

Tunowalność algorytmów uczenia maszynowego

Klaudia Kwoka, Pola Mościcka, Maciej Wach

Agenda

1. Cel projektu
2. Zbiory danych
3. Algorytmy i zakresy hiperparametrów
4. Wybrane wyniki eksperymentów

Cel projektu

- Porównanie Grid Search, Random Search oraz Bayes Optimisation
- Sprawdzenie tunowalności algorytmów
- Przeanalizowanie stabilizacji wyników optymalizacji
- Zbadanie czy występuje bias sampling

Zbiory danych

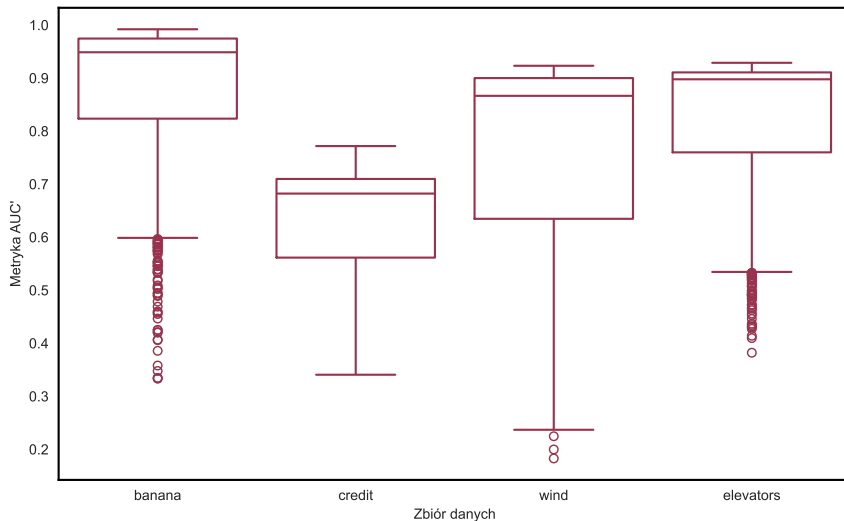
Zbiór danych	l. obserwacji	l. zmiennych	stosunek klas 1:0	źródło
banana	8000	7	50:50	kaggle
credit	1548	18	11:89	kaggle
wind	6574	14	53:47	openml
elevators	5000	18	69:31	openml

Algorytmy i zakresy hiperparametrów

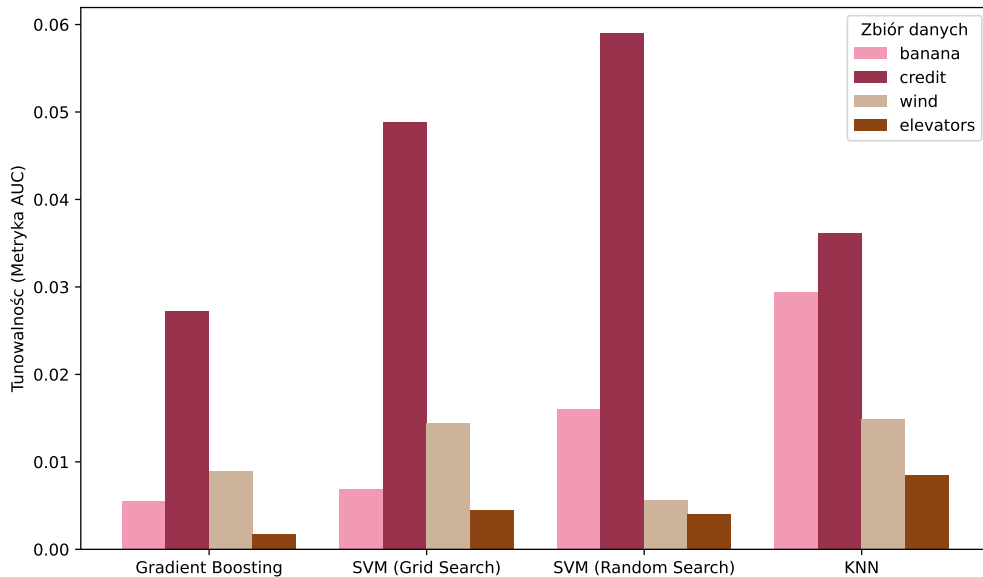
Algorytm	Hiperparametr	Typ	Lower	Upper
Gradient Boosting	n_estimators	int	100	500
	learning_rate	float	0.001	1
	subsample	float	0.1	1
	max_depth	int	4	10
KNN	n_neighbors	int	1	45
	p	float	1	2
SVM	C	float	0.001	1000
	gamma	float	0.001	1000

Rozkład wartości miary AUC podczas krosvalidacji

Algorytm Gradient Boosting

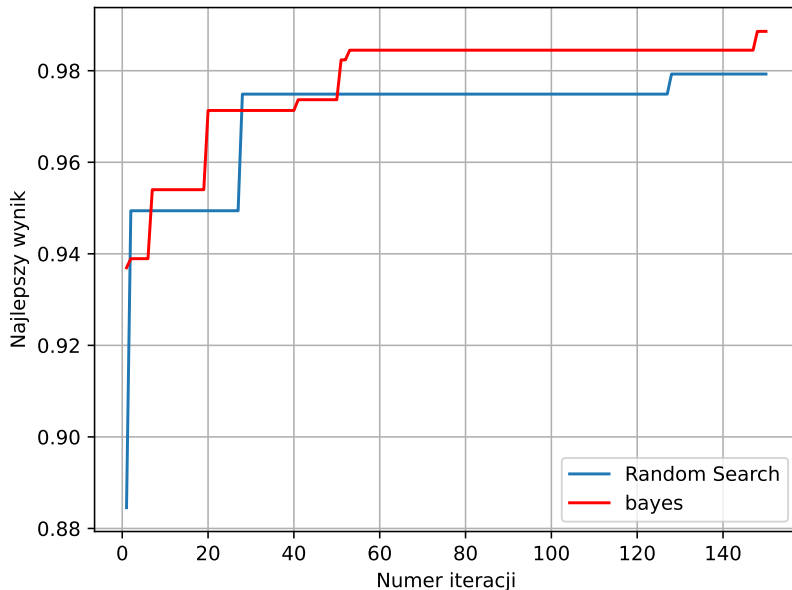


Tunowalność algorytmów



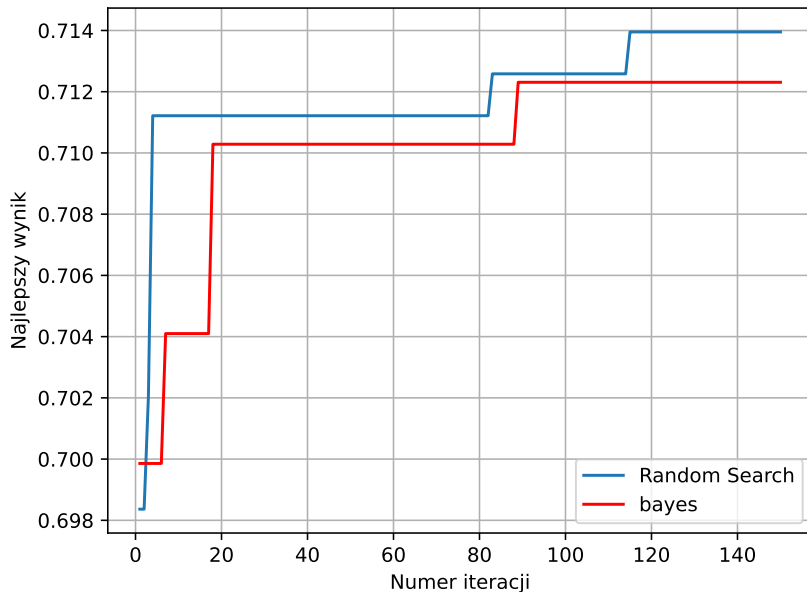
Iteracje Random Search oraz Bayes Optimization

Algorytm KNN, Zbiór 'banana'



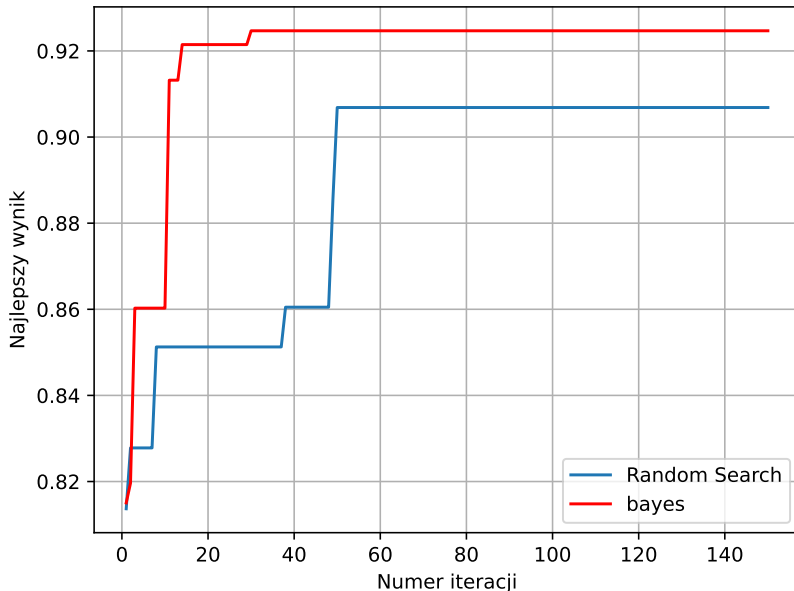
Iteracje Random Search oraz Bayes Optimization

Algorytm KNN, Zbiór 'credit'



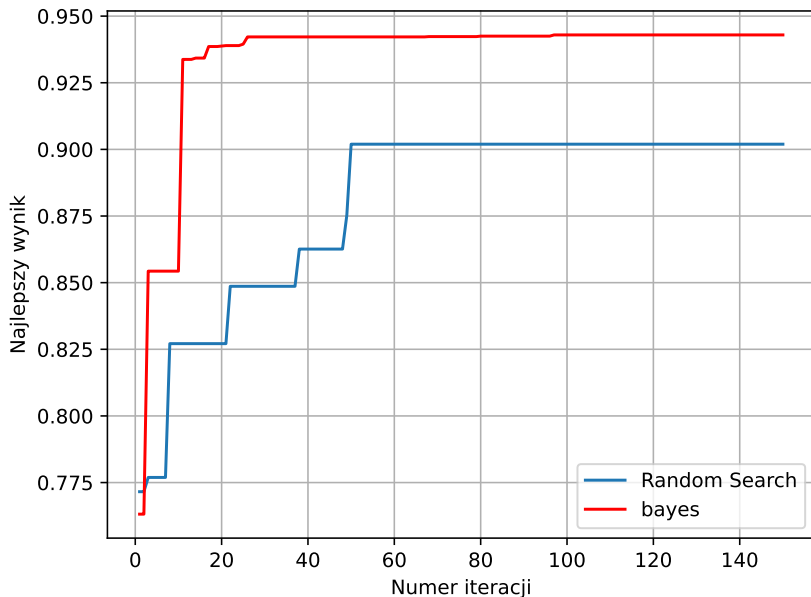
Iteracje Random Search oraz Bayes Optimization

Algorytm KNN, Zbiór 'wind'



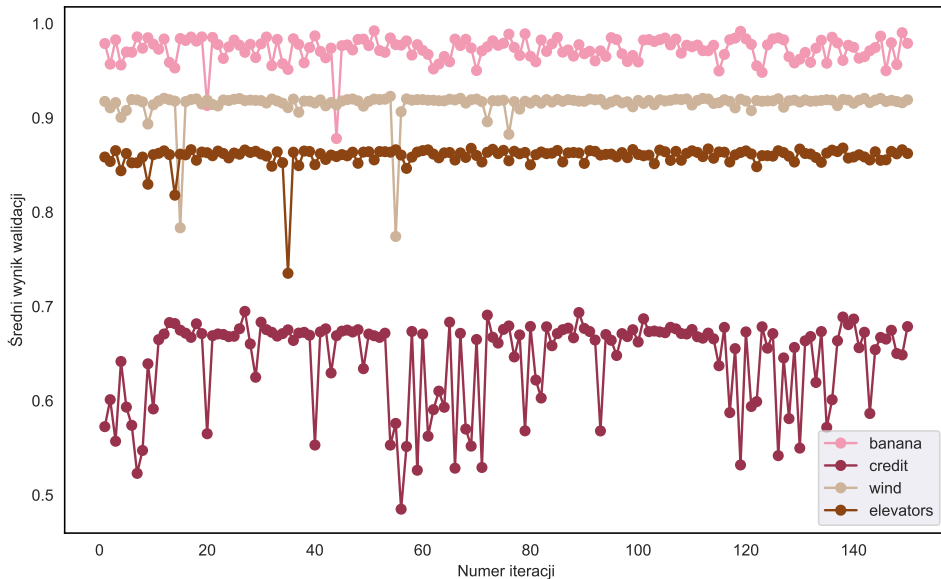
Iteracje Random Search oraz Bayes Optimization

Algorytm KNN, Zbiór 'elevators'



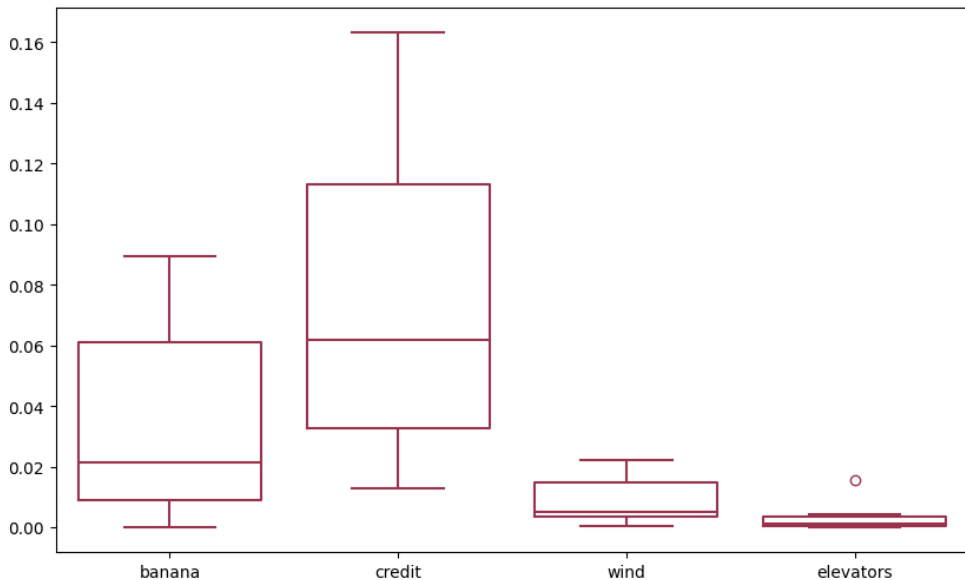
Średni wynik walidacji metryką AUC

Kolejne iteracje Bayes Optimisation dla algorytmu KNN



Występowanie bias sampling

Rozkład tunowalności uzyskanej dla algorytmu SVM (Grid Search) na różnych podziałach z krosvalidacji



Dziękujemy za uwagę