hidden locations associated with the crow's movement while hunting for food.

4.2.2 Outline for BA based Feature Selection

The BA method is a bio-inspired optimization technique that uses echolocation and other microbat-related behavior with variable pulse rates of emission and loudness. The properties of echolocation are as follows: Some bats have developed an incredibly sophisticated sense of hearing; they make noises that bounce off of anything in their path, sending echoes back to the bats. The bats can determine the size of the items, the distance between them, and their speed based on the echoes. Xin-She Yang (2010) [7] created the BA based on these echolocation behaviors of bats.

4.2.3 Classifier Training

Each instance has been preprocessed and divided into 60% training and 40% testing sets. For feature selection, a wrapper technique with two bio-inspired algorithms, CSA and BA, and the classification accuracy of SGD Classifier as the fitness function was utilized. Using SGD, the features chosen by each bio-inspired algorithm are utilized to train two BPNNs separately. Each BPNN has one hidden layer, and the activation function in the hidden layer is linear. The learning rate is 1e-07, with a maximum of 100 repetitions. Because the classification is binary, each BPNN has just one output node, and the output layer's activation function is linear. Figure 3 depicts the training of BPNN classifiers.

Gradient Descent (GD) is the most fundamental but widely used optimization method. It is extensively utilized in linear regression and classification methods. A gradient descent approach is also used in neural networks for back propagation. The loss is passed from one layer to the next, and the model's parameters, also known as weights, are altered based on the losses in attempt to minimize the loss. SGD is a more advanced variant of gradient descent. Frequent model parameter changes converge in less time and need less memory since no loss function values must be stored.

Name of the Dataset	Name of the Algorithm	Number of Instances	Number of Total Attributes	Number of Selected Attributes	Number of Training Instances (80%)	Number of Testing Instances (20%)
Autistic Screening Disorder	BA	1226	22	11	980	246
	CSA			11		
Hypothyroid	BA	4684	25	13	3747	937
	CSA			13		

Table 2: Outline of training and testing set

Table 2 shows the number of training examples for the CSA and BA classifiers. Despite the fact that the majority of the characteristics chosen by each bio-inspired algorithm overlap, it has been deduced that the number of features chosen by each method is not the same. The following steps show how to train the BPNN classifier using two SGD classifiers and the CSA and BA methods.

Input: Training set (selected feature of CSA and BA)

Step 1: Initialize the BPNN's settings, including weights and bias, the number of hidden layers, and the learning rate.

Step 2: The no. of hidden nodes is computed using Equation.

$$H = 2n \tag{10}$$

H is the number of hidden nodes in the above formula, while n is the number of input nodes.

Step 3: Equation 11 is used to compute the input of the hidden layer.

$$Y_k = \sum_{j=1}^{h} w_j \, \varphi(x_j) + \beta_N \tag{11}$$

In the above formula, Y_k is the input of the hidden layer; w_j is the weight of each input nodes; β_N is the bias.

Step 4: Equation 12 is used to compute the output of the hidden layer.

$$O_j = \frac{1}{1} + e^{-Y_k} \tag{12}$$

Where O_j is the output of the j^{th} hidden layer, and Y_k is the input to the neuron from the previous layer.

Step 5: Equation 13 is used to compute the error rate in the predicted output.

$$\text{Error Rate (MSE)} = E_k = \frac{1}{n}\sum_{i=0}^{n}(Y_k - y_k)^2 \tag{13}$$

In the above formula, Y_k is the expected output value, and y_k is the desired output value.

Step 6: Using SGD, update the new weights and bias based on the error rate and learning rate.

Step 7: Repeat steps 2–5 until the error rate converges and When SGD training is completed, the candidate with the highest fitness is chosen as the final model for classification tasks. All SGD upgrades and fitness evaluations have been completed.

Output: Classifiers trained on CSA and BA feature subsets using BPNN.

4.3 Classification of ensemble algorithms

The main aim of the classifier in the CDSS is to correctly predict the result based on the new data. Ensemble classifier operates as a meta-classifier, employing various underlying algorithms such as Naive Bayes (NB), K-nearest neighbor (k-NN), and Support Vector Machine (SVM) on subsamples of the dataset before voting on the best probable class based on the mode classification of the entire individual. The training data and the optimal feature subset acquired from the feature selection subsystem are used to train the classifier, which is then validated using the testing data. The classifier then looks at the majority outputs to determine the best option. The transit of the subset via the classification subsystem is depicted in figure 4.

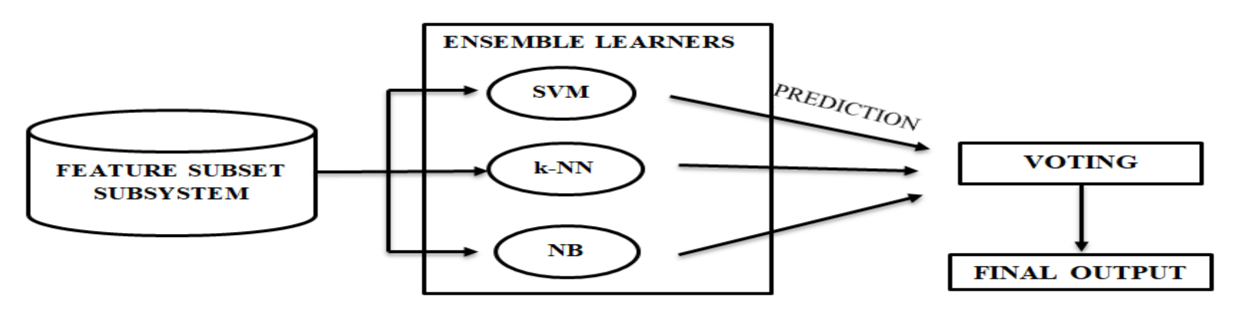

Figure 4: Classification of Ensemble Learner

4.3.1 Ensemble Classifier with voting

Ensemble learning approaches are useful because they can increase a prediction model's performance. The usefulness of an ensemble classifier is that by combining the predictions of numerous classifiers, it may adjust for mistakes caused by any particular classifier, resulting in improved overall accuracy. Ensemble learning approaches are based on the concept that combining the predictions of many classifiers would result in improved performance by either increasing prediction accuracy or decreasing bias and variance. In this part, an ensemble of machine learning techniques was applied, including Support Vector Machine, Naive Bayes, and k-Nearest Neighbor. To improve accuracy, the aforementioned algorithms were combined with a soft voting classifier.

The majority vote classifier is a meta-classifier that uses majority voting to combine any classifier. The final class label would be the one predicted by the majority of the classifiers. The final class label d_J is defined as,

$$d_J = mode\{C_1, C_2, \ldots., C_n\} \tag{15}$$

Where $\{C_1, C_2, \ldots., C_n\}$ represents, the individual classifiers that participate in the voting.

5. Experimental Evaluation

This section discusses the dataset utilized for the proposed work's implementation and the results gained when compared to other classification methods.

5.1 Experiment Design

The confusion matrix, precision, recall, and F-Score are employed as assessment measures in this experiment. Classification accuracy and the number of chosen characteristics are used to assess performance and compare results. The Back Propagation Neural Network with SGD classifier technique is used to assess the correctness of the selected features. To get better results, the algorithm was run multiple times, and the Ensemble Classifier is utilized as a classification model for the selected features in the proposed technique, which aids in model evaluation by dividing the original sample into training and test sets. The fitness function for the SGD classification algorithm is Mean Squared Error.

5.2 Result and Analysis

Tables 3 and 4 provide the confusion matrices for the Hypothyroid and Autistic Screening for Adult datasets, which show the number of true positives (TP), true negatives (TN), false positives (FP), and false negatives (FN) found for both datasets. Tables 3 and 4 show the assessment measures such as precision, recall, and f-measure, which are computed from true positives, true negatives, false positives, and false negatives.

Classifier	Approach	True positive	False Positive	True Negative	False Negative	Accuracy (%)
k-NN	BA + BPNN	342	128	302	165	69
	CSA+BPNN	315	132	30.6	184	67
NB	BA + BPNN	390	68	146	333	58
	CSA+BPNN	414	70	150	303	61
SVM	BA + BPNN	328	130	299	180	71
	CSA+BPNN	417	67	189	264	68
Voting Classifier	BA + BPNN	431	39	128	339	73
	CSA+BPNN	419	65	155	298	71

Table 3: Confusion Matrix for Hypothyroid dataset

Classifier	Approach	True positive	False Positive	True Negative	False Negative	Accuracy (%)
k-NN	BA + BPNN	79	140	106	26	77
	CSA+BPNN	76	32	113	25	82
NB	BA + BPNN	83	37	82	44	64
	CSA+BPNN	67	41	96	42	70
SVM	BA + BPNN	79	41	103	23	74
	CSA+BPNN	92	24	124	6	89
Voting Classifier	BA + BPNN	86	34	110	16	79
	CSA+BPNN	74	34	121	17	79

Table 4: Confusion Matrix for Autistic Screening Dataset

Tables 3 and 4 demonstrate accuracy comparisons of the proposed technique with various current approaches. Figures 5 and 6 illustrate the accuracy gained by several machine-learning algorithms. According to the comparison table, the voting classifier has decent accuracy when compare to other classifier techniques.

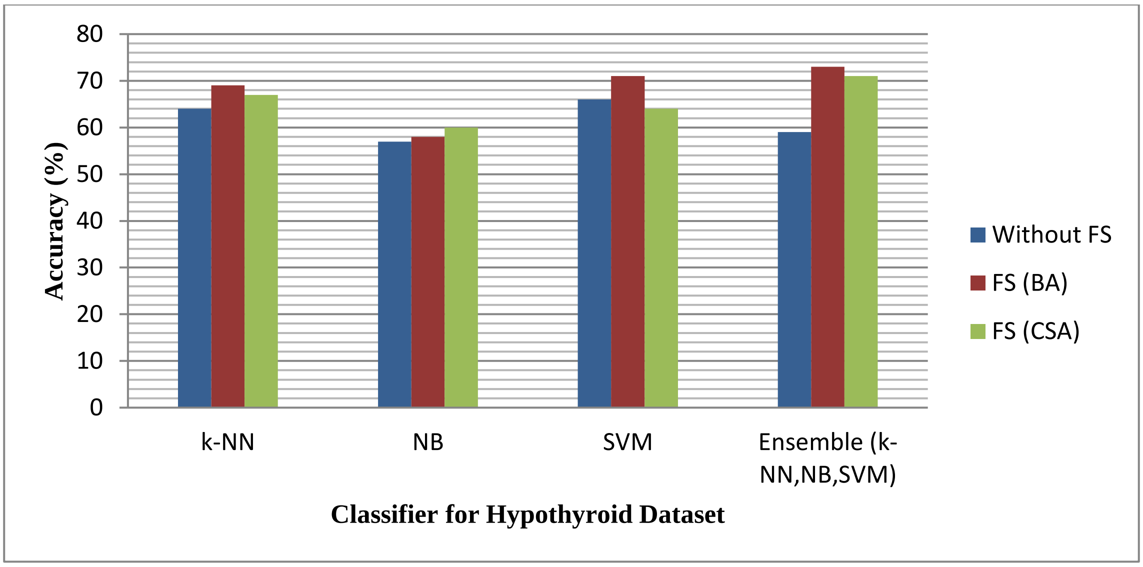

Figure 5: Accuracy obtained by various approaches of Hypothyroid dataset

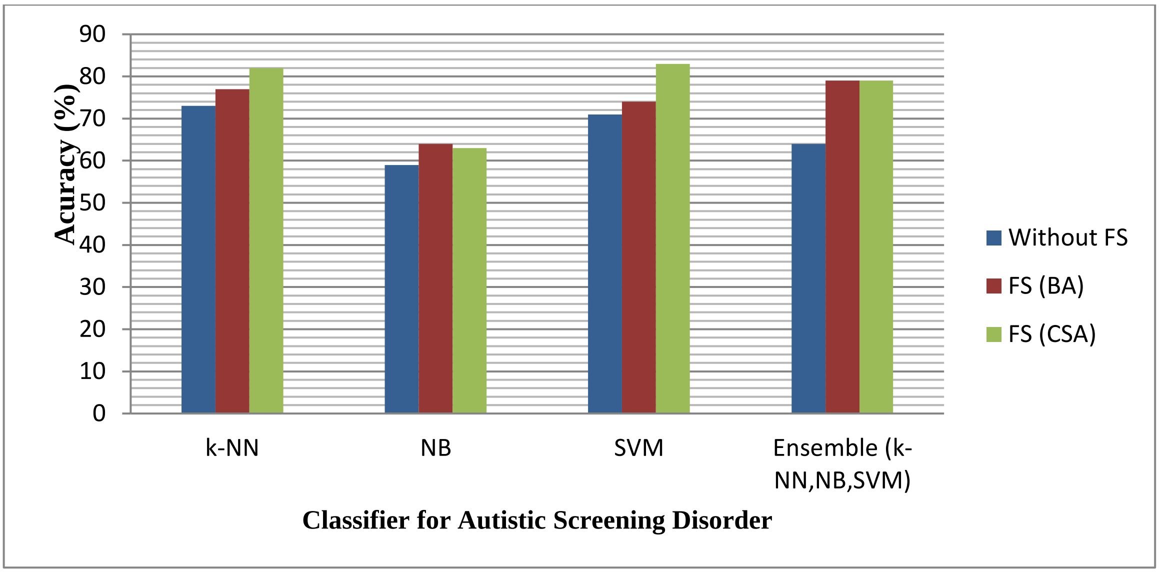

Figure 6: Accuracy obtained by various approaches of Autistic Screening dataset

The proposed framework's performance for feature selection and classification is compared to the performance of existing classifiers, and the results are shown in tables 3 and 4. Figures 5 and 6 show that when the Voting-based Ensemble model is used, the combined performance outperforms the individual performance.

6. Conclusion and Future Work

In this research work, we investigate the predictive accuracy of thyroid illness and autism in adults using an ensemble of classifiers. The UCI machine learning repository's Autistic Screening dataset was used for training and testing. For the experiments, the ensemble algorithms k-NN, NB, SVM, and majority voting were used. The accuracy of CSA was enhanced by up to 82% when k-NN was utilized. When NB was applied, the accuracy of CSA was enhanced by up to 70%. When SVM was utilized, the accuracy of CSA increased by up to 89%. When these classifiers were combined using majority voting, the accuracy for both BA and CSA was 79%. A comparison of findings revealed that majority voting improves accuracy the most. Feature selection techniques were used to improve the performance even further. The feature selection strategies aided in improving the ensemble algorithms' accuracy. Majority voting yielded the best accuracy. Larger datasets and alternative fitness functions can be investigated in future studies. A new approach of adapting swarm intelligence algorithms to discrete issues is another area of future research. We want to improve on this research by incorporating feature selection techniques to minimize the number of characteristics in the Hypothyroid and Autistic Screening dataset. It is expected that the amount of space and time consumed would be reduced further, improving classification accuracy. Other optimization algorithms may also be introduced and compared for application.

7. References

[1] Chaudhuri, A., & Sahu, T. P. (2021). Feature selection using Binary Crow Search Algorithm with time varying flight length. *Expert Systems with Applications, 168, 114288.*

[2] Yaghoubzadeh, R., Kamel, S. R., Barzgar, H., & Moshajeri San'ati, B. (2021). The Use of the Binary Bat Algorithm in Improving the Accuracy of Breast Cancer Diagnosis. *Multidisciplinary Cancer Investigation, 5(1), 1-7.*

[3] Anter, A. M., & Ali, M. (2020). Feature selection strategy based on hybrid crow search optimization algorithm integrated with chaos theory and fuzzy c-means algorithm for medical diagnosis problems. *Soft Computing, 24(3), 1565-1584.*

[4] Naik, S. M., Jagannath, R. P. K., & Kuppili, V. (2020). Bat algorithm-based weighted Laplacian probabilistic neural network. *Neural Computing and Applications, 32(4), 1157-1171.*

[5] Jeyasingh, S., & Veluchamy, M. (2017). Modified bat algorithm for feature selection with the Wisconsin Diagnosis Breast Cancer (WDBC) dataset. *Asian Pacific journal of cancer prevention: APJCP, 18(5), 1257.*

[6] Yang, B., Lu, Y., Zhu, K., Yang, G., Liu, J., & Yin, H. (2017). Feature selection based on modified bat algorithm. *IEICE TRANSACTIONS on Information and Systems, 100(8), 1860-1869.*

[7] Yang, X. S., & Gandomi, A. H. (2012). Bat algorithm: a novel approach for global engineering optimization. *Engineering computations.*

[8] Ozcift, A., & Gulten, A. (2011). Classifier ensemble construction with rotation forest to improve medical diagnosis performance of machine learning algorithms. *Computer methods and programs in biomedicine, 104(3), 443-451.*

[9] Yang, X. S. (2010). A new metaheuristic bat-inspired algorithm. In Nature inspired cooperative strategies for optimization (NICSO 2010) (pp. 65-74). *Springer, Berlin, Heidelberg.*

[10] Hussien, A. G., Amin, M., Wang, M., Liang, G., Alsanad, A., Gumaei, A., & Chen, H. (2020). Crow search algorithm: theory, recent advances, and applications. *IEEE Access, 8, 173548-173565.*

[11] Isaac, A., Khanna Nehemiah, H., & Kannan, A. (2021). Computer-Aided Diagnosis System for Diagnosis of Cavitary and Miliary Tuberculosis Using Improved Artificial Bee Colony Optimization. *IETE Journal of Research*, 1-20.

[12] Sreejith, S., Nehemiah, H. K., & Kannan, A. (2020). 'Clinical data classification using an

enhanced SMOTE and chaotic evolutionary feature selection'. *Computers in Biology and Medicine, 126, 103991.*

[13] S. Murugesan, R. S. Bhuvaneswaran, H. Khanna Nehemiah, S. Keerthana Sankari, Y. Nancy Jane, "Feature Selection and Classification of Clinical Datasets Using Bioinspired Algorithms and Super Learner", Computational and Mathematical Methods in Medicine, vol. 2021, Article ID 6662420, 18 pages, 2021. https://doi.org/10.1155/2021/6662420.

[14] Pacha Shoba Rani ,A Vasantharaj, , Sirajul Huque, KS Raghuram, R Ganeshvkumar, Sebahadin Nasir Shafi "Automated Brain Imaging Diagnosis and Classification Model using Rat Swarm Optimization with Deep Learning based Capsule Network" Publication date2021/7/12 International Journal of Image and Graphics Pages 2240001 Publisher World Scientific Publishing Company.

[15] Ouadfel, S., & Abd Elaziz, M. (2020). Enhanced crow search algorithm for feature selection. *Expert Systems with Applications*, *159*, 113572.