

MIW Mini-projekt 4 – regresja liniowa

Zadanie:

1. Wczytaj dane z pliku *daneXX.txt*. Zaproponuj oraz zrealizuj podział danych na treningowe i testowe. XX to numer studenta na liście. Lista jest dostępna na ftp'ie prowadzącego w pliku z ocenami.
2. (4 punkty) Zaproponuj liniowy model parametryczny. Określ jego parametry stosując metodę najmniejszych kwadratów dla danych treningowych. Zweryfikuj poprawność modelu za pomocą danych testowych.
3. (4 punkty) Zaproponuj bardziej złożony model parametryczny. Określ jego parametry stosując metodę najmniejszych kwadratów dla danych treningowych. Zweryfikuj poprawność modelu za pomocą danych testowych.
4. (2 punkty) Porównaj oba modele: graficznie oraz statystycznie.

Uwaga (1): kod musi być napisany posługując się wyłącznie językiem angielskim, komentarze piszemy w języku polskim

Uwaga (2): za nieczytelny kod będę przyznawał -1 punkt – zgodnie z poprzednimi wytycznymi; niechlujne prace (wszystko poza kodem) zasłużą na dodatkowy -1 punkt

Uwaga (3): zabronione jest korzystanie z gotowych elementów/modułów/bibliotek realizujących zadanie regresji (0 punktów za mini-projekt)

Uwaga (4): wypisz równania, macierze, wektory oraz obliczone parametry modeli (do -4 punktów)

Forma:

1. Komentowanie kodu jest na ogół złą praktyką. Jednak w tym przypadku wynikowy kod wraz z wartościami wyjściowymi wypisanymi np. na konsolę stanowi dokumentację naszej pracy. Kluczowe elementy kodu i wartości wynikowe proszę skrótowo, aczkolwiek dostatecznie jasno, opisywać i komentować.
2. Mini-projekt można oddać jedynie w formie pliku ipynb wraz z jego wyeksportowaną wersją do html lub pdf. W sumie każdy student wysyła dokładnie dwa pliki do jednego sprawozdania. Proszę pamiętać o wszystkich wartościach wynikowych, czyli kod ipynb musi zostać wykonany przed eksportem.
3. Sprawozdanie należy wysłać na adres e-mail pprzestrzelski@pjawst.edu.pl:
 - a. nazwy plików: *NR_INDEKSU_Imię_Nazwisko.ext*, gdzie NR_INDEKSU to np. S999, a ext to rozszerzenie pliku,
 - b. tytuł maila: *[MIW] Mini-projekt 4*.

Termin:

1. 27.04.2020 r., godzina 15:45 (100% punktów)
2. 04.05.2020 r., godzina 15:45 (50% punktów)