1. Classifier parameters setting

Five classifiers are used to ensure that the evaluation of the performances is not limited to a specific classifier: Logistic Regression (LR), Decision Tree(DT), Light Gradient Boosting Machine (LightGBM), Gradient Boosting Decision Tree (GBDT), and eXtreme Gradient Boosting (XGBoost). The parameters of all classifiers are set default as follows:

- LR[1]: norm used in the penalization penalty='l2', tolerance for stopping criteria tol=1e-4, inverse of regularization strength C=1.0, algorithm used in the optimization process solver='lbfgs', maximum number of iterations max_iter=100, and others are set default as 'None' or 0;

- DT[2]: function to measure the quality of a split criterion='gini', strategy used to choose the split at each node splitter='best', minimum number of samples required to split an internal node min_samples_split=2, minimum number of samples required to be at a leaf node min_samples_leaf=1, and others are set default as 'None' or 0;

- LightGBM[3]: boosting_type='gbdt', maximum tree leaves for base learners num_leaves=31, maximum tree depth max_depth=-1, boosting learning rate learning_rate=0.1, number of boosted trees to fit n_estimators=100, and others are set default as 'None' or 0;

- GBDT[4]: loss function to be optimized loss='deviance', boosting learning rate learning_rate=0.1, number of boosting stages to perform n_estimators=100, function to measure the quality of a split criterion='friedman_mse', minimum number of samples required to split an internal node min_samples_split=2, minimum number of samples required to be at a leaf node min_samples_leaf=1,

[1] https://scikit-learn.org/stable/modules/generated/sklearn.linear_model.LogisticRegression.html
[2] https://scikit-learn.org/stable/modules/generated/sklearn.tree.DecisionTreeClassifier.html
[3] https://lightgbm.readthedocs.io/en/latest/Parameters.html
[4] https://scikit-learn.org/stable/modules/generated/sklearn.ensemble.GradientBoostingClassifier.html

maximum depth of the individual regression estimators max_depth=3, tolerance for the early stopping tol=1e-4, and others are set default as 'None' or 0;

- XGBoost[5]: number of boosting rounds n_estimators=100, feature importance type importance_type='gain', and others are set default as 'None' or 0.

[5]https://xgboost.readthedocs.io/en/latest/parameter.html