

# POLITECHNIKA KRAKOWSKA IM. TADEUSZA KOŚCIUSZKI

## KARTA PRZEDMIOTU

obowiązuje studentów rozpoczynających studia w roku akademickim 2023/2024

Wydział Inżynierii Środowiska i Energetyki

Kierunek studiów: Geoinformatyka

Profil: Ogólnoakademicki

Forma studiów: stacjonarne

Kod kierunku: 12

Stopień studiów: II

Specjalności: bez specjalności

### 1 INFORMACJE O PRZEDMIOCIE

|                                         |                             |
|-----------------------------------------|-----------------------------|
| NAZWA PRZEDMIOTU                        | Statystyka II i Data Mining |
| NAZWA PRZEDMIOTU<br>W JĘZYKU ANGIELSKIM | Statistics and Data Mining  |
| KOD PRZEDMIOTU                          | WIŚIE GI oIIS C6 23/24      |
| KATEGORIA PRZEDMIOTU                    | Przedmioty kierunkowe       |
| LICZBA PUNKTÓW ECTS                     | 7.00                        |
| SEMESTRY                                | 1 2                         |

### 2 RODZAJ ZAJĘĆ, LICZBA GODZIN W PLANIE STUDIÓW

| SEMESTR | WYKŁAD | CWICZENIA | LABORATORIA | LABORATORIA<br>KOMPUTERO-<br>WE | PROJEKT | SEMINARIUM |
|---------|--------|-----------|-------------|---------------------------------|---------|------------|
| 1       | 15     | 0         | 0           | 15                              | 15      | 0          |
| 2       | 15     | 0         | 0           | 15                              | 15      | 0          |

### 3 CELE PRZEDMIOTU

**Cel 1** Przekazanie wiedzy w zakresie algorytmów ogólnej analizy danych

**Cel 2** Przekazanie umiejętności zastosowania algorytmów ogólnej analizy danych do rozwiązywania problemów inżynierskich

## 4 WYMAGANIA WSTĘPNE W ZAKRESIE WIEDZY, UMIEJĘTNOŚCI I INNYCH KOMPETENCJI

1 Zaliczone kursy w zakresie analizy matematycznej oraz metod probabilistycznych i statystyki

## 5 EFEKTY KSZTAŁCENIA

**EK1 Wiedza** Student zna i rozumie metody analizy danych jakościowych

**EK2 Wiedza** Student zna i rozumie metody analizy danych ilościowych

**EK3 Umiejętności** Student potrafi zastosować metody analizy danych jakościowych do rozwiązywania wybranych zagadnień inżynierskich

**EK4 Umiejętności** Student potrafi zastosować metody analizy danych ilościowych do rozwiązywania wybranych zagadnień inżynierskich

## 6 TREŚCI PROGRAMOWE

| WYKŁAD    |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      |                  |
|-----------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|------------------|
| LP        | TEMATYKA ZAJĘĆ<br>OPIS SZCZEGÓŁOWY BLOKÓW TEMATYCZNYCH                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               | LICZBA<br>GODZIN |
| <b>W1</b> | Tematy (podlegają zmianom): Cechy i ograniczenia danych ilościowych i jakościowych. Skale pomiarowe Stevensa. Regresja wieloraka. Regresja wielowymiarowa. Liniowa redukcja wymiarowości. Liniowa analiza dyskryminacyjna. Drzewa klasyfikacyjne i regresyjne. Metody regresji liniowej. Uogólnione modele liniowe. Analiza ryzyka. Analiza skupień. Analityka predykcyjna. Analiza koszykowa. Funkcje predykcyjne k-najbliższych sąsiadów. Naiwny algorytm Bayesa. Prognozowanie. Analityka w czasie rzeczywistym. Wizualizacja danych jedno-, dwu- i wielowymiarowych. Imputacja danych. Metody resamplingu. Mapowanie danych. Słowniki danych. Algorytmy skalowalne i statystyki asocjatywne. Dane jedno- i wielowymiarowe. Systemy składująco-przetwarzające Hadoop i MapReduce. | 30               |

| PROJEKT   |                                                                                                                                                  |                  |
|-----------|--------------------------------------------------------------------------------------------------------------------------------------------------|------------------|
| LP        | TEMATYKA ZAJĘĆ<br>OPIS SZCZEGÓŁOWY BLOKÓW TEMATYCZNYCH                                                                                           | LICZBA<br>GODZIN |
| <b>P1</b> | Realizacja projektów z zakresu: a) ogólnej analizy danych ilościowych, b) ogólnej analizy danych jakościowych we wskazanym obszarze tematycznym. | 30               |

| LABORATORIA KOMPUTEROWE |                                                        |                  |
|-------------------------|--------------------------------------------------------|------------------|
| LP                      | TEMATYKA ZAJĘĆ<br>OPIS SZCZEGÓŁOWY BLOKÓW TEMATYCZNYCH | LICZBA<br>GODZIN |

| LABORATORIA KOMPUTEROWE |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     |                  |
|-------------------------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|------------------|
| LP                      | TEMATYKA ZAJĘĆ<br>OPIS SZCZEGÓŁOWY BLOKÓW TEMATYCZNYCH                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              | LICZBA<br>GODZIN |
| K1                      | Tematy (podlegają zmianom): Regresja liniowa wraz z analityką. Regresja wieloraka wraz z analityką. Redukcja wymiarowości zadanego datasetu ilościowego za pomocą PCA. Regresja wielowymiarowa wraz z analityką. Identyfikacja drzewa regresyjnego. Analiza skupień dla zadanego datasetu ilościowego. Naprawa datasetu za pomocą imputacji. Identyfikacja rozkładu za pomocą metody bootstrap. Identyfikacja klasyfikatora dla zadanego datasetu jakościowego. Identyfikacja regresji logistycznej dla zadanego datasetu jakościowego. Analiza korpusu tekstowego. | 30               |

## 7 NARZĘDZIA DYDAKTYCZNE

N1 Wykłady

N2 Ćwiczenia laboratoryjne

N3 Ćwiczenia projektowe

## 8 OBCIĄŻENIE PRACĄ STUDENTA

| FORMA AKTYWNOŚCI                                                                                 | ŚREDNIA LICZBA GODZIN<br>NA ZREALIZOWANIE<br>AKTYWNOŚCI |
|--------------------------------------------------------------------------------------------------|---------------------------------------------------------|
| <b>Godziny kontaktowe z nauczycielem akademickim, w tym:</b>                                     |                                                         |
| Godziny wynikające z planu studiów                                                               | 90                                                      |
| Konsultacje przedmiotowe                                                                         | 11                                                      |
| Egzaminy i zaliczenia w sesji                                                                    | 4                                                       |
| <b>Godziny bez udziału nauczyciela akademickiego wynikające z nakładu pracy studenta, w tym:</b> |                                                         |
| Przygotowanie się do zajęć, w tym studiowanie zalecanej literatury                               | 40                                                      |
| Opracowanie wyników                                                                              | 0                                                       |
| Przygotowanie raportu, projektu, prezentacji, dyskusji                                           | 30                                                      |
| <b>SUMARYCZNA LICZBA GODZIN DLA PRZEDMIOTU WYNIKAJĄCA Z<br/>CAŁEGO NAKŁADU PRACY STUDENTA</b>    | <b>175</b>                                              |
| SUMARYCZNA LICZBA PUNKTÓW ECTS DLA PRZEDMIOTU                                                    | 7.00                                                    |

## 9 SPOSOBY OCENY

OCENA FORMUJĄCA

F1 Kolokwium

**F2** Ćwiczenie praktyczne

**F3** Projekt indywidualny

### **OCENA PODSUMOWUJĄCA**

**P1** Średnia ważona ocen formujących

### **WARUNKI ZALICZENIA PRZEDMIOTU**

**W1** Pozytywna ocena z kolokwium egzaminacyjnego

**W2** Pozytywne oceny z laboratoriów komputerowych

**W3** Pozytywna ocena z projektu

**W4** Obecność na co najmniej 66% laboratoriów komputerowych

### **KRYTERIA OCENY**

| EFEKT KSZTAŁCENIA 1 |                                                                                |
|---------------------|--------------------------------------------------------------------------------|
| NA OCENĘ 2.0        | Student zaliczył sprawdzian na poziomie poniżej 50% maksymalnej liczby punktów |
| NA OCENĘ 3.0        | Student zaliczył sprawdzian na poziomie 50% maksymalnej liczby punktów         |
| NA OCENĘ 3.5        | Student zaliczył sprawdzian na poziomie 60% maksymalnej liczby punktów         |
| NA OCENĘ 4.0        | Student zaliczył sprawdzian na poziomie 70% maksymalnej liczby punktów         |
| NA OCENĘ 4.5        | Student zaliczył sprawdzian na poziomie 80% maksymalnej liczby punktów         |
| NA OCENĘ 5.0        | Student zaliczył sprawdzian na poziomie 90% maksymalnej liczby punktów         |
| EFEKT KSZTAŁCENIA 2 |                                                                                |
| NA OCENĘ 2.0        | Student zaliczył sprawdzian na poziomie poniżej 50% maksymalnej liczby punktów |
| NA OCENĘ 3.0        | Student zaliczył sprawdzian na poziomie 50% maksymalnej liczby punktów         |
| NA OCENĘ 3.5        | Student zaliczył sprawdzian na poziomie 60% maksymalnej liczby punktów         |
| NA OCENĘ 4.0        | Student zaliczył sprawdzian na poziomie 70% maksymalnej liczby punktów         |
| NA OCENĘ 4.5        | Student zaliczył sprawdzian na poziomie 80% maksymalnej liczby punktów         |
| NA OCENĘ 5.0        | Student zaliczył sprawdzian na poziomie 90% maksymalnej liczby punktów         |
| EFEKT KSZTAŁCENIA 3 |                                                                                |
| NA OCENĘ 2.0        | Student zaliczył sprawdzian na poziomie poniżej 50% maksymalnej liczby punktów |
| NA OCENĘ 3.0        | Student zaliczył sprawdzian na poziomie 50% maksymalnej liczby punktów         |

|                     |                                                                                |
|---------------------|--------------------------------------------------------------------------------|
| NA OCENĘ 3.5        | Student zaliczył sprawdzian na poziomie 60% maksymalnej liczby punktów         |
| NA OCENĘ 4.0        | Student zaliczył sprawdzian na poziomie 70% maksymalnej liczby punktów         |
| NA OCENĘ 4.5        | Student zaliczył sprawdzian na poziomie 80% maksymalnej liczby punktów         |
| NA OCENĘ 5.0        | Student zaliczył sprawdzian na poziomie 90% maksymalnej liczby punktów         |
| EFEKT KSZTAŁCENIA 4 |                                                                                |
| NA OCENĘ 2.0        | Student zaliczył sprawdzian na poziomie poniżej 50% maksymalnej liczby punktów |
| NA OCENĘ 3.0        | Student zaliczył sprawdzian na poziomie 50% maksymalnej liczby punktów         |
| NA OCENĘ 3.5        | Student zaliczył sprawdzian na poziomie 60% maksymalnej liczby punktów         |
| NA OCENĘ 4.0        | Student zaliczył sprawdzian na poziomie 70% maksymalnej liczby punktów         |
| NA OCENĘ 4.5        | Student zaliczył sprawdzian na poziomie 80% maksymalnej liczby punktów         |
| NA OCENĘ 5.0        | Student zaliczył sprawdzian na poziomie 90% maksymalnej liczby punktów         |

## 10 MACIERZ REALIZACJI PRZEDMIOTU

| EFEKT KSZTAŁCENIA | ODNIESIENIE DANEGO EFEKTU DO SZCZEGÓŁOWYCH EFEKTÓW ZDEFINIOWANYCH DLA PROGRAMU | CELE PRZEDMIOTU | TREŚCI PROGRAMOWE | NARZĘDZIA DYDAKTYCZNE | SPOSOBY OCENY |
|-------------------|--------------------------------------------------------------------------------|-----------------|-------------------|-----------------------|---------------|
| EK1               | K_W07 K_W08                                                                    | Cel 1           | W1 P1 K1          | N1 N2 N3              | F1 F2 F3 P1   |
| EK2               | K_W07 K_W08                                                                    | Cel 1           | W1 P1 K1          | N1 N2 N3              | F1 F2 F3 P1   |
| EK3               | K_U06                                                                          | Cel 2           | W1 P1 K1          | N1 N2 N3              | F1 F2 F3 P1   |
| EK4               | K_U06                                                                          | Cel 2           | W1 P1 K1          | N1 N2 N3              | F1 F2 F3 P1   |

## 11 WYKAZ LITERATURY

### LITERATURA PODSTAWOWA

- [1] | **S. Brandt** — *Analiza danych*, Warszawa, 1998, PWN
- [2] | **S. Heinz** — *Mathematical modeling*, , 2011, Springer

- [3] | **A.J. Izenman** — *Modern multivariate statistical techniques. Regression, classification, and manifold learning*, 2008, Springer
- [4] | **A. Agresti** — *Categorical data analysis*, Hoboken, 2002, Wiley
- [5] | **I.T. Jolliffe** — *Principal component analysis. Second edition*, 2010, Springer
- [6] | **B.S. Everitt, S. Landau, M. Leese, D. Stahl** — *Cluster analysis. 5th Edition*, Hoboken, 2012, Wiley
- [7] | **B. Steele, J. Chandler, S. Reddy**, — *Algorithms for Data Science*, 2016, Springer

## 12 INFORMACJE O NAUCZYCIELACH AKADEMICKICH

### OSOBA ODPOWIEDZIALNA ZA KARTĘ

dr hab. inż., prof. PK Jacek Pietraszek (kontakt: [jacek.pietraszek@pk.edu.pl](mailto:jacek.pietraszek@pk.edu.pl))

### OSOBY PROWADZĄCE PRZEDMIOT

1 pracownicy Katedry Informatyki Stosowanej (kontakt: [m-7@pk.edu.pl](mailto:m-7@pk.edu.pl))

## 13 ZATWIERDZENIE KARTY PRZEDMIOTU DO REALIZACJI

---

(miejscowość, data)

(odpowiedzialny za przedmiot)

(dziekan)

PRZYJMUJĘ DO REALIZACJI (data i podpisy osób prowadzących przedmiot)

.....