

AutoML - praca domowa nr 1

Aleksandra Buchowicz
Filip Pazio

Wydział Matematyki i Nauk Informacyjnych

listopad 2023

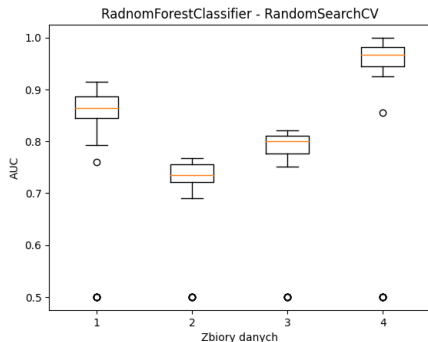
- 1 Random Forest
- 2 XGBoost
- 3 GradientBoosting

parametr	typ	kres dolny	kres górny
bootstrap	logical	-	-
max_features	numeric	0.01	1
min_samples_split	numeric	0.01	1
n_estimators	integer	1	2000

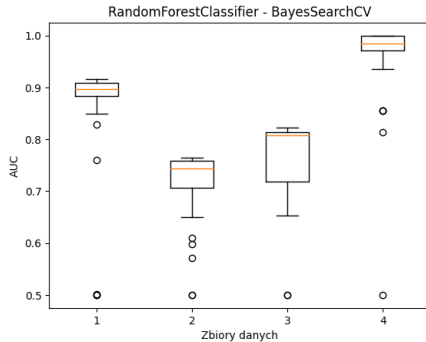
Tabela: Zakresy parametrów dla metody RandomForest.

	parametr	Zbiór 1	Zbiór 2	Zbiór 3	Zbiór 4
RandomizedSearchCV	bootstrap	True	True	False	True
	max_features	0.743	0.934	0.411	0.743
	min_samples_split	0.012	0.08	0.241	0.012
	n_estimators	1894	706	1377	1894
BayesSearchCV	bootstrap	True	True	True	False
	max_features	0.47	0.823	0.258	0.117
	min_samples_split	0.01	0.458	0.041	0.01
	n_estimators	2000	2000	1046	694

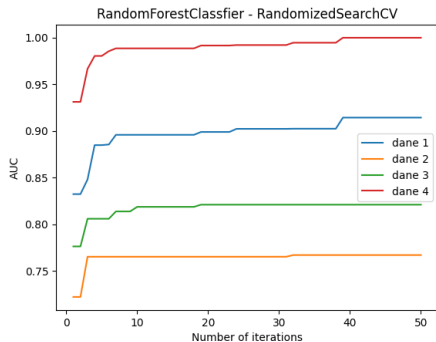
Tabela: Parametry zoptymalizowanych estymatorów.



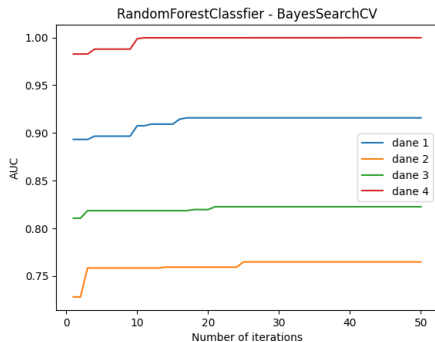
Rysunek: Wykresy skrzynkowe przedstawiające rozłożenie wartości AUC dla metody RandomSearchCV na zbiorach.



Rysunek: Wykresy skrzynkowe przedstawiające rozłożenie wartości AUC dla metody BayesSearchCV na zbiorach.



Rysunek: Liczba iteracji potrzebna RandomSearchCV do osiągnięcia optymalnego rozwiązania.



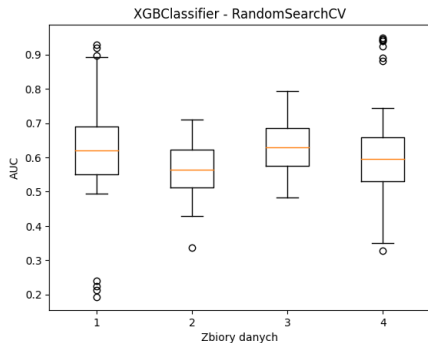
Rysunek: Liczba iteracji potrzebna BayesSearchCV do osiągnięcia optymalnego rozwiązania.

parametr	typ	kres dolny	kres górny
n_estimators	integer	1	5000
learning_rate	numeric	0	10
max_depth	integer	1	15
min_child_weight	numeric	0	7

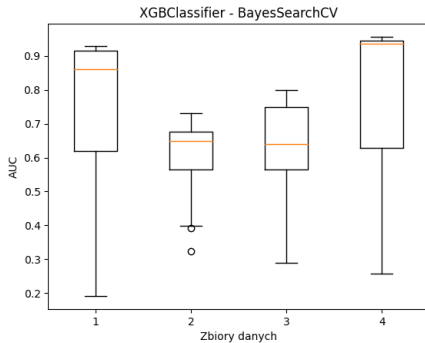
Tabela: Zakresy parametrów dla metody Xgboost.

	parametr	Zbiór 1	Zbiór 2	Zbiór 3	Zbiór 4
RandomizedSearchCV	n_estimators	537	1469	1005	427
	learning_rate	1.463	3.286	2.483	0.387
	max_depth	1	5	11	5
	min_child_weight	0.853	5.639	0.002	0.601
BayesSearchCV	n_estimators	3623	4	2329	4670
	learning_rate	0.806	1.132	0.873	1.009
	max_depth	1	1	15	14
	min_child_weight	0.186	7.000	0.000	0.000

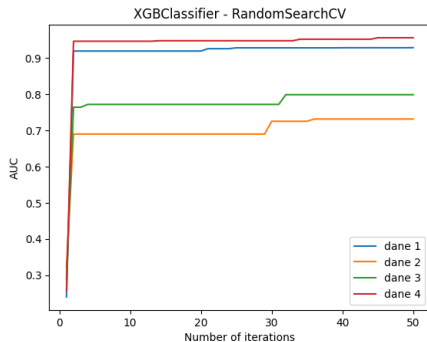
Tabela: Parametry zoptymalizowanych estymatorów.



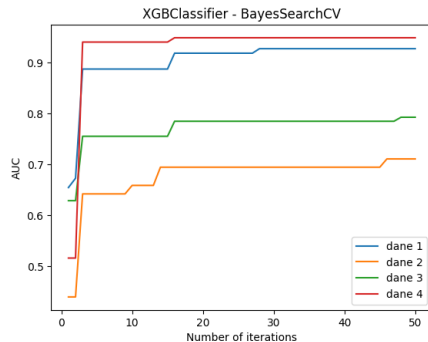
Rysunek: Wykresy skrzynkowe przedstawiające rozłożenie wartości AUC dla metody RandomSearchCV na zbiorach.



Rysunek: Wykresy skrzynkowe przedstawiające rozłożenie wartości AUC dla metody BayesSearchCV na zbiorach.



Rysunek: Liczba iteracji potrzebna RandomSearchCV do osiągnięcia optymalnego rozwiązania.



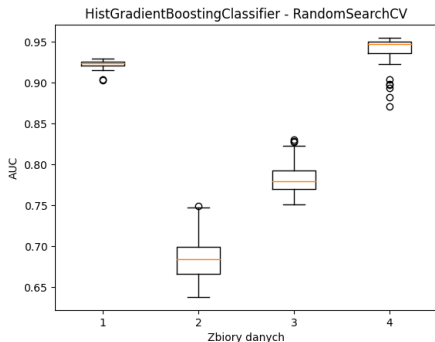
Rysunek: Liczba iteracji potrzebna BayesSearchCV do osiągnięcia optymalnego rozwiązania.

parametr	typ	kres dolny	kres górny
max_iter	integer	1	500
learning_rate	numeric	0.01	1
max_leaf_nodes	integer	2	50
min_samples_leaf	integer	1	50
max_depth	integer	1	15

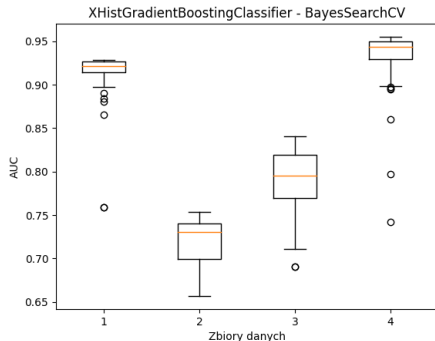
Tabela: Zakresy parametrów dla metody GradientBoosting.

	parametr	Zbiór 1	Zbiór 2	Zbiór 3	Zbiór 4
RandomizedSearchCV	max_iter	438	191	42	386
	learning_rate	0.087	0.043	0.081	0.555
	max_leaf_nodes	16	21	22	45
	min_samples_leaf	18	36	30	6
	max_depth	5	1	5	12
BayesSearchCV	max_iter	377	500	183	179
	learning_rate	0.097	0.010	0.097	0.423
	max_leaf_nodes	34	50	2	39
	min_samples_leaf	50	19	2	23
	max_depth	9	1	1	15

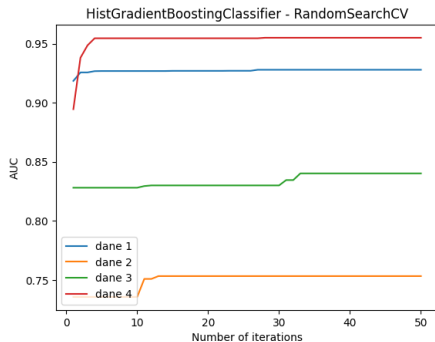
Tabela: Parametry zoptymalizowanych estymatorów.



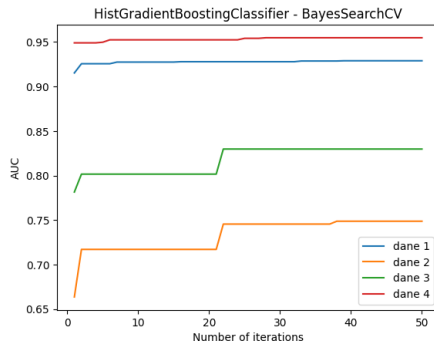
Rysunek: Wykresy skrzynkowe przedstawiające rozłożenie wartości AUC dla metody RandomSearchCV na zbiorach.



Rysunek: Wykresy skrzynkowe przedstawiające rozłożenie wartości AUC dla metody BayesSearchCV na zbiorach.



Rysunek: Liczba iteracji potrzebna RandomSearchCV do osiągnięcia optymalnego rozwiązania.



Rysunek: Liczba iteracji potrzebna BayesSearchCV do osiągnięcia optymalnego rozwiązania.

Dziękujemy za uwagę