

# Projekt 1

## Automatyczne uczenie maszynowe

Julita Kulesza, Agata Osmałek

Wydział Matematyki i Nauk Informatycznych

22.11.2023

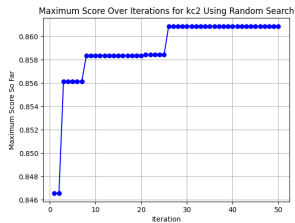
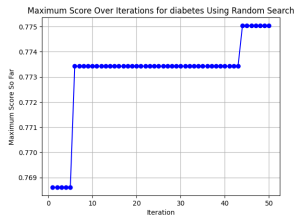
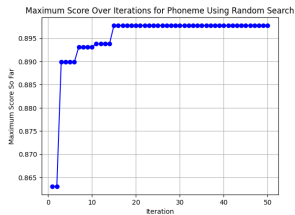
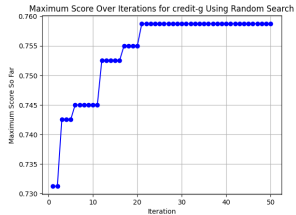
Celem naszego projektu było przeanalizowanie tunowalności - wpływu doboru hiperparametrów na wynik algorytmu - dla trzech wybranych algorytmów uczenia maszynowego, tj.

- GradientBoostingClassifier (GBC),
- RandomForestClassifier (RF),
- Support Vector Classification (SVC).

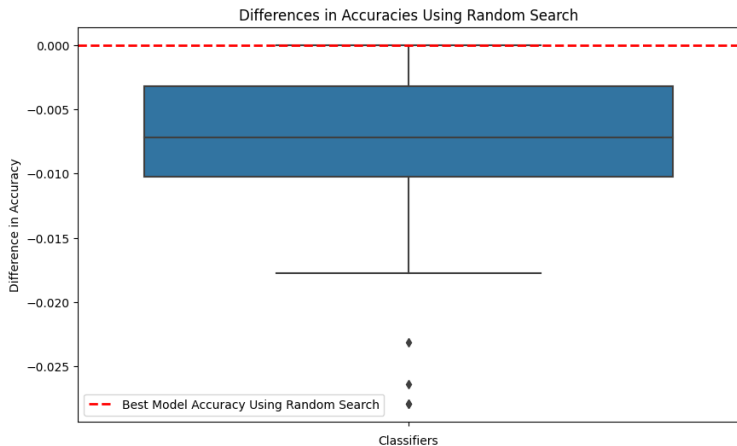
Dlaczego akurat taki wybór? **Przed wszystkim duża liczba hiperparametrów** (zwłaszcza w przypadku GBC oraz RFC).

Każdy z tych algorytmów został wykorzystany w problemie klasyfikacji binarnej na zbiorach z platformy openml o podobnej wielkości:

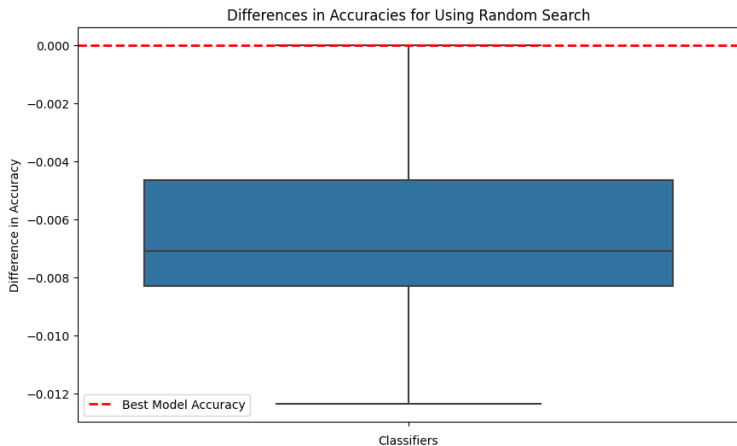
- *credit-g*,
- *Phoneme*,
- *diabetes*,
- *kc2*.



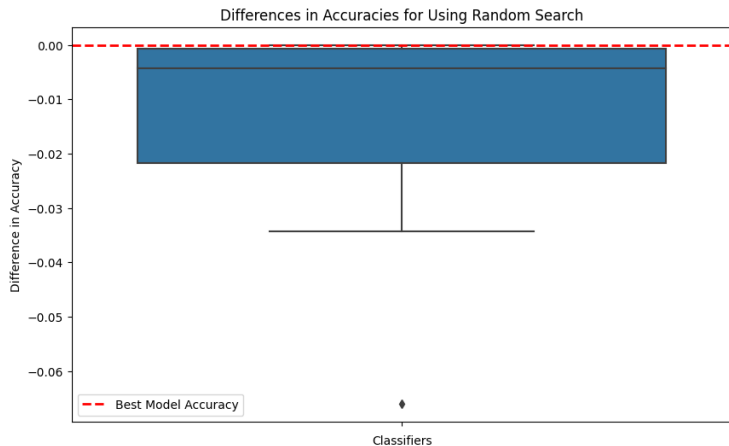
**Rysunek:** Ile iteracji potrzebujemy żeby uzyskać stabilne wyniki optymalizacji przy użyciu RandomizedSearchCV dla klasyfikatora GBC



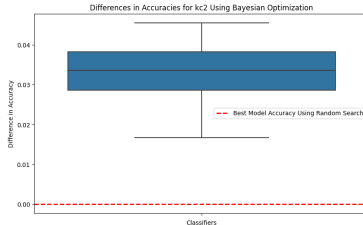
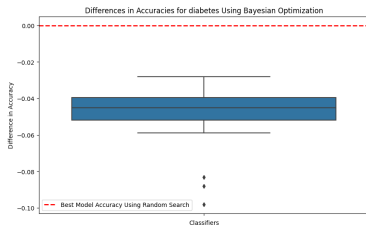
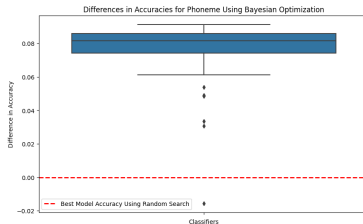
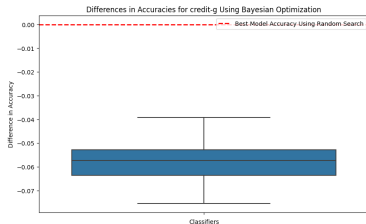
**Rysunek:** Różnice w średniej dokładności dla  $\theta^*$  wybranej przy pomocy RandomizedSearchCV oraz pozostałych średnich dokładności dla algorytmu GBC



**Rysunek:** Różnice w średniej dokładności dla  $\theta^*$  wybranej przy pomocy RandomizedSearchCV oraz pozostałych średnich dokładności dla algorytmu Random Forest

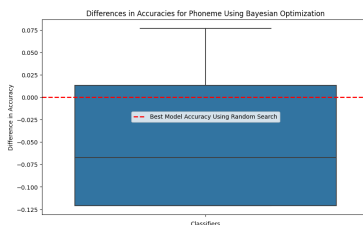
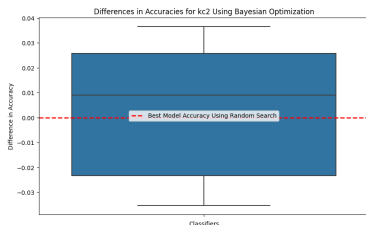
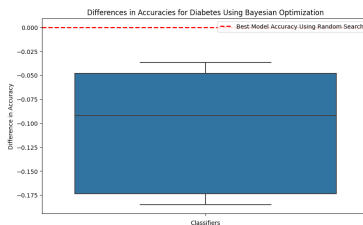
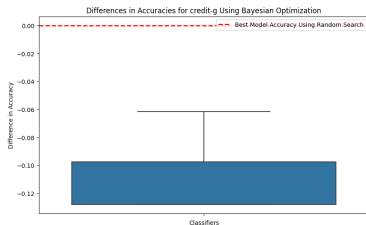


**Rysunek:** Różnice w średniej dokładności dla  $\theta^*$  wybranej przy pomocy RandomizedSearchCV oraz pozostałych średnich dokładności dla algorytmu SVC



**Rysunek:** Różnice w dokładności dla  $\theta^*$  i kolejnych iteracji BayesSearchCVSearchCV dla klasyfikatora GBC

# BayesSearchCV



**Rysunek:** Różnice w dokładności dla  $\theta^*$  i kolejnych iteracji BayesSearchCV dla klasyfikatora RF



# Projekt 1

## Automatyczne uczenie maszynowe

Julita Kulesza, Agata Osmałek

Wydział Matematyki i Nauk Informatycznych

22.11.2023