

Komputerowa analiza szeregów czasowych – raport I

Analiza danych z wykorzystaniem regresji liniowej

Deadline: 22 grudnia 2023 (piątek), 23:59

Raport można przygotować samodzielnie lub w dwuosobowych grupach.

1. Znalezienie zbioru danych rzeczywistych (min. 100 obserwacji) złożonego z dwóch zmiennych.
2. Opis danych (czego dotyczą, liczba obserwacji, z jakiego okresu, **źródło** itp.), wykres danych.
3. Statystyki opisowe (średnia, wariancja, itp.) dla każdej ze zmiennych.
4. Dobranie prostej regresji dla danych (jedna zmienna: x , druga zmienna: y), wyznaczenie R^2 .
5. Wyznaczenie przedziałów ufności dla parametrów β_0, β_1 .
6. Analiza residuów: czy układają się losowo wokół zera, czy mają stałą wariancję, czy mają rozkład normalny, czy występują obserwacje odstające (nieobowiązkowo: zbadać czy residua są nieskorelowane przy użyciu funkcji autokorelacji)
7. Predykcje: prognoza przyszłej wartości $Y(x_0)$ lub prognoza wartości średniej zmiennej $Y(x_0)$

Dodatkowe informacje:

- Proszę zwrócić szczególną uwagę na wnioski.
- Gdzie szukać danych? Kaggle, Yahoo Finance, Stooq, ...
- W raporcie powinny znaleźć się wszystkie użyte wzory oraz definicje.
- Do każdego rysunku/tabeli powinno być odwołanie w tekście, ponadto wszystkie rysunki/tabele powinny być podpisane.
- Raport w formie pliku PDF proszę przesłać przez ePortal (w przypadku dwuosobowych grup: wystarczy wysłanie przez jedną osobę).
- Za raport można otrzymać maksymalnie 10 punktów.