**Supplementary material**

**Supplementary Table S1 -** Gridsearch search hyperparameters.

|  |  |
| --- | --- |
| **Model** | **Param\_ grids** |
| GaussianNB | var\_smoothing = [1e-9, 1e-8, 1e-7, 1e-6] |
| DecisionTree | criterion = [‘gini’, ‘entropy’], max\_depth = [None, 5, 10, 15], min\_samples\_split = [2, 5, 10], min\_samples\_leaf = [1, 3, 5], max\_features = [‘sqrt’, ‘log2’, None] |
| Adaboost | n\_estimators = [50,100,200], learning\_rate = [0.01, 0.1, 1], algorithm = ’SAMME’, estimator = [DecisionTreeClassifier(max\_depth=1, random\_state=42), DecisionTreeClassifier(max\_depth=3, random\_state=42),  DecisionTreeClassifier(max\_depth=5, random\_state=42)] |
| KNN | n\_neighbors = [3 ,5, 7, 11], weights = [‘uniform’, ‘distance’], algorithm = [‘auto’, ‘ball\_tree’, ‘kd\_tree’, ‘brute’], p = [1, 2] |
| RF | n\_estimators = [100, 200, 300], criterion = [‘gini’, ‘entropy’], max\_depth = [None, 10, 20], min\_samples\_split = [2, 5], min\_samples\_leaf = [1, 3], max\_features = [‘sqrt’, ‘log2’] |
| ExtraTree | n\_estimators = [100, 200, 300], criterion = [‘gini’, ‘entropy’], max\_depth = [None, 10, 20], min\_samples\_split = [2, 5], min\_samples\_leaf = [1, 3], max\_features = [‘sqrt’, ‘log2’] |
| GradientBoosting | n\_estimators = [100, 200, 300], learning\_rate = [0.01, 0.1], max\_depth = [3, 5, 7], min\_samples\_split = [2, 5], min\_samples\_leaf = [1, 3], subsample = [0.8, 1.0] |
| MLP | hidden\_layer\_sizes = [(50,), (100,). (50, 50)], activation = [‘relu’, ‘tanh’, ‘logistic’], solver = [‘adam’, ‘sgd’, ‘lbfgs’], alpha = [0.0001, 0.001], learning\_rate = [‘constant’, ‘adaptative’], max\_iter = 300 |
| KMeans | n\_clusters = [2, 3, 4], init = [‘k-means++’, ‘random’], n\_init = [10, 20] |
| XGB | n\_estimators = [100, 200], learning\_rate = [0.01, 0.1], max\_depth = [3, 5], subsample = [0.8, 1.0], colsample\_bytree = [0.8, 1.0], eval\_metric = ‘logloss’ |
| Logistic Regression | penalty = [‘l1’, ‘l2’, ‘elasticnet’, None], C = [0.001, 0.01, 0.1, 1, 10, 100], solver = [‘newton-cg’, lbfgs’, ‘liblinear’, ‘sag’, ‘saga’], l1\_ratio = [None, 0.25, 0.5, 0.75], |
| SVM | C = [0.1, 1. 10], kernel = [‘linear’, rbf’], gamma = [‘scale’, ‘auto’] |