

**WYDZIAŁ MATEMATYKI  
KARTA PRZEDMIOTU**

**Nazwa w języku polskim: Eksploracja Danych**

**Nazwa w języku angielskim: Data Mining**

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

**Stopień studiów i forma: 1 stopień, stacjonarna /niestacjonarna\***

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

**Kod przedmiotu: MAT001631**

**Grupa kursów: TAK / NIE**

	Wykład	Ćwiczenia	Laboratorium	Projekt	Seminarium
Liczba godzin zajęć zorganizowanych w Uczelni (ZZU)	30		30		
Liczba godzin całkowitego nakładu pracy studenta (CNPS)	120				
Forma zaliczenia	<del>Egzamin</del> / zaliczenie na ocenę				
Dla grupy kursów zaznaczyć kurs końcowy	X				
Liczba punktów ECTS	4				
w tym liczba punktów odpowiadająca zajęciom o charakterze praktycznym (P)	2				
W tym liczba punktów ECTS odpowiadająca zajęciom wymagającym bezpośredniego kontaktu (BK)	2				

**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 eksploracji danych (data mining).
- 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 Umiejętność stosowania 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 eksploracji danych,  
PE\_W02 zna podstawowe metody klasyfikacji, redukcji wymiaru, analizy skupień (grupowania) 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 podstawowe metody/algorytmy redukcji wymiaru, klasyfikacji i grupowania danych,  
PE\_U03 potrafi weryfikować własności stosowanych metod eksploracji danych.

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 eksploracji danych. Cel i rodzaje zadań eksploracji.	2
Wy2	Podstawy analizy danych wielowymiarowych.	2
Wy3	Przygotowanie danych do eksploracji. Obsługa brakujących danych. Podstawowe metody identyfikacji obserwacji odstających. Niezbędne przekształcenia danych.	4
Wy4	Metody redukcji wymiaru. Analiza składowych głównych (PCA). Skalowanie wielowymiarowe (MDS).	4
Wy5	Klasyfikacja danych. Idea klasyfikacji i przegląd wybranych algorytmów (m.in.: metoda k najbliższych sąsiadów (k-nn), drzewa klasyfikacyjne, naiwny klasyfikator bayesowski).	6
Wy6	Analiza skupień (grupowanie). Cel analizy skupień. Metody grupujące i hierarchiczne (m.in. algorytmy: k-means, PAM, AGNES, DIANA).	4
Wy7	Metody stosowane w ocenie jakości klasyfikacji i analizy skupień.	2
Wy8	Maszyny wektorów wspierających (SVM).	2
Wy9	Rodziny klasyfikatorów. Algorytmy: bagging, boosting i lasy losowe (random forest).	2
Wy10	Wprowadzenie do odkrywania reguł asocjacyjnych. Algorytm Apriori.	2
	Suma godzin	<b>30</b>

<b>Forma zajęć - laboratorium</b>		<b>Liczba godzin</b>
La1	Metody analizy opisowej i wizualizacji danych wielowymiarowych. Poznanie danych i wybór interesującego podzbioru do dalszych analiz.	3
La2	Przygotowanie (wstępna obróbka) danych. Obsługa brakujących danych. Identyfikacja obserwacji odstających. Niezbędne przekształcenia danych.	4
La3	Metody redukcji wymiaru. Algorytmy PCA i MDS.	4
La4	Klasyfikacja. Algorytm k najbliższych sąsiadów (k-nn), drzewa klasyfikacyjne, naiwny klasyfikator bayesowski.	4
La5	Analiza skupień - metody grupujące (k-means, PAM).	2
La6	Analiza skupień - metody hierarchiczne (AGNES, DIANA, MONA)	2
La7	Ocena jakości klasyfikacji i analizy skupień.	4
La8	Maszyny wektorów podpierających (SVM).	2
La9	Rodziny klasyfikatorów: bagging, boosting i lasy losowe (random forest).	3
La10	Podstawy odkrywania reguł asocjacyjnych.	2
	Suma godzin	<b>30</b>

<b>STOSOWANE NARZĘDZIA DYDAKTYCZNE</b>
<ol style="list-style-type: none"> <li>1. Wykład problemowy – metoda tradycyjna,</li> <li>2. Zajęcia laboratoryjne w pracowni komputerowej.</li> <li>3. Konsultacje,</li> <li>4. Praca własna studenta – przygotowanie do zajęć laboratoryjnych.</li> </ol>

### **OCENA OSIĄGNIĘCIA EFEKTÓW KSZTAŁCENIA**

<b>Oceny</b> (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] G. James, D. Witten, T. Hastie, R. Tibshirani, An Introduction to Statistical Learning with Applications in R. Springer, 2013.
- [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] T. Morzy, Eksploracja danych: metody i algorytmy. PWN, 2013.
- [3] W.N. Venables, B.D. Ripley, Modern Applied Statistics With S, Springer, 2001.
- [4] M. Walesiak, E. Gatnar, Statystyczna analiza danych z wykorzystaniem programu R. PWN, 2011.

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

dr Adam Zagdański ([Adam.Zagdanski@pwr.edu.pl](mailto:Adam.Zagdanski@pwr.edu.pl))

**MACIERZ POWIĄZANIA EFEKTÓW KSZTAŁCENIA DLA PRZEDMIOTU  
EKSPLOACJA DANYCH  
Z EFEKTAMI KSZTAŁCENIA NA KIERUNKU MATEMATYKA I STATYSTYKA**

<b>Przedmiotowy efekt kształcenia</b>	<b>Odniesienie przedmiotowego efektu do efektów kształcenia zdefiniowanych dla kierunku studiów i specjalności (o ile dotyczy)</b>	<b>Cele przedmiotu**</b>	<b>Treści programowe**</b>	<b>Numer narzędzia dydaktycznego**</b>
<b>PEK_W01 (wiedza)</b>	K1MIS_W07, K1MIS_W22_SAD, K1MIS_W23_SAD	C1	Wy1, Wy2	1,3
<b>PEK_W02</b>	K1MIS_W07, K1MIS_W22_SAD, K1MIS_W23_SAD	C2, C3, C4	Wy3-Wy6, Wy8-Wy10	1,3
<b>PEK_W03</b>	K1MIS_W07, K1MIS_W22_SAD, K1MIS_W23_SAD	C5	Wy5-Wy7	1,3
<b>PEK_U01 (umiejętności)</b>	K1MIS_U16, K1MIS_U19, K1MIS_U27, K1MIS_U32SAD, K1MIS_U34SAD, K1MIS_U35SAD	C1, C2, C3, C4, C6	La3-La10	2,3,4
<b>PEK_U02</b>	K1MIS_U16, K1MIS_U19, K1MIS_U27, K1MIS_U32SAD, K1MIS_U34SAD, K1MIS_U35SAD	C2, C3, C4, C6	La1-La10	2,3,4
<b>PEK_U03</b>	K1MIS_U16, K1MIS_U19, K1MIS_U27, K1MIS_U32SAD, K1MIS_U34SAD, K1MIS_U35SAD	C5, C6	La3-La10	2,3,4
<b>PEK_K01 (kompetencje)</b>	K1MIS_K01, K1MIS_K05	C1, C2, C3, C4, C5, C6	Wy1-Wy10 La1-La10	1,2,3,4
<b>PEK_K02</b>	K1MIS_K01, K1MIS_K05	C1, C2, C3, C4, C5, C6	Wy1-Wy10 La1-La10	1,2,3,4

\*\* - z tabeli powyżej