Prajwal Kailas
14CO132
9632080195
prajwal967@gmail.com

Feature Selection Inspired Classifier Ensemble Reduction

Finalized Work:

- Show the working of machine learning algorithms such as linear regression, logistic regression, MLP, decision trees and random forest.
- Show the working of ensemble strategies such as blending, stacking, weighted average. These strategies will be used to build the ensemble classifier.
- Automate the ensemble building process. The training, constructing of base classifiers and ensemble classifiers should be automated.
- Apply harmony search to the ensemble classifier to perform classifier ensemble reduction. Harmony search will be used to perform feature selection.
- The feature selection is used to select a subset of the base classifiers to reduce system overhead and memory usage.
- Run tests on the final ensemble classifier. The final classifier will be a reduced set of the ensemble classifier.
- Set up a GUI for the user to use.

Completed Work:

- Dataset : Bank Dataset (UCI Machine Learning Repository) . The dataset was used in "A Data-Driven Approach to Predict the Success of Bank Telemarketing."
- Dataset contains 41188 examples and 20 inputs. The classification task is to predict if the client will subscribe to a term deposit.
- Built and integrated linear regression, logistic regression, MLP, decision trees and random forest algorithms. These algorithms will form the base classifiers.
- Constructed a framework for performing stacking, blending and weighted average. This will form the ensemble classifiers.
- Automating the ensemble process is partially completed. The building and training of the base classifiers has been automated.
- The automation will include selecting the base models, the selected base models will be trained. The next step is to select the ensemble methods, upon which the ensemble classifier will be trained.
- Libraries: functools, numpy, sklearn, pandas, keras

Pending Work:

- Complete the automation of the ensemble process. Automate building the new dataset and training the ensemble classifiers.
- Build, code and test the harmony search algorithm. The harmony search should select a subset of the base classifiers in the ensemble classifier such that the accuracy and diversity remain the same.
- Obtain the reduced ensemble classifier, this will be the final ensemble classifier and can be used to predict the outputs of the test dataset holdout.
- Setting up the GUI.

Additional Information:

- The paper constructs the ensemble classifier manually.
- The contribution to the paper is the automation of the ensemble classifier building and training process.
- Harmony search is an optimisation algorithm. The focus will be on the optimisation portion.