**Nazwa przedmiotu:**

Metody statystyki obliczeniowej

**Koordynator przedmiotu:**

Dr hab. Przemysław Grzegorzewski, prof. PW, Prof. Jan Mielniczuk

**Status przedmiotu:**

Obowiązkowy

**Poziom kształcenia:**

Studia I stopnia

**Program:**

Inżynieria i Analiza Danych

**Grupa przedmiotów:**

Wspólne

**Kod przedmiotu:**

1120-DS000-ISP-0354

**Semestr nominalny:**

5 / rok ak. 2020/2021

**Liczba punktów ECTS:**

5

**Liczba godzin pracy studenta związanych z osiągnięciem efektów uczenia się:**

1. godziny kontaktowe – 70 h; w tym
a) obecność na wykładach – 30 h
b) obecność na laboratoriach – 30 h
c) obecność na egzaminie – 5 h
d) konsultacje – 5 h
2. praca własna studenta – 55 h; w tym
a) przygotowanie do laboratoriów i do kolokwiów – 20 h
b) rozwiązywanie zadań domowych – 20 h
c) przygotowanie do egzaminu – 15 h
Razem 127h, co odpowiada 5 pkt. ECTS

**Liczba punktów ECTS na zajęciach wymagających bezpośredniego udziału nauczycieli akademickich:**

1. obecność na wykładach – 30 h
2. obecność na laboratoriach – 30 h
3. obecność na egzaminie – 5 h
4. konsultacje – 5h
Razem 70 h, co odpowiada 3 pkt. ECTS

**Język prowadzenia zajęć:**

polski

**Liczba punktów ECTS, którą student uzyskuje w ramach zajęć o charakterze praktycznym:**

1. obecność na laboratoriach – 30 h
2. rozwiązywanie zadań domowych – 20 h
3 przygotowanie do laboratoriów i do kolokwiów – 20 h
Razem 70 h, co odpowiada 3 pkt. ECTS

**Formy zajęć i ich wymiar w semestrze:**

|  |  |
| --- | --- |
| Wykład: | 30h |
| Ćwiczenia: | 0h |
| Laboratorium: | 30h |
| Projekt: | 0h |
| Lekcje komputerowe: | 0h |

**Wymagania wstępne:**

Rachunek Prawdopodobieństwa, Statystyka Matematyczna

**Limit liczby studentów:**

.

**Cel przedmiotu:**

Celem przedmiotu jest zapoznanie studentów z podstawowymi metodami nieparametrycznymi wnioskowania statystycznego oraz z metodami analizy regresji, klasyfikacji i szeregów czasowych

**Treści kształcenia:**

Podstawowe testy nieparametryczne dla pojedynczej próbki i wielu próbek. Nieparametryczne metody oceny stopnia zależności między cechami.
Model regresji liniowej prostej i wielokrotnej: metoda MNK, własności estymatora MNK, podstawowe testy, diagnostyka dopasowania. Metoda Lasso i regresji grzbietowej, regresja metodą składowych głównych.
Model parametryczny regresji nieliniowej i nieliniowa metoda najmniejszych kwadratów. Estymacja nieparametryczna funkcji regresji: estymator średniej ruchomej i lokalnie liniowy.
Modele regresji binarnej regresja logistyczna i probitowa. Estymatory największej wiarogodności w modelu logistycznym.
Klasyfikacja pod nadzorem i liniowe metody klasyfikacyjne: LDA, logistyczna, SVM. Nieliniowe metody klasyfikacyjne: QDA i drzewa klasyfikacyjne.
Podstawowe charakterystyki stacjonarnego szeregu czasowego: funkcja autokowariancji, autokorelacji, dystrybuanta i gęstość spektralna. Modele szeregu czasowego: proces autoregresyjny, średniej ruchomej, ARMA. Problem optymalnej prognozy liniowej i jego rozwiązanie. Dekompozycja nie-stacjonarnego procesu czasowego.

**Metody oceny:**

Podstawą do zaliczenia laboratoriów są wyniki dwóch kolokwiów (max 16 punktów każde) i aktywność na zajęciach (max 8 punktów) = 40 punktów max. Nie przewiduje się kolokwiów poprawkowych. Studenci którzy uzyskają > 30 punkty są zwolnieni z części pisemnej egzaminu. Zaliczenie ćwiczeń: min 20 punktów.
Egzamin składa się z dwóch części pisemnej i ustnej. W czasie egzaminu pisemnego można uzyskać max 50 punktów. Część pisemna egzaminu jest jednocześnie poprawkowym zaliczaniem laboratoriów. Studenci, którzy uzyskają min 30 punktów łącznie z ćwiczeń i części pisemnej dopuszczeni są do egzaminu ustnego.
Końcowa ocena ustalana jest po egzaminie ustnym po uwzględnieniu wyników z laboratoriów i części pisemnej.

**Egzamin:**

tak

**Literatura:**

1. Faraway, Linear models with R, Chapman and Hall
2. J. Koronacki, J. Cwik, Statystyczne systemy uczące się, Exit

**Witryna www przedmiotu:**

.

**Uwagi:**

.

## Charakterystyki przedmiotowe

### Profil ogólnoakademicki - wiedza

**Charakterystyka W01:**

Zna podstawowe testy nieparametryczne dla pojedynczej próbki i wielu próbek oraz nieparametryczne metody oceny stopnia zależności między cechami.

Weryfikacja:

egzamin pisemny/ustny

**Powiązane charakterystyki kierunkowe:** DS\_W03

**Powiązane charakterystyki obszarowe:** I.P6S\_WG

**Charakterystyka W02:**

Zna podstawy analizy regresji liniowej prostej i wielokrotnej, metodę najmniejszych kwadratów oraz sposoby diagnostyki dopasowania. Zna metodę Lasso, regresję grzbietową, metody regresji nieliniowej, regresji logitowej oraz probitowej.

Weryfikacja:

egzamin pisemny/ustny

**Powiązane charakterystyki kierunkowe:** DS\_W03, DS\_W04

**Powiązane charakterystyki obszarowe:** I.P6S\_WG

**Charakterystyka W03:**

Zna podstawowe metody liniowe i nieliniowe klasyfikacji: LDA, klasyfikację logistyczna, SVM, QDA i drzewa klasyfikacyjne

Weryfikacja:

egzamin pisemny/ustny

**Powiązane charakterystyki kierunkowe:** DS\_W04

**Powiązane charakterystyki obszarowe:** I.P6S\_WG

**Charakterystyka W04:**

Zna podstawowe charakterystyki i modele szeregu czasowego. Zna metody dekompozycji szeregu czasowego oraz metody prognozowania.

Weryfikacja:

egzamin pisemny/ustny

**Powiązane charakterystyki kierunkowe:** DS\_W02, DS\_W03, DS\_W04

**Powiązane charakterystyki obszarowe:** I.P6S\_WG

### Profil ogólnoakademicki - umiejętności

**Charakterystyka U01:**

Potrafi dobrać test nieparametryczny właściwy do badanego zagadnienia i umie zastosować ów test w praktyce. Potrafi ocenić stopień zależności cech.

Weryfikacja:

ocena punktowa aktywności na zajęciach, kolokwium

**Powiązane charakterystyki kierunkowe:** DS\_U05, DS\_U06

**Powiązane charakterystyki obszarowe:** I.P6S\_UW

**Charakterystyka U02:**

Umie dobrać właściwy model regresji oraz przeprowadzić jego diagnostykę. Umie przeprowadzić selekcję zmiennych modelu.

Weryfikacja:

ocena punktowa aktywności na zajęciach, kolokwium

**Powiązane charakterystyki kierunkowe:** DS\_U05, DS\_U06

**Powiązane charakterystyki obszarowe:** I.P6S\_UW

**Charakterystyka U03:**

Umie zastosować właściwą metodę klasyfikacji i ocenić jej skuteczność.

Weryfikacja:

ocena punktowa aktywności na zajęciach, kolokwium

**Powiązane charakterystyki kierunkowe:** DS\_U06

**Powiązane charakterystyki obszarowe:** I.P6S\_UW

**Charakterystyka U04:**

Umie dopasować i przeprowadzić diagnostykę dopasowania podstawowych klas szeregów czasowych. Umie dokonać predykcji i ocenić jej błąd.

Weryfikacja:

ocena punktowa aktywności na zajęciach, kolokwium

**Powiązane charakterystyki kierunkowe:** DS\_U06, DS\_U08

**Powiązane charakterystyki obszarowe:** I.P6S\_UW

### Profil ogólnoakademicki - kompetencje społeczne

**Charakterystyka K01:**

Rozumie potrzebę uczenia się przez całe życie i podnoszenia kompetencji zawodowych

Weryfikacja:

ocena punktowa aktywności na zajęciach

**Powiązane charakterystyki kierunkowe:** DS\_K01

**Powiązane charakterystyki obszarowe:** I.P6S\_KK