



Wstęp do uczenia maszynowego - projekt

Oliwia Trzcińska

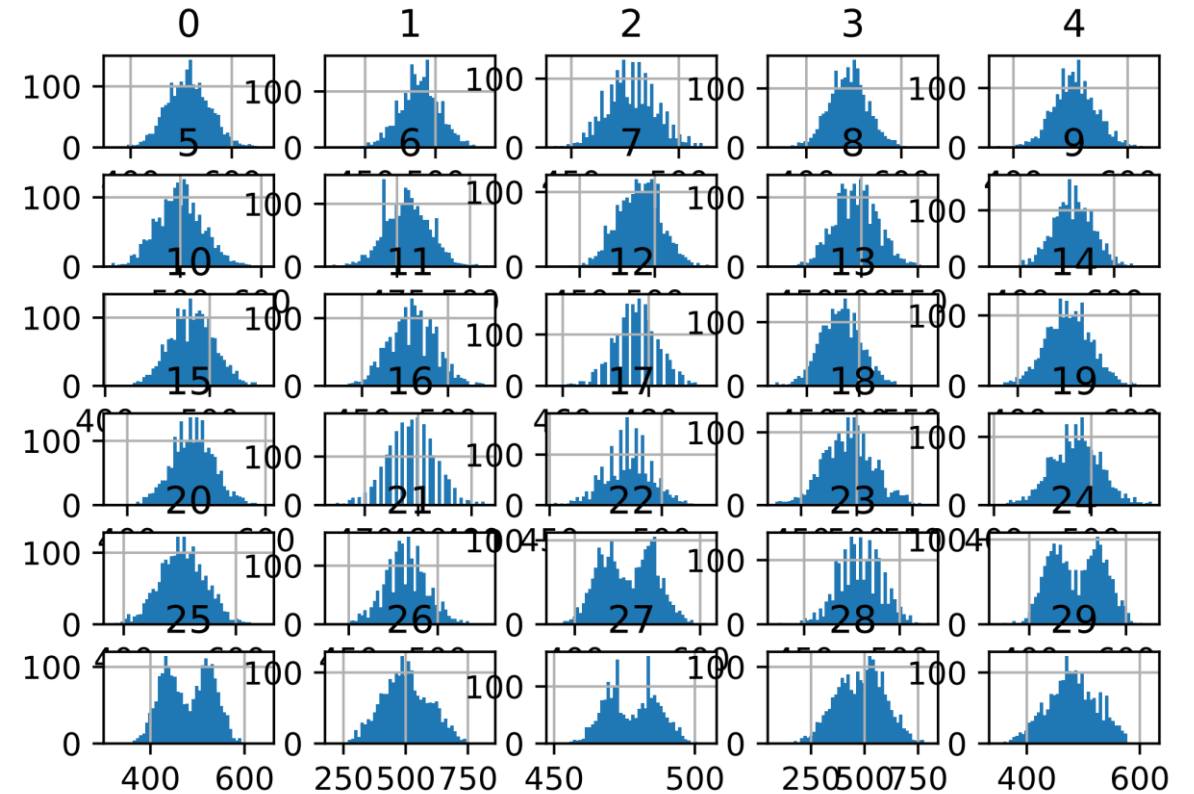


Cel projektu

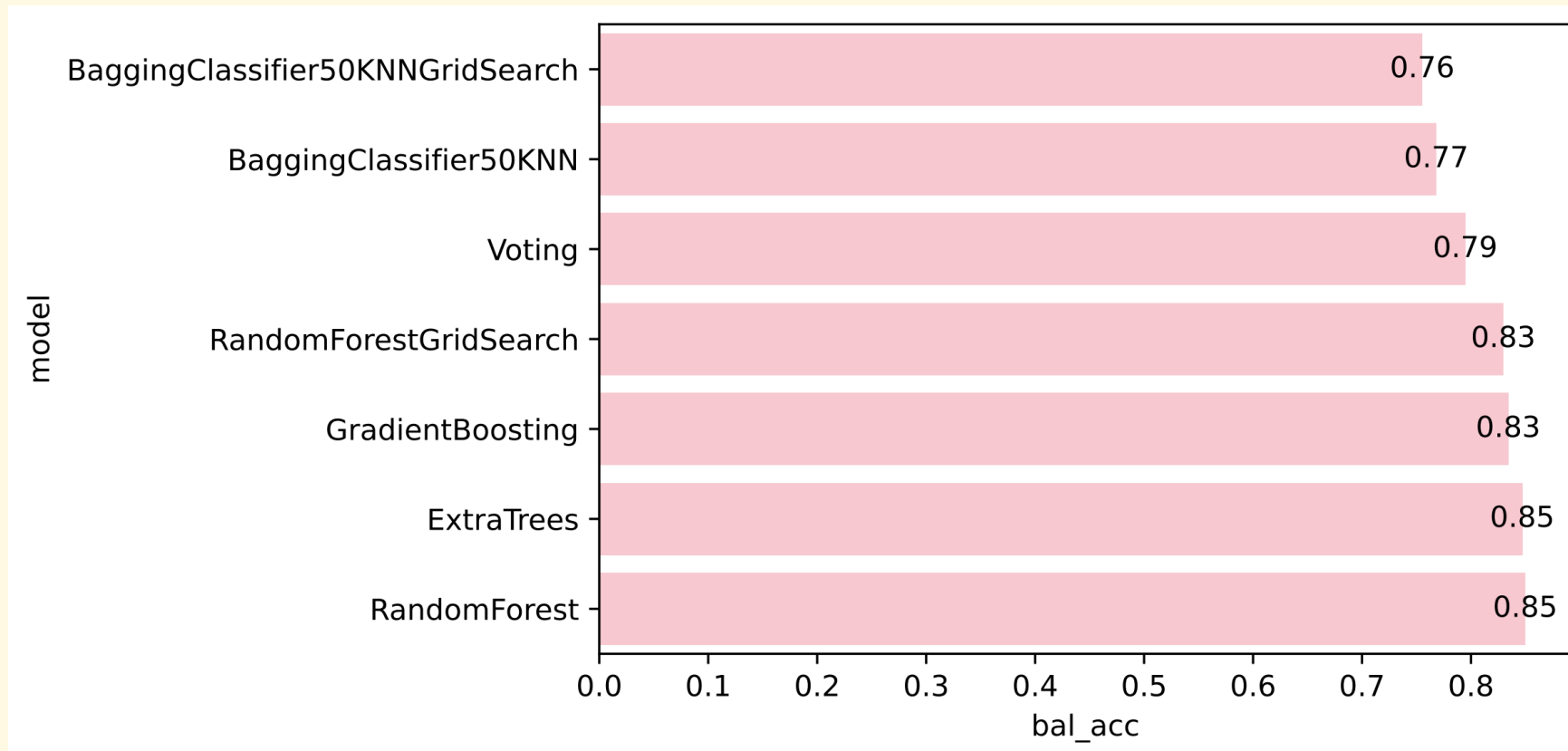
- Stworzenie modelu klasyfikacji o jak największej mocy predykcyjnej. Będziemy zajmować się danymi ze sztucznie wygenerowanego zbioru ,artifical’.

Dane

- Braki danych, obserwacje odstające, rozkłady zmiennych
- Podział na zbiory treningowy i testowy



Wybór modelu



Model ExtraTrees

Najlepsza wartość zrównoważonej dokładności na części testowej: 0,893 (8 istotnych zmiennych)



Model RandomForest

- Najlepsza wartość zrównoważonej dokładności na części testowej: 0,902 (8 istotnych zmiennych)



Wynik

```
forest_all = RandomForestClassifier(random_state = 320584)
forest_all.fit(X_train, y_train["0"])
sfm2 = SelectFromModel(forest_all)
X_train_sel = sfm2.fit_transform(X_train, y_train["0"])
forest_all.fit(X_train_sel, y_train["0"])
y_train_pred = forest_all.predict(X_train_sel)
```


A serene forest landscape featuring a small stream that flows over large, moss-covered rocks. The water is clear and creates small white rapids as it moves through the forest. The surrounding trees are tall and slender, with dense green foliage that filters the sunlight, creating a dappled light effect on the forest floor. The overall atmosphere is peaceful and natural.

Podsumowanie