

## WYDZIAŁ MATEMATYKI

### KARTA PRZEDMIOTU

**Nazwa w języku polskim:** Pozyskiwanie wiedzy

**Nazwa w języku angielskim:** Data Mining

**Kierunek studiów (jeśli dotyczy):** MATEMATYKA

**Specjalność (jeśli dotyczy):** Statystyka matematyczna

**Stopień studiów i forma:** 2 stopień, stacjonarna / ~~niestacjonarna~~\*

**Rodzaj przedmiotu:** ~~obowiązkowy~~ / wybieralny / ~~ogólnouczelniany~~\*

**Kod przedmiotu** INP1908

**Grupa kursów** TAK / ~~NIE~~

|   | Wykład                        | Ćwiczenia | Laboratorium | Projekt | Seminarium |
|---|-------------------------------|-----------|--------------|---------|------------|
| Liczba godzin zajęć zorganizowanych w Uczelni (ZZU)                                       | 30                            |           | 15           |         |            |
| Liczba godzin całkowitego nakładu pracy studenta (CNPS)                                   | 150                           |           |              |         |            |
| Forma zaliczenia  | Egzamin / zaliczenie na ocenę |           |              |         |            |
| Dla grupy kursów zaznaczyć kurs końcowy   | X                             |           |              |         |            |
| Liczba punktów ECTS   | 5                             |           |              |         |            |
| w tym liczba punktów odpowiadająca zajęciom o charakterze praktycznym (P)                 | 3                             |           |              |         |            |
| W tym liczba punktów ECTS odpowiadająca zajęciom wymagającym bezpośredniego kontaktu (BK) | 3                             |           |              |         |            |

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

1. Wstęp do rachunku prawdopodobieństwa,
2. Wstęp do statystyki matematycznej.
3. Wstęp do programowania.

#### CELE PRZEDMIOTU

- C1 Poznanie podstawowych rodzajów zadań data mining (pozyskiwania wiedzy).
- C2 Opanowanie podstawowej wiedzy na temat metod eksploracji danych oraz ich własności.
- C3 Poznanie klasycznych i nowoczesnych metod klasyfikacji, redukcji wymiaru oraz analizy skupień.
- C4 Poznanie podstawowych algorytmów stosowanych w odkrywaniu reguł asocjacyjnych.
- C5 Poznanie metod stosowanych w ocenie jakości klasyfikacji i analizy skupień.
- C6 Stosowanie zdobytej wiedzy do rozwiązywania zagadnień praktycznych z różnych dziedzin nauki, techniki i ekonomii.

### PRZEDMIOTOWE EFEKTY KSZTAŁCENIA

Z zakresu wiedzy student:

PE\_W01 ma wiedzę dotyczącą podstawowych rodzajów zadań data mining,

PE\_W02 zna podstawowe metody klasyfikacji, redukcji wymiaru, analizy skupień i odkrywania reguł asocjacyjnych oraz ich własności,

PE\_W03 zna podstawowe metody oceny jakości klasyfikacji i analizy skupień.

Z zakresu umiejętności student:

PE\_U01 potrafi odpowiednio dobierać metody umożliwiające realizację określonego zadania eksploracji danych,

PE\_U02 potrafi stosować metody/algorytmy uczenia nadzorowanego i uczenia bez nadzoru,

PE\_U03 potrafi weryfikować efektywność stosowanych metod.

Z zakresu kompetencji społecznych student:

PE\_K01 potrafi korzystać z literatury naukowej, w tym docierać do materiałów źródłowych oraz dokonywać ich przeglądu,

PE\_K02 rozumie konieczność systematycznej i samodzielnej pracy nad opanowaniem materiału kursu.

### TREŚCI PROGRAMOWE

| Forma zajęć - wykłady |  | Liczba godzin |
|-----------------------|--|---------------|
| Wy1                   | Wprowadzenie do problematyki pozyskiwania wiedzy/data mining   | 2             |
| Wy2                   | Podstawy analizy danych wielowymiarowych   | 2             |
| Wy3                   | Wprowadzenie do środowiska R   | 2             |
| Wy4                   | Przygotowanie danych do analiz data mining   | 2             |
| Wy5                   | Metody redukcji wymiaru: metoda składowych głównych (PCA), skalowanie wielowymiarowe (MDS)   | 2             |
| Wy6                   | Metody klasyfikacji danych: metoda k-sąsiadów (k-nn), drzewa klasyfikacyjne, naiwny klasyfikator bayesowski, analiza dyskryminacyjna, regresja logistyczna | 6             |
| Wy7                   | Analiza skupień - metody grupujące i hierarchiczne   | 4             |
| Wy8                   | Ocena jakości klasyfikacji i analizy skupień   | 2             |
| Wy9                   | Maszyny wektorów wspierających (SVM)   | 2             |
| Wy10                  | Sieci neuronowe  | 2             |
| Wy11                  | Odkrywanie reguł asocjacyjnych   | 2             |
| Wy12                  | Rodziny klasyfikatorów: bagging, boosting, random forest   | 2             |
|                       | Suma godzin  | <b>30</b>     |
|                       |  |               |

| Forma zajęć - laboratorium |   | Liczba godzin |
|----------------------------|---|---------------|
| Lal                        | Wprowadzenie do środowiska R, elementy programowania, import/eksport danych | 2             |

|     |  |           |
|-----|--|-----------|
| La2 | Metody analizy opisowej i wizualizacji danych wielowymiarowych   | 2         |
| La3 | Metody redukcji wymiaru (PCA, MDS)   | 2         |
| La4 | Metody klasyfikacji danych (k-nn, drzewa klasyfikacyjne, analiza dyskryminacyjna i regresja logistyczna) | 4         |
| La5 | Analiza skupień - metody grupujące (k-means, PAM) i metody hierarchiczne (AGNES, DIANA, MONA)            | 3         |
| La6 | Analiza reguł asocjacyjnych  | 2         |
|     | Suma godzin  | <b>15</b> |

### STOSOWANE NARZĘDZIA DYDAKTYCZNE

1. Wykład problemowy – metoda tradycyjna,
2. Zajęcia laboratoryjne w pracowni komputerowej.
3. Konsultacje,
4. Praca własna studenta – przygotowanie do zajęć laboratoryjnych.

### OCENA OSIĄGNIĘCIA EFEKTÓW KSZTAŁCENIA

| Oceny (F – formująca (w trakcie semestru), P – podsumowująca (na koniec semestru)) | Numer efektu kształcenia                   | Sposób oceny osiągnięcia efektu kształcenia                 |
|--|--|---|
| F1   | PEK_U01,PEK_U02, PEK_U03,PEK_K01, PEK_K02, | Odpowiedzi ustne, raporty z zadań laboratoryjnych, projekty |
| F2   | PEK_W01,PEK_W02, PEK_W03,PEK_K01, PEK_K02, | Kolokwium zaliczeniowe na wykładzie.                        |
| P = 60%F1 + 40%F2  |  |   |

### LITERATURA PODSTAWOWA I UZUPEŁNIAJĄCA

#### **LITERATURA PODSTAWOWA:**

- [1] P.-N. Tan, M. Steinbach, V. Kumar, Introduction to Data Mining, Addison-Wesley, 2006.
- [2] T.Hastie, R.Tibshirani, J. Friedman, The Elements of Statistical Learning: Data Mining, Inference, and Prediction, Springer, 2011.
- [3] D.T. Larose, Odkrywanie wiedzy z danych. Wprowadzenie do eksploracji danych. PWN, 2006.
- [4] D.T. Larose, Metody i modele eksploracji danych, PWN, 2008.
- [5] D.J. Hand, H. Mannila, P. Smyth, Eksploracja danych, WNT, 2005.

#### **LITERATURA UZUPEŁNIAJĄCA:**

- [1] J. Koronacki, J. Ćwik, Statystyczne systemy uczące się, Exit, 2008.
- [2] W.N. Venables, B.D. Ripley, Modern Applied Statistics With S, Springer, 2001.
- [3] R.A. Johnson, D.W. Wichern, Applied multivariate statistical analysis, Pearson

Prentice Hall, 2002.

[4] P. Dalgaard, Introductory Statistics with R, Springer, 2001.

[5] Marek Walesiak, Eugeniusz Gatnar, Statystyczna analiza danych z wykorzystaniem programu R, PWN, 2011.

**OPIEKUN PRZEDMIOTU (IMIE, NAZWISKO, ADRES E-MAIL)**

dr Adam Zagdański ( [adam.zagdanski@pwr.wroc.pl](mailto:adam.zagdanski@pwr.wroc.pl))

dr Artur Suchwałko ( [artur.suchwalko@pwr.wroc.pl](mailto:artur.suchwalko@pwr.wroc.pl))

MACIERZ POWIĄZANIA EFEKTÓW KSZTAŁCENIA DLA PRZEDMIOTU  
**Pozyskiwanie wiedzy**  
 Z EFEKTAMI KSZTAŁCENIA NA KIERUNKU MATEMATYKA  
 I SPECJALNOŚCI **Statystyka matematyczna**

| Przedmiotowy efekt kształcenia   | Odniesienie przedmiotowego efektu do efektów kształcenia zdefiniowanych dla kierunku studiów i specjalności (o ile dotyczy) | Cele przedmiotu**         | Treści programowe**   | Numer narzędzia dydaktycznego** |
|----------------------------------|---|---------------------------|-----------------------|---------------------------------|
| <b>PEK_W01</b><br>(wiedza)       | K2MAT_W03, K2MAT_W05,<br>K2MAT_W07-K2MAT_W11  | C1                        | Wy1, Wy2              | 1,3                             |
| <b>PEK_W02</b>                   | K2MAT_W03, K2MAT_W05,<br>K2MAT_W07-K2MAT_W11  | C2, C3, C4                | Wy3-Wy17,<br>Wy9-Wy12 | 1,3                             |
| <b>PEK_W03</b>                   | K2MAT_W03, K2MAT_W05,<br>K2MAT_W07-K2MAT_W11  | C5                        | Wy6, Wy7,<br>Wy8      | 1,3                             |
| <b>PEK_U01</b><br>(umiejętności) | K2MAT_U02 K2MAT_U03<br>K2MAT_U04 K2MAT_U05<br>K2MAT_U06 K2MAT_U08   | C1, C2,C3,<br>C4, C6      | La3-La6               | 2,3,4                           |
| <b>PEK_U02</b>                   | K2MAT_U02 K2MAT_U03<br>K2MAT_U04 K2MAT_U05<br>K2MAT_U06 K2MAT_U08   | C2,C3, C4,<br>C6          | La1-La6               | 2,3,4                           |
| <b>PEK_U03</b>                   | K2MAT_U02 K2MAT_U03<br>K2MAT_U04 K2MAT_U05<br>K2MAT_U06 K2MAT_U08   | C5, C6                    | La3-La6               | 2,3,4                           |
| <b>PEK_K01</b><br>(kompetencje)  | K2MAT_K01, K2MAT_K02,<br>K2MAT_K04, K2MAT_K05,<br>K2MAT_K06   | C1, C2, C3,<br>C4, C5, C6 | Wy1-Wy12<br>La1-La6   | 1,2,3,4                         |
| <b>PEK_K02</b>                   | K2MAT_K01, K2MAT_K02<br>K2MAT_K04, K2MAT_K05,<br>K2MAT_K06  | C1, C2, C3,<br>C4, C5, C6 | Wy1-Wy12<br>La1-La6   | 1,2,3,4                         |

\*\* - z tabeli powyżej