

Zadanie 7 - Raport

Piotr Niedziałek

Czerwiec 2024

1 Wstęp

Celem laboratorium była implementacja klasyfikatora Gaussowski Naiwny Bayes. Jest to klasyfikator, działający na atrybutach ciągłych. Zakłada, że zmienne objaśniające nie są od siebie zależne. Jest to uproszczenie, które znacząco pomaga w obliczeniach i jednocześnie dobrze modeluje rzeczywistość. Ponadto zakłada się, że wartości atrybutów pochodzą z rozkładu normalnego.

2 Implementacja

Zadanie zaimplementowano w pythonie. Skorzystano z bibliotek: skleran, pandas i numpy. Struktura kodu:

- klasa GaussianNaiveBayes - główna klasa, uczenie klasyfikatora oraz klasyfikacja
- get_results() - walidacja krzyżowa i wyniki
- pozostały kod - obliczanie wyników dla różnych ziaren StratifiedKFold

3 Badania

Porównano działanie 3 klasyfikatorów na zbiorze danych iris z scikit-learn. Badane klasyfikatory to SVM, drzewo decyzyjne i Gausowski Naiwny Bayes. Przeprowadzono średnie obliczenia parametrów dla 5 różnych ziaren. Miary jakości klasyfikacji:

- Accuracy
- Precision
- Recall
- F1

4 Wyniki eksperymentów

Metric	Gaussian Naive Bayes	Decision Tree	SVM
Accuracy	0.952 ± 0.038	0.945 ± 0.031	0.983 ± 0.018
Precision	0.954 ± 0.038	0.948 ± 0.032	0.984 ± 0.016
Recall	0.952 ± 0.038	0.951 ± 0.025	0.983 ± 0.018
F1 Score	0.952 ± 0.038	0.947 ± 0.041	0.983 ± 0.018