

Table 1: Base Models

Model	Parameters	FS Percentile	F1 Train Score	F1 Validation Score
MLPClassifier	hidden_layer_sizes=(20, 20), learning_rate_init=0.01, max_iter=1000	90	0.39 \pm 0.02	0.37 \pm 0.01
MLPClassifier	hidden_layer_sizes=(20, 20), learning_rate_init=0.01, max_iter=1000	70	0.37 \pm 0.02	0.36 \pm 0.02
MLPClassifier	hidden_layer_sizes=(20, 20), learning_rate_init=0.01, max_iter=1000	50	0.36 \pm 0.02	0.34 \pm 0.01
LogisticRegression	class_weight='balanced', max_iter=1000	90	0.32 \pm 0.0	0.3 \pm 0.01
LogisticRegression	class_weight='balanced', max_iter=1000	70	0.31 \pm 0.01	0.3 \pm 0.01
LogisticRegression	class_weight='balanced', max_iter=1000	50	0.3 \pm 0.01	0.29 \pm 0.02
KNeighborsClassifier	n_neighbors=35, p=1, weights='distance'	50	1.0 \pm 0.0	0.34 \pm 0.01
KNeighborsClassifier	n_neighbors=35, p=1, weights='distance'	70	1.0 \pm 0.0	0.34 \pm 0.01
KNeighborsClassifier	n_neighbors=35, p=1, weights='distance'	90	1.0 \pm 0.0	0.34 \pm 0.0
GaussianNB	base model	50	0.27 \pm 0.01	0.27 \pm 0.01
GaussianNB	base model	70	0.18 \pm 0.04	0.18 \pm 0.03
GaussianNB	base model	90	0.13 \pm 0.01	0.12 \pm 0.01
DecisionTreeClassifier	class_weight='balanced', max_depth=15	90	0.46 \pm 0.01	0.34 \pm 0.03
DecisionTreeClassifier	class_weight='balanced', max_depth=15	70	0.46 \pm 0.01	0.34 \pm 0.03
DecisionTreeClassifier	class_weight='balanced', max_depth=15	50	0.45 \pm 0.01	0.33 \pm 0.04

Table 2: Ensembles

Model	Parameters	F1 Train Score	F1 Validation Score
RandomForestClassifier	max_depth: 10, min_samples_split: 2, n_estimators: 100	0.36 \pm 0.01	0.35 \pm 0.0
RandomForestClassifier	max_depth: 10, min_samples_split: 2, n_estimators: 200	0.36 \pm 0.01	0.35 \pm 0.0
RandomForestClassifier	max_depth: 10, min_samples_split: 5, n_estimators: 100	0.36 \pm 0.0	0.35 \pm 0.0
RandomForestClassifier	max_depth: 10, min_samples_split: 5, n_estimators: 200	0.36 \pm 0.01	0.35 \pm 0.0
RandomForestClassifier	max_depth: 5, min_samples_split: 2, n_estimators: 200	0.25 \pm 0.01	0.25 \pm 0.01
RandomForestClassifier	max_depth: 5, min_samples_split: 5, n_estimators: 100	0.25 \pm 0.01	0.25 \pm 0.01
RandomForestClassifier	max_depth: 5, min_samples_split: 5, n_estimators: 200	0.25 \pm 0.01	0.25 \pm 0.01
RandomForestClassifier	max_depth: 5, min_samples_split: 2, n_estimators: 100	0.25 \pm 0.01	0.25 \pm 0.0
GradientBoostingClassifier	learning_rate: 0.05, max_depth: 5, n_estimators: 100	0.52 \pm 0.03	0.4 \pm 0.01
GradientBoostingClassifier	learning_rate: 0.1, max_depth: 5, n_estimators: 50	0.46 \pm 0.01	0.4 \pm 0.01
GradientBoostingClassifier	learning_rate: 0.1, max_depth: 5, n_estimators: 100	0.48 \pm 0.01	0.4 \pm 0.01
GradientBoostingClassifier	learning_rate: 0.05, max_depth: 5, n_estimators: 50	0.49 \pm 0.02	0.39 \pm 0.01
GradientBoostingClassifier	learning_rate: 0.05, max_depth: 10, n_estimators: 50	0.7 \pm 0.01	0.39 \pm 0.01
GradientBoostingClassifier	learning_rate: 0.05, max_depth: 10, n_estimators: 100	0.75 \pm 0.01	0.39 \pm 0.01
GradientBoostingClassifier	learning_rate: 0.1, max_depth: 10, n_estimators: 50	0.72 \pm 0.01	0.39 \pm 0.01
GradientBoostingClassifier	learning_rate: 0.1, max_depth: 10, n_estimators: 100	0.8 \pm 0.01	0.38 \pm 0.02
BaggingClassifier	estimator=MLPClassifier(hidden_layer_sizes=(20, 20), learning_rate_init=0.01, max_iter=1000, max_features: 1.0, max_samples: 0.8, n_estimators=10)	0.4 \pm 0.01	0.38 \pm 0.02
BaggingClassifier	estimator=MLPClassifier(hidden_layer_sizes=(20, 20), learning_rate_init=0.01, max_iter=1000, max_features: 0.8, max_samples: 0.8, n_estimators=10)	0.38 \pm 0.01	0.37 \pm 0.01
BaggingClassifier	estimator=MLPClassifier(hidden_layer_sizes=(20, 20), learning_rate_init=0.01, max_iter=1000, max_features: 1.0, max_samples: 1.0, n_estimators=10)	0.39 \pm 0.02	0.37 \pm 0.0
BaggingClassifier	estimator=MLPClassifier(hidden_layer_sizes=(20, 20), learning_rate_init=0.01, max_iter=1000, max_features: 0.8, max_samples: 1.0, n_estimators=10)	0.36 \pm 0.01	0.36 \pm 0.01
AdaBoostClassifier	learning_rate: 0.05, n_estimators: 50	0.2 \pm 0.0	0.2 \pm 0.0
AdaBoostClassifier	learning_rate: 0.05, n_estimators: 100	0.2 \pm 0.0	0.2 \pm 0.0
AdaBoostClassifier	learning_rate: 0.1, n_estimators: 50	0.2 \pm 0.0	0.2 \pm 0.0
AdaBoostClassifier	learning_rate: 0.1, n_estimators: 100	0.2 \pm 0.0	0.2 \pm 0.0

Table 3: MLP Optimization (Top 15)

Activation Function	Architecture	Learning Rate	Initial Learning Rate	Solver	F1 Train Score	F1 Validation Score
relu	(30, 30)	adaptive	0.02	sgd	0.42 ± 0.01	0.4 ± 0.02
relu	(30, 30)	constant	0.02	sgd	0.42 ± 0.01	0.4 ± 0.01
tanh	(30, 30)	constant	0.01	sgd	0.41 ± 0.01	0.4 ± 0.01
relu	(20, 20, 20)	adaptive	0.02	sgd	0.41 ± 0.01	0.39 ± 0.02
relu	(30, 30)	adaptive	0.01	adam	0.41 ± 0.01	0.39 ± 0.01
tanh	(20, 20, 20)	adaptive	0.00	sgd	0.41 ± 0.01	0.39 ± 0.01
relu	(30, 30)	adaptive	0.01	sgd	0.41 ± 0.01	0.39 ± 0.02
relu	(30, 30)	adaptive	0.00	sgd	0.41 ± 0.01	0.39 ± 0.01
relu	(30, 30)	constant	0.00	sgd	0.41 ± 0.01	0.39 ± 0.01
tanh	(20, 20, 20)	constant	0.01	sgd	0.41 ± 0.01	0.39 ± 0.02
tanh	(20, 20, 20)	constant	0.02	sgd	0.41 ± 0.01	0.39 ± 0.02
relu	(30, 30)	constant	0.01	sgd	0.41 ± 0.01	0.39 ± 0.01
logistic	(30, 30)	adaptive	0.00	adam	0.41 ± 0.01	0.39 ± 0.01
tanh	(30, 30)	adaptive	0.00	sgd	0.41 ± 0.01	0.39 ± 0.02
tanh	(30, 30)	adaptive	0.02	sgd	0.42 ± 0.0	0.39 ± 0.01

Table 4: Ensembles Optimization - Bagging

Model	Parameters	F1 Train Score	F1 Validation Score
BaggingClassifier	estimator=MLPClassifier(activation=tanh, hidden_layer_sizes=(30, 30), learning_rate_init=0.01 max_iter=1000, solver=sgd , max_features: 1.0, max_samples: 0.2, n_estimators: 10, n_estimators=10	0.4 \pm 0.01	0.38 \pm 0.01
BaggingClassifier	estimator=MLPClassifier(activation=tanh, hidden_layer_sizes=(30, 30), learning_rate_init=0.01 max_iter=1000, solver=sgd , max_features: 1.0, max_samples: 0.4, n_estimators: 30, n_estimators=10	0.4 \pm 0.01	0.38 \pm 0.01
BaggingClassifier	estimator=MLPClassifier(activation=tanh, hidden_layer_sizes=(30, 30), learning_rate_init=0.01 max_iter=1000, solver=sgd , max_features: 1.0, max_samples: 0.4, n_estimators: 50, n_estimators=10	0.41 \pm 0.01	0.38 \pm 0.01
BaggingClassifier	estimator=MLPClassifier(activation=tanh, hidden_layer_sizes=(30, 30), learning_rate_init=0.01 max_iter=1000, solver=sgd , max_features: 0.8, max_samples: 0.4, n_estimators: 10, n_estimators=10	0.39 \pm 0.0	0.37 \pm 0.01
BaggingClassifier	estimator=MLPClassifier(activation=tanh, hidden_layer_sizes=(30, 30), learning_rate_init=0.01 max_iter=1000, solver=sgd , max_features: 0.8, max_samples: 0.4, n_estimators: 30, n_estimators=10	0.39 \pm 0.01	0.37 \pm 0.01
BaggingClassifier	estimator=MLPClassifier(activation=tanh, hidden_layer_sizes=(30, 30), learning_rate_init=0.01 max_iter=1000, solver=sgd , max_features: 1.0, max_samples: 0.4, n_estimators: 10, n_estimators=10	0.4 \pm 0.01	0.37 \pm 0.01
BaggingClassifier	estimator=MLPClassifier(activation=tanh, hidden_layer_sizes=(30, 30), learning_rate_init=0.01 max_iter=1000, solver=sgd , max_features: 0.8, max_samples: 0.4, n_estimators: 50, n_estimators=10	0.38 \pm 0.02	0.37 \pm 0.0
BaggingClassifier	estimator=MLPClassifier(activation=tanh, hidden_layer_sizes=(30, 30), learning_rate_init=0.01 max_iter=1000, solver=sgd , max_features: 1.0, max_samples: 0.2, n_estimators: 30, n_estimators=10	0.39 \pm 0.02	0.37 \pm 0.0
BaggingClassifier	estimator=MLPClassifier(activation=tanh, hidden_layer_sizes=(30, 30), learning_rate_init=0.01 max_iter=1000, solver=sgd , max_features: 1.0, max_samples: 0.2, n_estimators: 50, n_estimators=10	0.39 \pm 0.01	0.37 \pm 0.0
BaggingClassifier	estimator=MLPClassifier(activation=tanh, hidden_layer_sizes=(30, 30), learning_rate_init=0.01 max_iter=1000, solver=sgd , max_features: 0.8, max_samples: 0.2, n_estimators: 30, n_estimators=10	0.37 \pm 0.01	0.36 \pm 0.0
BaggingClassifier	estimator=MLPClassifier(activation=tanh, hidden_layer_sizes=(30, 30), learning_rate_init=0.01 max_iter=1000, solver=sgd , max_features: 0.8, max_samples: 0.2, n_estimators: 50, n_estimators=10	0.37 \pm 0.01	0.36 \pm 0.0
BaggingClassifier	estimator=MLPClassifier(activation=tanh, hidden_layer_sizes=(30, 30), learning_rate_init=0.01 max_iter=1000, solver=sgd , max_features: 0.8, max_samples: 0.2, n_estimators: 10, n_estimators=10	0.36 \pm 0.0	0.35 \pm 0.0

Table 5: Ensembles Optimization - GradientBoost

Model	Parameters	F1 Train Score	F1 Validation Score
GradientBoostingClassifier	learning_rate: 0.1, max_depth: 5, max_features: 0.5, n_estimators: 50, subsample: 1	0.47 \pm 0.01	0.39 \pm 0.01
GradientBoostingClassifier	learning_rate: 0.1, max_depth: 5, max_features: 0.5, n_estimators: 100, subsample: 1	0.45 \pm 0.05	0.38 \pm 0.03
GradientBoostingClassifier	learning_rate: 0.1, max_depth: 5, max_features: 0.5, n_estimators: 50, subsample: 0.5	0.4 \pm 0.03	0.38 \pm 0.01
GradientBoostingClassifier	learning_rate: 0.1, max_depth: 10, max_features: 0.5, n_estimators: 50, subsample: 1	0.74 \pm 0.01	0.37 \pm 0.03
GradientBoostingClassifier	learning_rate: 0.1, max_depth: 10, max_features: 0.5, n_estimators: 100, subsample: 1	0.77 \pm 0.03	0.37 \pm 0.02
GradientBoostingClassifier	learning_rate: 0.1, max_depth: 5, max_features: 0.5, n_estimators: 100, subsample: 0.5	0.43 \pm 0.01	0.37 \pm 0.01
GradientBoostingClassifier	learning_rate: 0.1, max_depth: 10, max_features: 0.5, n_estimators: 50, subsample: 0.5	0.53 \pm 0.0	0.37 \pm 0.01
GradientBoostingClassifier	learning_rate: 0.1, max_depth: 5, max_features: 1, n_estimators: 100, subsample: 1	0.43 \pm 0.04	0.36 \pm 0.01
GradientBoostingClassifier	learning_rate: 0.1, max_depth: 10, max_features: 1, n_estimators: 50, subsample: 1	0.59 \pm 0.05	0.36 \pm 0.01
GradientBoostingClassifier	learning_rate: 0.1, max_depth: 10, max_features: 1, n_estimators: 100, subsample: 1	0.55 \pm 0.12	0.36 \pm 0.01
GradientBoostingClassifier	learning_rate: 0.1, max_depth: 10, max_features: 1, n_estimators: 100, subsample: 0.5	0.45 \pm 0.0	0.34 \pm 0.02
GradientBoostingClassifier	learning_rate: 0.1, max_depth: 10, max_features: 1, n_estimators: 50, subsample: 0.5	0.42 \pm 0.01	0.34 \pm 0.01
GradientBoostingClassifier	learning_rate: 0.1, max_depth: 5, max_features: 1, n_estimators: 50, subsample: 1	0.37 \pm 0.02	0.33 \pm 0.0
GradientBoostingClassifier	learning_rate: 0.1, max_depth: 10, max_features: 0.5, n_estimators: 100, subsample: 0.5	0.52 \pm 0.05	0.31 \pm 0.06
GradientBoostingClassifier	learning_rate: 0.5, max_depth: 5, max_features: 1, n_estimators: 50, subsample: 1	0.34 \pm 0.03	0.31 \pm 0.04
GradientBoostingClassifier	learning_rate: 0.1, max_depth: 5, max_features: 1, n_estimators: 50, subsample: 0.5	0.35 \pm 0.02	0.31 \pm 0.01
GradientBoostingClassifier	learning_rate: 0.5, max_depth: 10, max_features: 1, n_estimators: 50, subsample: 1	0.42 \pm 0.03	0.29 \pm 0.03
GradientBoostingClassifier	learning_rate: 0.5, max_depth: 5, max_features: 0.5, n_estimators: 50, subsample: 1	0.28 \pm 0.08	0.27 \pm 0.07
GradientBoostingClassifier	learning_rate: 0.5, max_depth: 10, max_features: 1, n_estimators: 50, subsample: 0.5	0.33 \pm 0.01	0.26 \pm 0.03
GradientBoostingClassifier	learning_rate: 0.1, max_depth: 5, max_features: 1, n_estimators: 100, subsample: 0.5	0.27 \pm 0.15	0.25 \pm 0.15
GradientBoostingClassifier	learning_rate: 0.5, max_depth: 10, max_features: 1, n_estimators: 100, subsample: 1	0.33 \pm 0.01	0.25 \pm 0.03
GradientBoostingClassifier	learning_rate: 0.5, max_depth: 5, max_features: 1, n_estimators: 100, subsample: 1	0.27 \pm 0.06	0.24 \pm 0.04
GradientBoostingClassifier	learning_rate: 0.5, max_depth: 5, max_features: 0.5, n_estimators: 50, subsample: 0.5	0.22 \pm 0.12	0.21 \pm 0.11
GradientBoostingClassifier	learning_rate: 0.5, max_depth: 10, max_features: 1, n_estimators: 100, subsample: 0.5	0.26 \pm 0.05	0.2 \pm 0.02
GradientBoostingClassifier	learning_rate: 0.5, max_depth: 5, max_features: 1, n_estimators: 50, subsample: 0.5	0.22 \pm 0.06	0.18 \pm 0.05
GradientBoostingClassifier	learning_rate: 0.5, max_depth: 10, max_features: 0.5, n_estimators: 50, subsample: 1	0.21 \pm 0.04	0.18 \pm 0.01
GradientBoostingClassifier	learning_rate: 0.5, max_depth: 10, max_features: 0.5, n_estimators: 50, subsample: 0.5	0.2 \pm 0.12	0.16 \pm 0.06
GradientBoostingClassifier	learning_rate: 0.5, max_depth: 5, max_features: 1, n_estimators: 100, subsample: 0.5	0.19 \pm 0.08	0.16 \pm 0.04
GradientBoostingClassifier	learning_rate: 0.5, max_depth: 10, max_features: 0.5, n_estimators: 100, subsample: 0.5	0.17 \pm 0.03	0.15 \pm 0.04
GradientBoostingClassifier	learning_rate: 0.5, max_depth: 5, max_features: 0.5, n_estimators: 100, subsample: 1	0.14 \pm 0.11	0.14 \pm 0.1
GradientBoostingClassifier	learning_rate: 0.5, max_depth: 5, max_features: 0.5, n_estimators: 100, subsample: 0.5	0.13 \pm 0.03	0.12 \pm 0.05
GradientBoostingClassifier	learning_rate: 0.5, max_depth: 10, max_features: 0.5, n_estimators: 100, subsample: 1	0.14 \pm 0.03	0.12 \pm 0.01

Table 6: stacking

Base Learners	F1 Validation Score	F1 Train Score
Bagging(MLP)	0.31 ± 0.07	0.33 ± 0.08
Bagging(MLP), GBoost	0.31 ± 0.06	0.33 ± 0.07
Bagging(MLP), NB	0.31 ± 0.04	0.33 ± 0.04
MLP, NB, DT, Logistic Regression	0.30 ± 0.04	0.33 ± 0.03
GBoost, NB	0.30 ± 0.03	0.32 ± 0.04
MLP, Logistic Regression	0.31 ± 0.03	0.32 ± 0.03
MLP	0.29 ± 0.07	0.30 ± 0.08
MLP, DT, NB	0.27 ± 0.04	0.30 ± 0.05
MLP, DT	0.27 ± 0.07	0.29 ± 0.08
GBoost	0.27 ± 0.06	0.29 ± 0.07
MLP, NB	0.28 ± 0.05	0.29 ± 0.05

Table 7: oversampling

Model	Oversampling	Neighbours	F1 Train Score	F1 Validation Score
OneVsRestClassifier	None	-	0.39 ± 0.01	0.37 ± 0.0
MLPClassifier	None	-	0.42 ± 0.0	0.4 ± 0.01
OneVsRestClassifier	SMOTE	5	0.52 ± 0.01	0.37 ± 0.04
OneVsRestClassifier	SMOTE	10	0.52 ± 0.01	0.37 ± 0.03
MLPClassifier	SMOTE	5	0.46 ± 0.01	0.36 ± 0.02
MLPClassifier	SMOTE	10	0.44 ± 0.02	0.34 ± 0.05
OneVsRestClassifier	ADASYN	5	0.52 ± 0.02	0.35 ± 0.03
OneVsRestClassifier	ADASYN	10	0.52 ± 0.03	0.38 ± 0.02
MLPClassifier	ADASYN	5	0.44 ± 0.01	0.33 ± 0.03
MLPClassifier	ADASYN	10	0.43 ± 0.02	0.33 ± 0.06