Chapter 4

FFBAT-RBFL Algorithm for Heart Disease Classification

4.1 INTRODUCTION

In previous chapter a novel FFBAT-RBFL algorithm has been proposed and applied on diabetes dataset and a better classification accuracy has been achieved. To test the consistency of the proposed algorithm on other medical datasets, in this chapter FFBAT-RBFL algorithm is used for diagnosing heart disease. The Hungarian, Switzerland, and Cleveland heart disease datasets have been used to evaluate the accuracy performance of the algorithm. The obtained experimental results prove that FFBAT-RBFL classifier has outperformed the existing classifiers. In this chapter we used FFBAT-RBFL algorithm, which was proposed in chapter 3 to classify heart disease datasets. The algorithm is experimented on heart disease datasets. The proposed classifier has the following two steps, i) Feature or attribute reduction by employing LPP algorithm ii) Classification of Heart disease by FFBAT-RBFL classifier.

4.2 HEART DISEASE PREDICTION USING RULE BASED FFBAT-FUZZY CLASSIFIER

Heart disease classification is a major research area because it is the major cause of death universally over the past decade. Many researchers integrated fuzzy techniques with other methods like genetic algorithms, decision tree, neural networks for competent classification. Genetic algorithms, Neural network and decision tree are applied to produce the rules. The produced rules from these methods are next fed to fuzzy rule base for classification process. These methods dint attain the maximum prediction accuracy. Hence, we have applied FFBAT algorithm, which was proposed in chapter 3 to produce the fuzzy rules and the rule produced from the FFBAT algorithm are employed for classification. The method of the rule based fuzzy classifier (RBFL) is done in 2 steps, i) Feature reduction is performed by LPP algorithm ii) Heart disease classification is performed by using of FFBAT-RBFL. The architecture of the proposed system is illustrated in Figure 4.1. The steps given in section 3.2 are followed for heart disease prediction.
4.3. RESULTS AND DISCUSSION

Matlab version 8.4 is used for implementation of the proposed approach. Accuracy of the proposed method is the ratio of the total number of TP and TN to the total number of data.

4.3.1. DATASET DESCRIPTION

The widely-experimented datasets to experiment with the proposed system namely, Hungarian, Cleveland, and Switzerland heart disease datasets which are publicly available from UCI datasets are used. The detailed explanation of dataset description is followed here;

i) Cleveland data: 14 out of 76 features present in the Cleveland heart disease database are used. Machine Learning researchers employ Cleveland dataset widely to carry out their research. Class distributions is as follows: 54% of the tuples where heart disease is “absent”, while 46% of the tuples have heart disease label as “present”.

ii) Hungarian data: Three of the features have been removed due to a vast % of missing values. The format of the data is similar to Cleveland data. Due to missing values 34 examples of the database were removed. Class distributions in this dataset are 62.5% heart disease “not present” and 37.5% heart disease “present”.

iii) Switzerland data: This dataset has additional missing values. It encloses 14 features and 123 tuples. Class distributions are 93.5% heart disease “present” and 6.5% heart disease.

4.3.2. EVALUATION OF PERFORMANCE

The performance of the RBFL algorithm is evaluated based on accuracy. The input dataset is divided into 10, 15, and 20 samples respectively, where 7,10 and 14 are considered as training
dataset and the remaining 3, 5 and 6 samples as testing dataset. This procedure is carried out for all the iterations and the accuracy is computed for each time. Here size of the population is changed and accuracy is checked for different datasets using the proposed approach. Figure 4.2 to 4.4 shows accuracy performance of the datasets i.e., Cleveland, Hungarian, Switzerland datasets with different population size.

The figures 4.2 to 4.5 show the particular overall performance of various methods, including firefly, BAT and Hybrid FFBAT algorithm. Among these our proposed hybrid FFBAT usually outperforms various other approaches in terms of accuracy. The accuracy of firefly algorithm is 69.52%, BAT algorithm is 70.76% and the hybrid FFBAT achieves 76.51% which is better than other two methods. Figure 4.9 shows the hybrid approach achieves 73.62% accuracy which is better than other two individual approaches. Moreover, the figure 4.9 shows the performance of iteration vs. fitness.

Fig 4. 2 Performance Evaluation of Accuracy Plot of Population Size 10
Fig 4.3  Performance Evaluation of Accuracy Plot of Population Size 15

Fig 4.4  Performance Evaluation of Accuracy Plot of Population Size 20

Fig 4.5  Accuracy Plot between Proposed and Existing Approach of Population Size 10
Fig 4.6  Accuracy Plot between Proposed and Existing Approach of Population Size 15

Fig 4.7  Accuracy Plot Between Proposed and Existing Approach of Population Size 20

Fig 4.8  Performance of Iteration vs. Fitness
4.4.3 COMPARISON OF THE ACCURACY WITH OTHER REPORTED RESULTS

In this section, proposed RBFL based heart disease prediction approach is compared against one of the recently published literature (Srinivas, K. et al., 2014) where the authors developed a rough-fuzzy classifier for heart disease prediction. They combined fuzzy set with rough-set theory. They divided the overall process into two major steps such as (i) Rough set theory is used for rule generation and (ii) Fuzzy classifier is used for prediction.

![Accuracy Plot between Proposed and Existing Approach](image)

Fig 4.9  Accuracy Plot between Proposed and Existing Approach

The above figure 4.8 shows the performance of accuracy between proposed RBFL based heart disease predictions. In this proposed approach, we reduce the feature based on LPP algorithm and generate fuzzy rule based on FFBAT algorithm. We obtain the maximum accuracy of 77% which is 68% for using rough-set fuzzy. As per the analysis, our proposed approach achieves the maximum accuracy compare to the existing approach.

4.5. SUMMARY

In this chapter, the model proposed in Chapter 3 has been used to classify heart disease datasets. The heart disease prediction frameworks introduced in this chapter consist of two key components; mainly feature reduction, and heart disease prediction using rule based fuzzy classifier. At first, we selected the data from the dataset for further processing. After that we reduced the features of the data using LPP algorithm to improve the prediction accuracy. Finally, the data are classified using fuzzy classifier. The results of experimentation demonstrated that the proposed model achieved better accuracy as compared to the existing models used for heart disease classification. Hence the algorithm worked efficiently on heart disease dataset also. Further FFBAT algorithm can be modified or blended with other mathematical and optimization techniques to achieve better results.