**Nazwa przedmiotu:**

Statystyka w analizie danych

**Koordynator przedmiotu:**

Rajmund Kożuszek

**Status przedmiotu:**

Obowiązkowy

**Poziom kształcenia:**

Studia II stopnia

**Program:**

Informatyka

**Grupa przedmiotów:**

Przedmioty techniczne - zaawansowane

**Kod przedmiotu:**

SAD

**Semestr nominalny:**

2 / rok ak. 2021/2022

**Liczba punktów ECTS:**

5

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

1. liczba godzin kontaktowych – 49 godz., w tym
obecność na wykładach 30 godz.,
spotkania projektowe 15 godz.,
konsultacje 2 godz.
obecność na egzaminie 2 godz.
2. praca własna studenta – 80 godz., w tym
wykonanie projektów 40 godz.,
przygotowanie dokumentacji projektu 10 godz.
przygotowanie do wykładów: 10 godz.,
przygotowanie do egzaminu 20 godz.
Łączny nakład pracy studenta wynosi 129 godz., co odpowiada 5 pkt. ECTS.

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

1,89 pkt. ECTS, co odpowiada 49 godz. kontaktowym

**Język prowadzenia zajęć:**

polski

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

2,52 pkt. ECTS, co odpowiada 50 godz. poświęconym wykonaniu projektu i dokumentacji plus 15 godz. spotkań projektowych

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

|  |  |
| --- | --- |
| Wykład:  | 30h |
| Ćwiczenia:  | 0h |
| Laboratorium:  | 0h |
| Projekt:  | 15h |
| Lekcje komputerowe:  | 0h |

**Wymagania wstępne:**

Zakłada się znajomość podstawowych pojęć i metod probabilistyki: prawdopodobieństwo w ujęciu częstościowym i geometrycznym, prawdopodobieństwo warunkowe, zmienne losowe, własności zmiennych losowych, przekształcenia zmiennych losowych.
Wymagana jest również umiejętność rozwiązywania prostych zagadnień optymalizacyjnych: poszukiwanie ekstremów funkcji 1 i 2 zmiennych.

**Limit liczby studentów:**

32

**Cel przedmiotu:**

Celem przedmiotu jest zapoznanie studentów z wybranymi zagadnieniami statystyki mającymi zastosowanie w analizie danych.

**Treści kształcenia:**

Wykład:
1. (2h) Wprowadzenie do zagadnień wnioskowania statystycznego i statystyki opisowej. Modele parametrycznie.
2. (2h) Podstawowe rozkłady zmiennych losowych. Estymacja metodą momentów i jej własności.
3. (4h) Estymatory największej wiarygodności oraz ich własności. Zgodność, nieobciążoność i asymptotyczna normalność estymatorów. Informacja Fishera. Nierówność Cramera-Rao.
4. (2h) Estymatory efektywne. Wykładnicze rodziny rozkładów.
5. (2h) Estymacja parametrów rozkładu normalnego. Estymacja przedziałowa. Bootstrap parametryczny i nieparametryczny.
6. (3h) Estymacja bayesowska. Rozkłady sprzężone.
7. (1h) Estymatory jądrowe gęstości.
8. (2h) Testowanie hipotez prostych. Testy najmocniejsze. Testy randomizowane.
9. (2h) Testowanie hipotez jednostronnych.
10. (4h) Twierdzenie Pearsona. Testowanie zgodności rozkładu. Testowanie niezależności oraz jednorodności zmiennych losowych.
11. (4h) Regresja liniowa i jej własności. Modele liniowe.
12. (2h) Analiza wariancji
Projekt:
W ramach kursu przewiduje się dwa mini projekty wykonywane w kilkuosobowych zespołach. Tematem pierwszego projektu będą wybrane metody prezentacji danych. Drugi projekt dotyczył będzie wybranych metod estymacji, modelowania statystycznego lub testowania hipotez statystycznych.

**Metody oceny:**

Projekty wykonywane będą w kilkuosobowych (3-5 osób w zależności od liczby osób zapisanych na przedmiot). Skład grup będzie losowany do każdego z dwóch projektów. Każdy z projektów będzie poddawany dwuetapowej procedurze oceny: na etapie wstępnej analizy problemu oraz na etapie ostatecznego sprawozdania z wykonania projektu.

**Egzamin:**

tak

**Literatura:**

Biecek P., Analiza danych z programem R, PWN, Warszawa 2013,
 • Dudley, Richard, 18.443 Statistics for Applications, Spring 2009. (Massachusetts Institute of Technology: MIT OpenCourseWare), http://ocw.mit.edu
 • Klonecki Witold, Statystyka dla inżynierów, PWN, Warszawa 1999,
 • Rice John A., Mathematical Statistics and Data Analysis. 3rd ed. Belmont, CA: Duxbury Press, 2006,
 • Zieliński Ryszard, Siedem wykładów wprowadzających do statystyki matematycznej, http://www.impan.gov.pl/~rziel/7ALL.pdf

**Witryna www przedmiotu:**

https://usosweb.usos.pw.edu.pl/kontroler.php?\_action=katalog2/przedmioty/pokazPrzedmiot&prz\_kod=103A-INSZI-MSP-SAD

**Uwagi:**

(-)

## Charakterystyki przedmiotowe