

Appendix Table 1. Functions, Packages, and Tuning Parameters in the Anaconda Software Used for Each Machine Learning Algorithm

Algorithm	Classifier	Package	Tuning Parameters
Logistic regression	LogisticRegression	from sklearn.linear_model import LogisticRegression	penalty='l2',tol=0.0001,C=0.7,fit_intercept=True,intercept_scaling=1,class_weight=None,max_iter=100,multi_class='ovr',verbose=0,warm_start=False,n_jobs=-1
DecisionTree	DecisionTreeClassifier	from sklearn.tree import DecisionTreeClassifier	criterion='gini',splitter='best',max_depth=7,min_samples_split=20,min_samples_leaf=1,min_weight_fraction_leaf=0.0,max_features=None,random_state=None,max_leaf_nodes=None,min_impurity_decrease=0.0,min_impurity_split=None,class_weight=None,preresort=False
forest	RandomForestClassifier	from sklearn.ensemble import RandomForestClassifier	n_estimators= 200, n_jobs = 100,min_samples_split=40,min_samples_leaf=10
GradientBoosting	GradientBoostingClassifier	from sklearn.ensemble import GradientBoostingClassifier	learning_rate= 0.01,n_estimators= 100,min_samples_split=10,min_samples_leaf=1,subsample= 0.5,max_depth=5
gbm	lgb.LGBMClassifier	lightgbm 2.2.0	boosting_type= 'gbdt', reg_alpha= 0.001, reg_lambda= 0.8, learning_rate=0.1, max_depth=1, n_estimators= 100, objective='binary'

Appendix Figure 1 Flow chart of data processing by machine learning

