MIT2018050,MIT2018061,MIT2018072

Report on implementation results of "Feature Selection Method Based on Grey Wolf Optimization for Coronary Artery Disease Classification"

In this paper, we used Grey Wolf Optimization [1] along with Support Vector Machine (GWO-SVM) classifier for effective feature selection to help with diagnosis of Coronary Artery Disease (CAD). GWO-SVM works in two stages where in the first stage we use GWO for effective feature selection from the Cleveland heart dataset. Initially GWO produce the initial positions of the population and then current positions with each iteration until a stopping criteria is satisfied. In second stage, the SVM is used for classification on the optimal feature subset obtained from the first stage.

The dataset has been used from [2]. It has thirteen features and one target variable. Firstly SVM is implemented after preprocessing and the accuracy obtained is 59.8%. In the suggested method, the feature selection of GWO has been applied on the dataset to get the best features in terms of alpha solution of Grey Wolf optimiser. A certain threshold is decided based on the numbers obtained which helps in feature selection. Also, for the GWO optimisation, a benchmark function has been predecided. The accuracy obtained after implementing SVM on the selected feature is 62.2%, which is increased. The following table suggests the comparison:

Accuracy(%)	SVM	GWO-SVM
Result obtained	59.8	62.2
Paper suggested	76.57	89.83

There are several factors that decide the accuracy obtained in GWO-SVM. These include:

1. Benchmark function
2. Number of wolves
3. Iterations
4. Threshold for alpha solution.
5. SVM kernel

A certain combination of above produces the maximum accuracy which has been mentioned in the paper. The threshold of alpha solution can't be fixed as it changes in each iteration and certainly acts as a deciding factor of the final result obtained.

References:

[1] Al-Tashi, Qasem & Rais, Helmi & Jadid Abdulkadir, Said. (2018). "Feature Selection Method Based on Grey Wolf Optimization for Coronary Artery Disease Classification." 257-266. 10.1007/978-3-319-99007-1_25.

[2] Cleveland dataset. http://archive.ics.uci.edu/ml/machine-learning-databases/heart-disease/processed.cleveland.data.