**Nazwa przedmiotu:**

Statystyczna Eksploracja Danych

**Koordynator przedmiotu:**

Dr inż. Julian Sienkiewicz

**Status przedmiotu:**

Obowiązkowy

**Poziom kształcenia:**

Studia II stopnia

**Program:**

Fizyka Techniczna

**Grupa przedmiotów:**

Obowiązkowe

**Kod przedmiotu:**

1050-FTEDM-MSP-2SED

**Semestr nominalny:**

1 / rok ak. 2020/2021

**Liczba punktów ECTS:**

3

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

1. godziny kontaktowe – 35 h; w tym
 a) obecność na wykładach – 30 h
 b) obecność na egzaminie – 2 h
 c) uczestniczenie w konsultacjach – 3 h
2. praca własna studenta – 35 h; w tym
 a) przygotowanie do wykładów – 15 h
 b) zapoznanie się z literaturą – 10 h
 c) przygotowanie do egzaminu/kolokwiów – 10 h

Razem w semestrze 70 h, co odpowiada 3 pkt. ECTS

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

1. obecność na wykładach – 30 h
2. obecność na egzaminie – 2 h
3. uczestniczenie w konsulatacjach – 3 h

Razem w semestrze 35 h, co odpowiada 1.5 pkt. ECTS

**Język prowadzenia zajęć:**

polski

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

Dla chętnych studentów

1. przygotowanie projektów – 30h
Razem w semestrze 30 h, co odpowiada 1 pkt. ECTS

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

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

**Wymagania wstępne:**

Probabilistyka, Komputerowa Analiza Danych Doświadczalnych, Wprowadzenie do Data Science

**Limit liczby studentów:**

brak

**Cel przedmiotu:**

Poznanie i zrozumienie podstaw eksploracji danych, ich ograniczeń oraz obszarów zastosowań.

**Treści kształcenia:**

1. Wprowadzenie, dyskryminacja fisherowska dla g=2
2. Dyskryminacja fisherowska - przykłady, dyskryminacja fisherowska dla wyższych g
3. Klasyfikator Bayesa, LDA, QDA, skuteczność klasyfikatorów
4. Metoda najbliższych sąsiadów
5. Drzewa klasyfikacyjne
6. Zespoły klasyfikatorów. Bagging, boosting, lasy losowe.
7. Analiza skupień
8. Analiza składowych głównych i skalowanie wielowymiarowe
9. Analiza czynnikowa
10. Maszyny wektorów podpierających
11. Jądra – metody nieliniowe (kernels) w SVM, CCA i PCA.
12. Modele graficzne: sieci bayesowskie, sieci Markowa.

**Metody oceny:**

Podstawą oceny końcowej jest egzamin. Studenci, którzy w trakcie semestru wezmą udział w trzech kolokwiach oraz przedstawią wyniki projektu i otrzymają sumaryczną ocenę db, db1/2 lub bdb są zwolnieni z taką oceną z egzaminu.

**Egzamin:**

tak

**Literatura:**

1. L. Rutkowski, Metody i techniki sztucznej inteligencji, PWN, Warszawa 2005
2. D. Larose, Metody i modele eksploracji danych, PWN, Warszawa 2008.
3. D. Hand, H. Manilla, P. Smyth, Eksploracja danych, WNT, Warszawa 2005.
4. J. Koronacki, J. Ćwik, Systemy uczące się: rozpoznawanie wzorców, analiza skupień i redukcja wymiarowości, WNT, Warszawa 2008

**Witryna www przedmiotu:**

http://www.fizyka.pw.edu.pl/~julas/SED

**Uwagi:**

brak

## Efekty przedmiotowe

### Profil ogólnoakademicki - wiedza

**Efekt SED\_W01:**

Ma podstawową wiedzę w zakresie eksploracji danych.

Weryfikacja:

Kolokwium / egzamin

**Powiązane efekty kierunkowe:** FT2\_W02

**Powiązane efekty obszarowe:** X2A\_W02, T2A\_W01, T2A\_W02

**Efekt SED\_W02:**

Ma świadomość współczesnych kierunków badań w zakresie eksploracji danych

Weryfikacja:

Kolokwium / egzamin

**Powiązane efekty kierunkowe:** FT2\_W04

**Powiązane efekty obszarowe:** X2A\_W06, T2A\_W05, T2A\_W07

### Profil ogólnoakademicki - umiejętności

**Efekt SED\_U01:**

Potrafi zastosować zdobytą wiedzę do opisu różnych, z góry niedefiniowanych zjawisk.

Weryfikacja:

Kolokwium / egzamin

**Powiązane efekty kierunkowe:** FT2\_U03, FT2\_U06

**Powiązane efekty obszarowe:** X2A\_U08, X2A\_U09, T2A\_U03, T2A\_U04, X2A\_U02, X2A\_U04, T2A\_U09

**Efekt SED\_U02:**

Umie ocenić przydatność poszczególnych metod i ich wykorzystanie do różnie zdefiniowanych problemów.

Weryfikacja:

Kolokwium/
egzamin

**Powiązane efekty kierunkowe:** FT2\_U12

**Powiązane efekty obszarowe:** X2A\_U04, T2A\_U13

### Profil ogólnoakademicki - kompetencje społeczne

**Efekt SED\_K01:**

Potrafi myśleć w sposób przedsiębiorczy oraz określić priorytet zawiązany z realizacją wybranego przez siebie zadania.

Weryfikacja:

Kolokwium/
egzamin

**Powiązane efekty kierunkowe:** FT2\_K05

**Powiązane efekty obszarowe:** X2A\_K03, T2A\_K04, T2A\_K05

**Efekt SED\_K02:**

Ma świadomość różnicy pomiędzy popularnym zrozumieniem metod „data mining” a ich faktyczną rolą w nauce

Weryfikacja:

Kolokwium/
egzamin

**Powiązane efekty kierunkowe:** FT2\_K06

**Powiązane efekty obszarowe:** X2A\_K04, T2A\_K07