WSI sprawozdanie zadanie 4

Wstęp

W uczeniu maszynowym, Support Vector Machines (SVM) oraz drzewa decyzyjne są jednymi z najbardziej znanych i szeroko stosowanych algorytmów klasyfikacyjnych.

SVM to narzędzie służące do klasyfikacji, szczególnie efektywne w przypadksch wysokiej wymiarowości i przy małych zestawach danych. SVM działa na zasadzie znalezienia optymalnej hiperpłaszczyzny, która maksymalizuje margines między dwiema klasami danych. Margines ten to odległość między najbliższymi punktami danych (zwanymi wektorami wspierającymi) z każdej klasy a hiperpłaszczyzną.

Drzewa decyzyjne do metoda klasyfikacji, która używa struktury drzewa do modelowania decyzji i ich konsekwencji. Każdy węzeł wewnętrzny reprezentuje test na atrybucie, każda gałąź wynik testu, a każdy liść reprezentuje etykietę klasy.

Opis programu

Całość projektu składa się z 5 skryptów python oraz pliku konfiguracyjnego args.yaml. Po uruchomieniu skryptu main.py ładowane są dane konfiguracyjne a następnie uruchamianych jest 6 eksperymentów. Eksperymenty są zaimplementowane w skrypcie experiments.py. Każdy eksperyment po zakończeniu zapisuje wyniki precision, accuracy, recall oraz f1 wraz z odchyleniami standardowymi w plikach csv. W skrypcie plot\_experiments.py zaimplementowane są eksperymenty zapisujące wyniki precision dla 50 wartości takich parametrów jak siła regularyzacji czy głębokość drzewa decyzyjnego. Skrypt plot\_results.py odpowiedzialny jest za utworzenie wykresów z danych zapisanych przez plot\_experiments.py.

Eksperymenty:

Pierwszy eksperyment badano zależność skuteczności SVM od wartości parametru regularyzacji. Miarą skuteczności były wskaźniki precision, accuracy, recall oraz f1. Eksperyment był uruchamiany dla jądra linear.

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| C | precision | precision std | accuracy | accuracy std | recall | recall std | f1 | f1 std |
| 0,1 | 0,976 | 0,012 | 0,973 | 0,013 | 0,973 | 0,013 | 0,973 | 0,013 |
| 1 | 0,979 | 0,014 | 0,977 | 0,015 | 0,977 | 0,015 | 0,977 | 0,015 |
| 10 | 0,979 | 0,021 | 0,976 | 0,026 | 0,976 | 0,026 | 0,975 | 0,026 |
| 100 | 0,977 | 0,025 | 0,973 | 0,031 | 0,973 | 0,031 | 0,973 | 0,031 |
| 1000 | 0,976 | 0,028 | 0,972 | 0,034 | 0,972 | 0,034 | 0,972 | 0,034 |

W drugim eksperymencie badano wpływ funkcji jądra na skuteczność SVM.

W