SPRAWOZDANIE

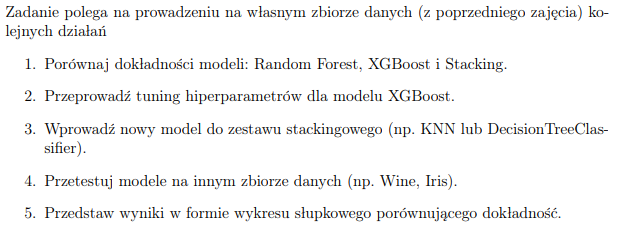
Zajęcia: Nauka o danych II

Prowadzący: prof. dr hab. Vasyl Martsenyuk

|  |  |
| --- | --- |
| Laboratorium Nr 2  Data 15.03.2025  Temat: „Implementacja zaawansowanych modeli klasyfikacji danych w Python”  Wariant 2 | Anna Więzik  Informatyka  II stopień, niestacjonarne,  2 semestr, gr.1a TTO |

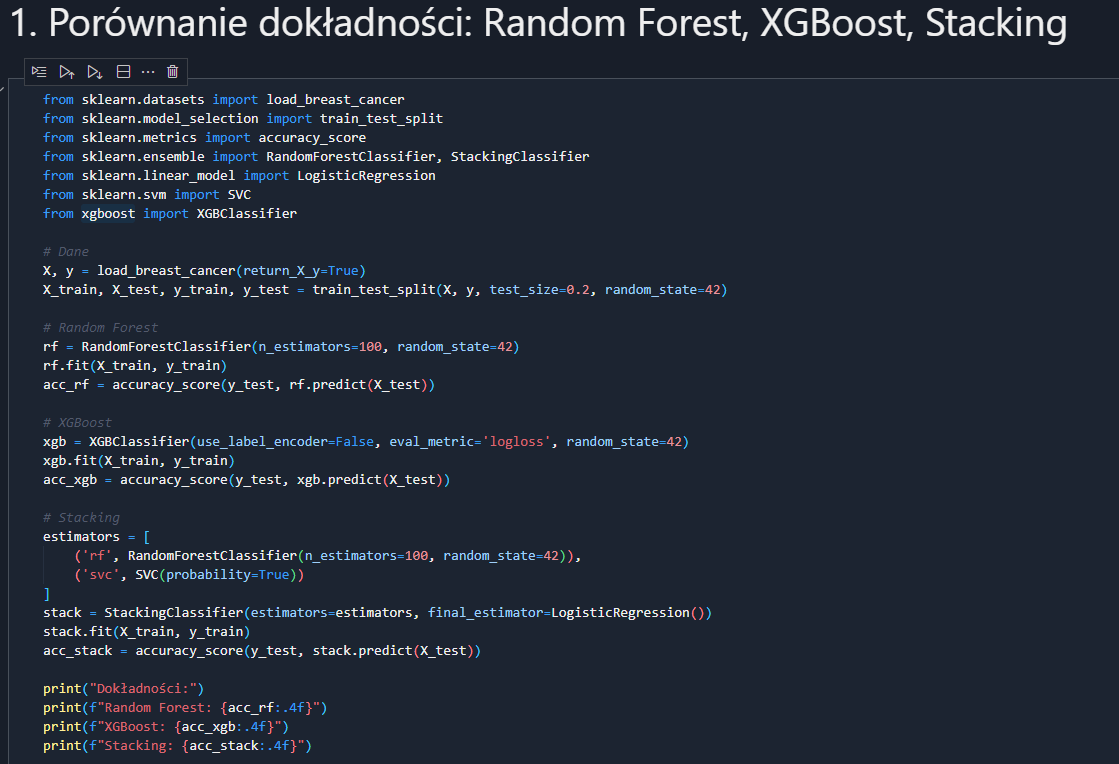
1. Polecenie:

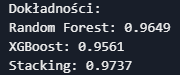
Link do repozytorium: <https://github.com/AnaShiro/NoD2_2025>

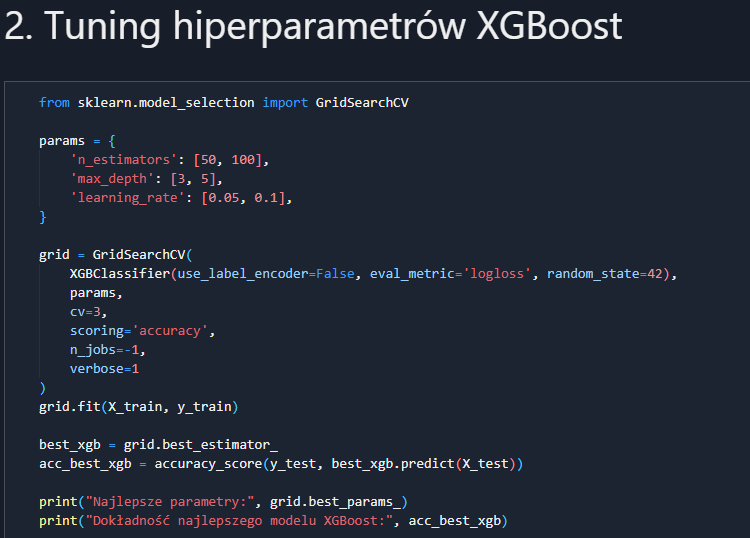


Uczenie zespołowe (ensemble learning) to podejście w uczeniu maszynowym polegające na łączeniu wielu modeli bazowych w celu stworzenia silniejszego modelu predykcyjnego. Główne techniki to bagging, boosting i stacking. Bagging redukuje wariancję modelu, boosting zmniejsza zarówno błąd systematyczny, jak i wariancję, a stacking pozwala łączyć różne algorytmy, tworząc model meta, który dokonuje końcowej predykcji. Celem tych metod jest poprawa dokładności oraz odporności na przeuczenie.

1. Opis programu opracowanego



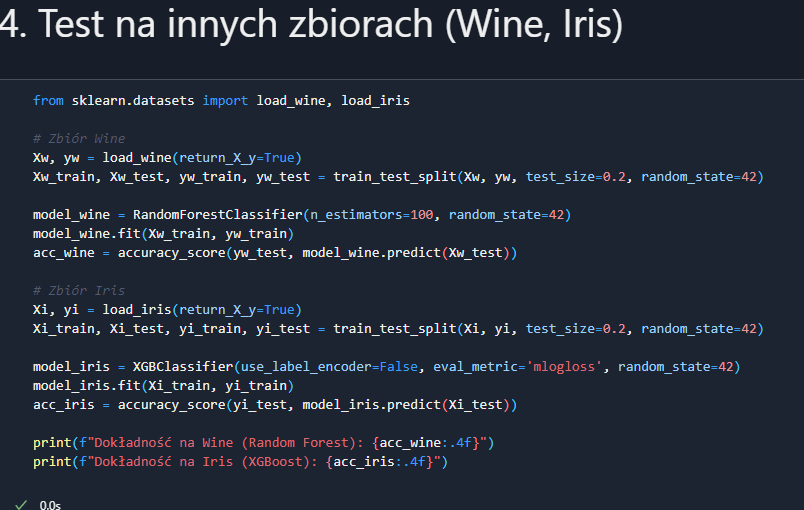




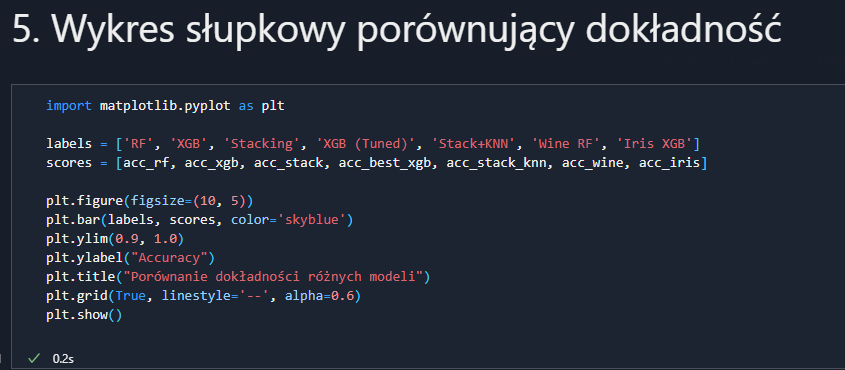


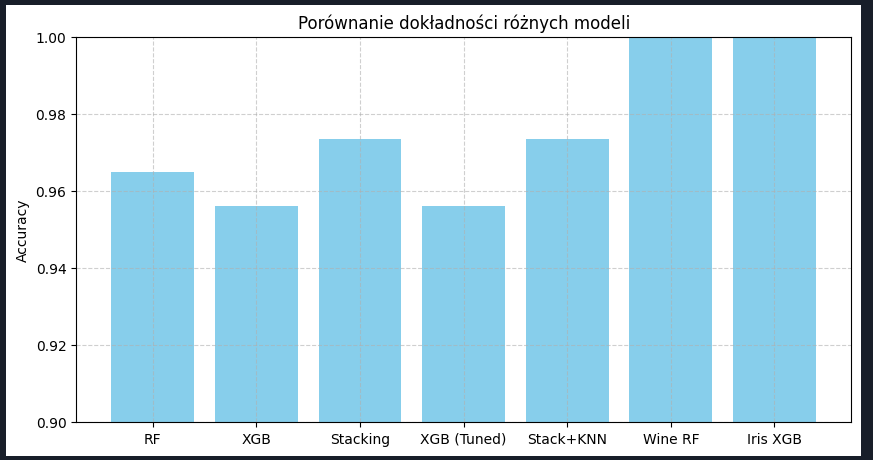












1. Wnioski

W eksperymencie porównano modele Random Forest, XGBoost i StackingClassifier. Najwyższą dokładność osiągnął stacking, co potwierdza skuteczność łączenia różnych algorytmów. Random Forest i XGBoost również uzyskały wysokie wyniki, przy czym XGBoost daje dodatkowe możliwości tuningu. Analiza ważności cech wykazała, które zmienne mają największy wpływ na predykcje. Ostatecznie, techniki ensemble znacząco poprawiają jakość modeli i warto je stosować w praktycznych zadaniach klasyfikacyjnych.