**Nazwa przedmiotu:**

Statystyka obliczeniowa

**Koordynator przedmiotu:**

Prof. dr hab. Jan Mielniczuk

**Status przedmiotu:**

Obowiązkowy

**Poziom kształcenia:**

Studia II stopnia

**Program:**

Informatyka

**Grupa przedmiotów:**

Wspólne

**Kod przedmiotu:**

1120-INPAD-MSP-0112

**Semestr nominalny:**

1 / rok ak. 2017/2018

**Liczba punktów ECTS:**

5

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

.

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

.

**Język prowadzenia zajęć:**

polski

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

.

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

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

**Wymagania wstępne:**

Znajomość rachunku prawdopodobieństwa w zakresie podstawowym

**Limit liczby studentów:**

.

**Cel przedmiotu:**

Celem przedmiotu jest zapoznanie studentów z podstawowymi metodami statystyki obliczeniowej.

**Treści kształcenia:**

Wykład:
Metody eksploracyjnej analizy danych w jednym wymiarze, detekcja obserwacji odstających.
Przestrzeń statystyczna. Problem estymacji, testowania i predykcji. Próby losowe i metody próbkowania.
Metody estymacji i własności estymatorów: obciążenie, MSE, efektywność, zgodność, rozkład asymptotyczny, pivot.
Konstrukcja estymatorów parametrycznych, metodą momentów, największej wiarogodności, własności, metoda delta.
Ocena jakości estymatorów: ryzyko, przedziały ufności , klasyczny i typu bootstrap, walidacja krzyżowa.
Metody estymacji nieparametrycznej: dystrybuanta empiryczna, histogram, estymatory jądrowe.
Metody estymacji funkcji regresji: model regresji liniowej jednokrotnej, estymator MNK, średniej ruchomej, lokalnie liniowy.
Problem testowania: postawienie problemu testowania, statystyka testowa i jej rozkład, obszar krytyczny, błąd I i II rodzaju, specyficzność, czułość testu, p-wartość, krzywa ROC i LIFT, podstawowe testy dla jednej i dwóch prób.
Techniki konstrukcji testów (np. ilorazu wiarogodności) i związane własności, testy wielokrotne. Kontrola FDR, Procedura Bonferroniego, Holma, Benjaminiego-Hochberga.
Testy nieparametryczne.
Pomiar zależności stochastycznej na różnych skalach pomiarowych (rho Pearsona, Spearmana, tau Kendalla). Testowanie niezależności. Modele zależności stochastycznej i ich reprezentacje graficzne.
Analiza tablic kontyngencji, modele logliniowe.
Wprowadzenie do statystyki bayesowskiej; estymator bayesowski, estymator MAP.
Laboratorium:
Praktyczna realizacja tematów omawianych na wykładzie w oparciu o system R w oparciu o rzeczywiste i symulowane zbiory danych.

**Metody oceny:**

Laboratoria maks. 30 p., egzamin maks. 70 p., łącznie maks. 100 p. Na zaliczenie konieczne jest uzyskanie łącznie ponad 50 p. na 100 p. możliwych. Ostateczna ocena z przedmiotu wynika z sumy zdobytych punktów: [0, 50] – 2,0; (50, 60] – 3,0; (60, 70] – 3,5; (70, 80] – 4,0; (80, 90] – 4,5; (90, 100] – 5,0.

**Egzamin:**

tak

**Literatura:**

1. J. Johnson, Probability and Statistics for Computer Science
2. J. Gentle, Computational Statistics
3. G. Givens, J. Hoeting, Computational Statistics

**Witryna www przedmiotu:**

e.mini.pw.edu.pl

**Uwagi:**

.

## Efekty przedmiotowe

### Profil ogólnoakademicki - wiedza

**Efekt W2\_01:**

Zna podstawowe metody reprezentacji danych jednowymiarowych z uwzględnieniem skali pomiarowej, umie zidentyfikować obserwacje odstające na podstawie wykresu pudełkowego i zinterpretować wykres kwantylowy; odróżnia zagadnienia estymacji, testowania i predykcji i zna związki między nimi; zna metody generacji prostych próby losowych z podstawowych rozkładów prawdopodobieństwa; zna metody oceny jakości estymatora na podstawie jego rozkładu, konstrukcji przedział ufności dla odpowiedniego parametru i obliczenia jego błąd średniokwadratowy; zna postawienie problemu testowania dla danych, rozróżnia błędy obu rodzajów, umie zinterpretować p-wartość, umie skonstruować krzywą ROC i LIFT; zna metodę budowy testu LRT dla prostego problemu parametrycznego

Weryfikacja:

egzamin, ocena z zajęć laboratoryjnych

**Powiązane efekty kierunkowe:** PD\_W01, PD\_W03, PD\_W08, PD\_W09, PD\_W13

**Powiązane efekty obszarowe:** , , , ,

**Efekt W2\_02:**

Zna podstawowe metody rangowe i wie jak symulacyjnie wyznaczać rozkład statystyk rangowych; zna podstawowe miary zależności stochastycznej i obszar ich stosowalności; zna podstawowe modele zależności w modelach wielodzielczych i umie skonstruować ich testy w modelach logliniowych

Weryfikacja:

egzamin, ocena z zajęć laboratoryjnych

**Powiązane efekty kierunkowe:** PD\_W01, PD\_W03, PD\_W08, PD\_W09, PD\_W13

**Powiązane efekty obszarowe:** , , , ,

### Profil ogólnoakademicki - umiejętności

**Efekt U2\_01:**

Umie generować próbki pseudolosowe z różnych rozkładów prawdopodobieństwa; umie stosować metody Monte Carlo do całkowania i zagadnień optymalizacyjnych; umie dobrać test właściwy do badanego zagadnienia i potrafi stosować ów test w praktyce

Weryfikacja:

egzamin, ocena z zajęć laboratoryjnych

**Powiązane efekty kierunkowe:** PD\_U06, PD\_U16, PD\_U19

**Powiązane efekty obszarowe:** , ,

**Efekt U2\_02:**

Umie ocenić jakość estymatora parametru i skonstruować przedział ufności (klasyczny i metodą bootstrap)

Weryfikacja:

egzamin, ocena z zajęć laboratoryjnych

**Powiązane efekty kierunkowe:** PD\_U06, PD\_U16, PD\_U19

**Powiązane efekty obszarowe:** , ,

**Efekt U2\_03:**

Umie dokonać oceny kroswalidacyjnej jakości estymatora i metody prognozy

Weryfikacja:

egzamin, ocena z zajęć laboratoryjnych

**Powiązane efekty kierunkowe:** PD\_U06, PD\_U16, PD\_U19

**Powiązane efekty obszarowe:** , ,

**Efekt U2\_04:**

Umie przeprowadzić test niezależności i warunkowej niezależności w tablicach kontyngencji

Weryfikacja:

egzamin, ocena z zajęć laboratoryjnych

**Powiązane efekty kierunkowe:** PD\_U06, PD\_U19

**Powiązane efekty obszarowe:** ,

### Profil ogólnoakademicki - kompetencje społeczne

**Efekt K2\_01:**

Posiada świadomość wagi społecznej rzetelnej analizy statystycznej

Weryfikacja:

egzamin, ocena z zajęć laboratoryjnych

**Powiązane efekty kierunkowe:** PD\_K03

**Powiązane efekty obszarowe:**