



Tunability AutoML

HUBERT BUJAKOWSKI
MIKOŁAJ GAŁKOWSKI

Dane

- Klasyfikacja binarna
- Czyste dane

Dataset	Liczba obserwacji	Liczba zmiennych
spambase (44)	4601	58
steel-plates-fault(1504)	1941	34
diabetes(37)	768	9
qsar-biodeg(1494)	1055	107

Algorytmy

Parametr	Wartości
n_estimators	1, 2, ..., 150
max_depth	1, 2, ..., 15
learning_rate	100 wartości z $U \sim (0.01, 1)$
booster	"gbtree", "gblinear", "dart"
gamma	1000 wartości z $U \sim (0.001, 1024)$
subsample	0.1, 0.2, ..., 1
colsample_bytree	10 wartości z $U \sim (0, 1)$
colsample_bylevel	10 wartości z $U \sim (0, 1)$
reg_alpha	1000 wartości z $U \sim (0.001, 1024)$
reg_lambda	1000 wartości z $U \sim (0.001, 1024)$

XGBoost Classifier

Decision Tree Classifier

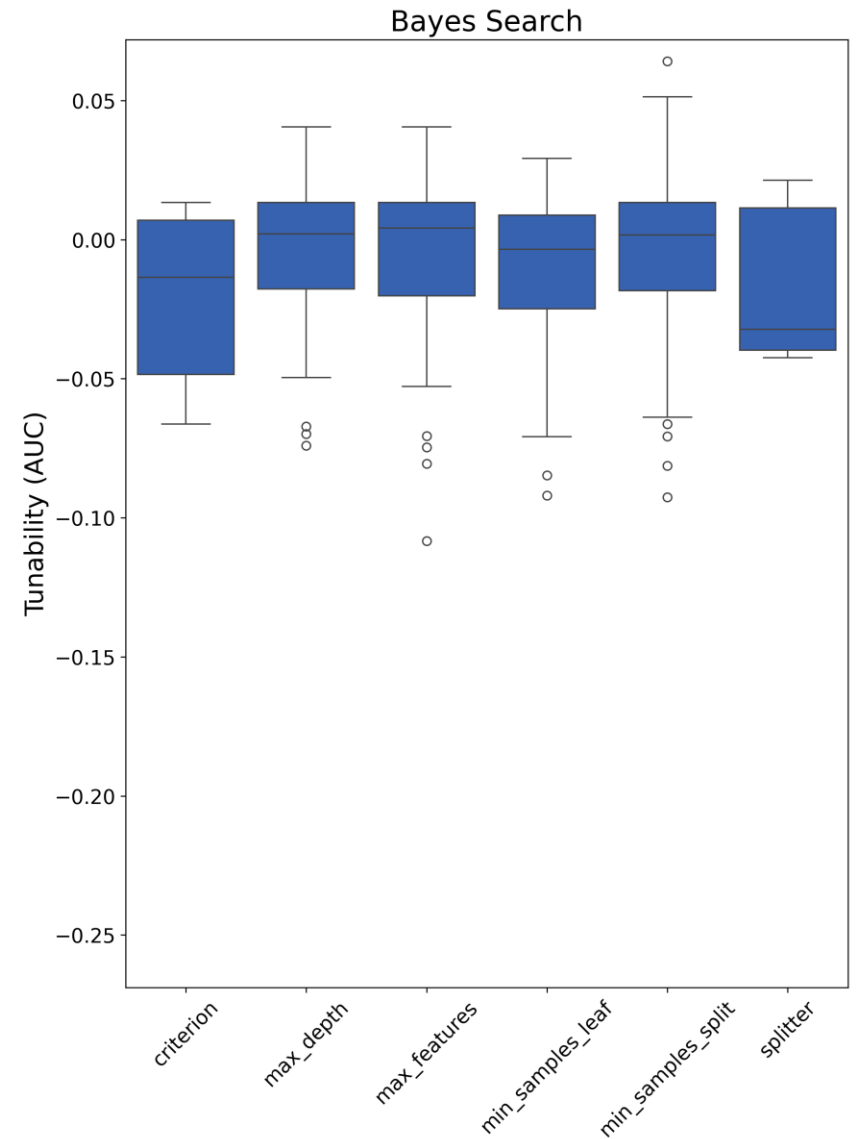
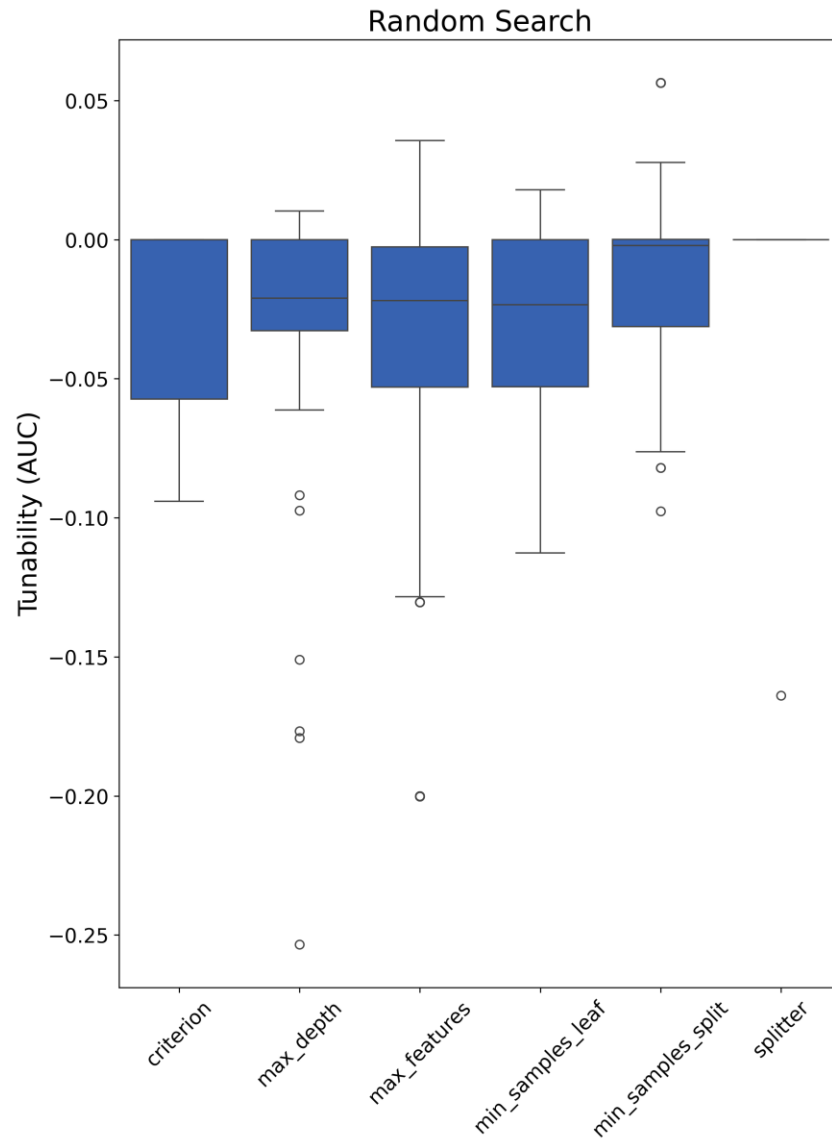
Parametr	Wartości
criterion	"gini", "entropy"
splitter	"best", "random"
max_depth	1, 2, ..., 30
min_samples_split	2, 3, ..., 30
min_samples_leaf	1, 2, ..., 30
max_features	100 wartości z $U \sim (0.01, 1)$

Parametr	Wartości
n_estimators	1, 2, ..., 200
max_depth	10, 11, ..., 50
min_samples_split	2, 3, ..., 20
min_samples_leaf	1, 2, ..., 10
max_features	100 wartości z $U \sim (0.01, 1)$
criterion	"gini", "entropy"
class_weight	None, "balanced"
max_samples	100 wartości z $U \sim (0.01, 1)$

Random Forest Classifier

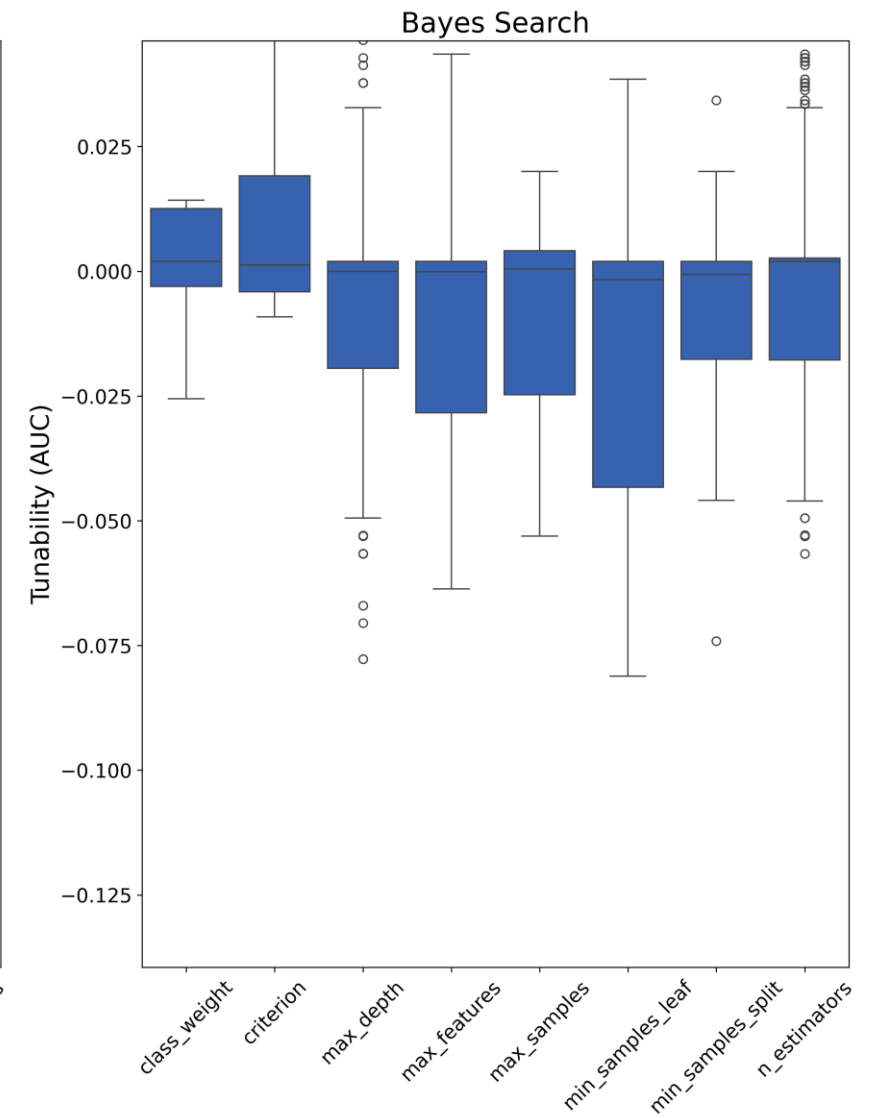
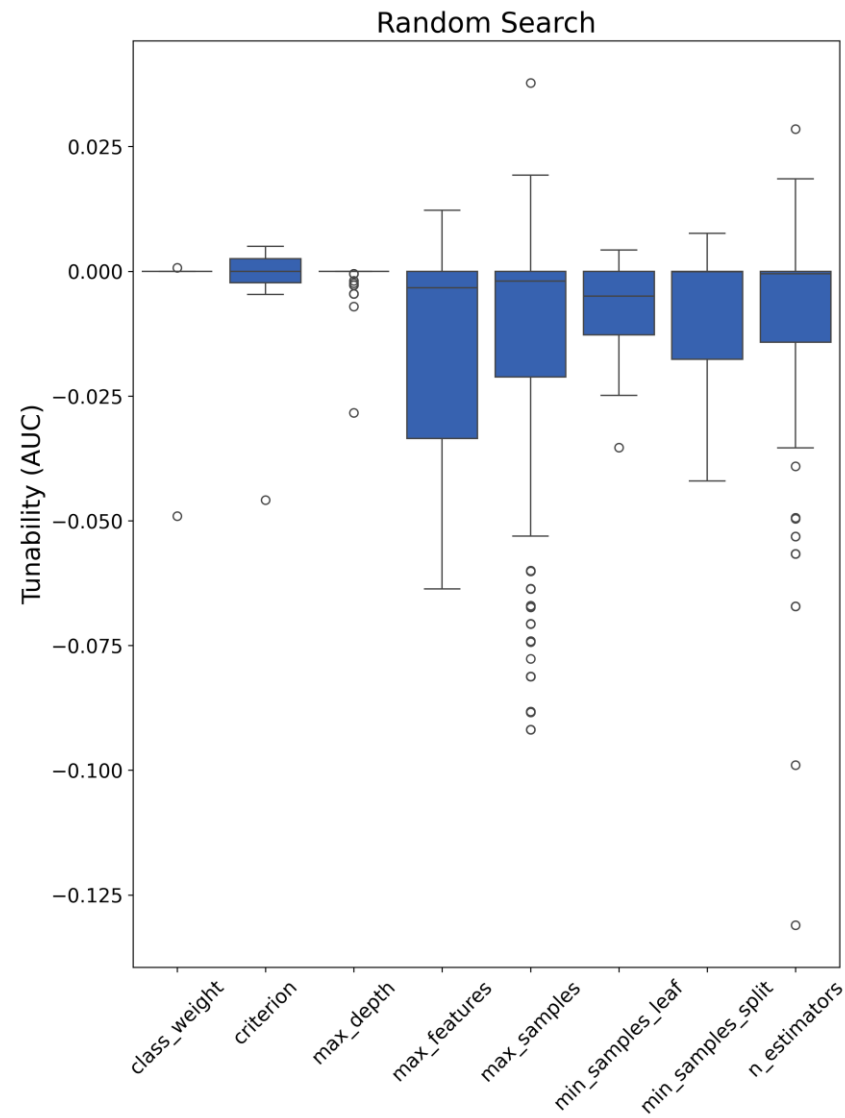
Decision Tree Classifier

Wyniki



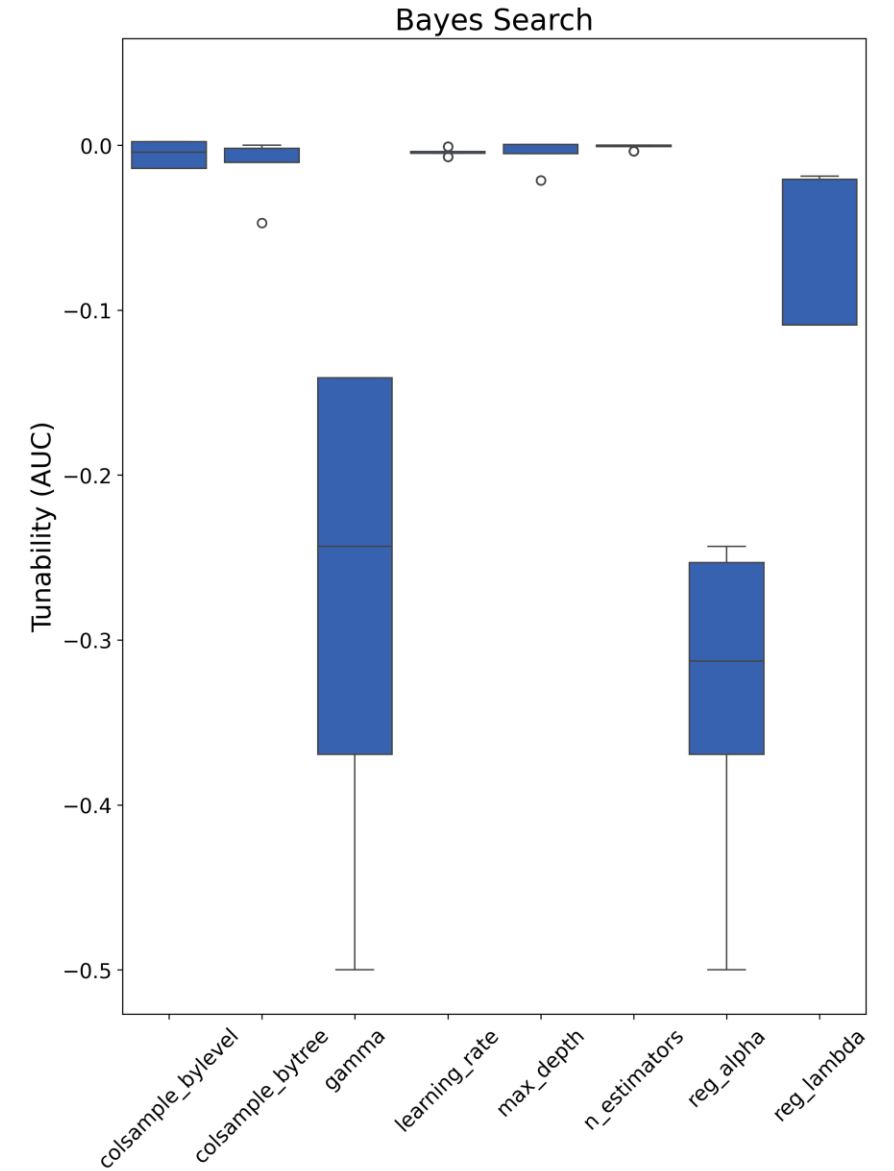
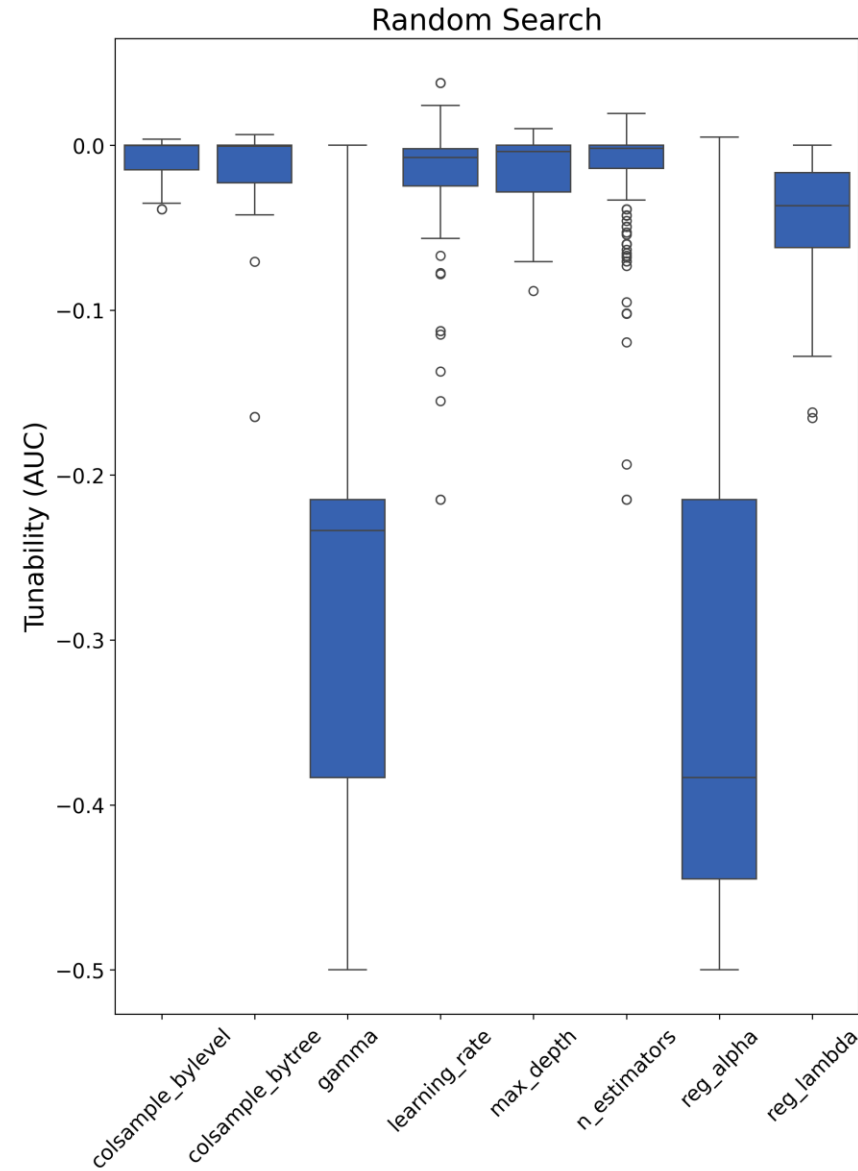
Random Forest Classifier

Wyniki



XGBoost Classifier

Wyniki





Dziękujemy za uwagę!