

Nazwa przedmiotu <i>Metody statystyczne w eksploracji danych</i> <i>Statistical methods in data mining</i>		Kod ECTS <i>3.1.KRK.12TY.MSED</i>		
Nazwa jednostki prowadzącej przedmiot Uniwersytet Opolski, Wydział Matematyki, Fizyki i Informatyki, Instytut Matematyki i Informatyki				
Studia				
	Kierunek	stopień	tryb	specjalność
	<i>Matematyka</i>	<i>Drugi</i>	<i>Stacjonarne</i> <i>Niestacjonarne *</i>	
Nazwisko osoby prowadzącej (osób prowadzących) Pracownicy Zakładu Metod Stochastycznych				
Formy zajęć, sposób ich realizacji i przypisana im liczba godzin		Liczba punktów ECTS: 6		
A. Formy zajęć • wykład (W), • konwersatorium (K),		<i>Bilans nakładu pracy przeciętnego studenta:</i> • 4 godz. – wstępny przegląd literatury [^{*)} 4] • 15×2 godz. = 30 godz. – udział w wykładach [^{*)} 18] • 15×2 godz. = 30 godz. – udział w konwersatoriach [^{*)} 18] • 15×1 godz. = 15 godz. – analiza i przyswojenie treści wykładu [^{*)} 21] • 7 × 1 godz. = 7 godz. – udział w konsultacjach do wykładu [^{*)} 2] • 15×2 godz. = 30 godz. – przygotowanie do konwersatoriów [^{*)} 36] • 8× 1 godz.= 8 godz. – udział w konsultacjach do konwersatorium [^{*)} 3] • 12 godz. – przygotowanie do sprawdzianów pisemnych na konwersatoriach [^{*)} 28] • 14 godz. – przygotowanie referatu/raportu/projektu zaliczeniowego [^{*)} 20]		
B. Sposób realizacji • zajęcia w sali wykładowej/dydaktycznej				
C. Liczba godzin Wykład – 30 godzin Konwersatorium – 30 godzin *) Studia niestacjonarne: Wykład – 18 godz. (2T+16Z) Konwersatorium – 18 godzin		Łączny nakład pracy studenta: 150 godzin, co odpowiada 6 pkt. ECTS w tym • nakład pracy związany z zajęciami wymagającymi bezpośredniego udziału nauczycieli akademickich: 30+30+7+8=75 godz., co odpowiada 3 pkt. ECTS; • nakład pracy związany z zajęciami o charakterze praktycznym: 30+30+8+12+14 = 94 godz., co odpowiada 3,5 pkt. ECTS *) na studiach niestacjonarnych: • nakład pracy związany z zajęciami wymagającymi bezpośredniego udziału nauczycieli akademickich: 18+18+2+3=41 godz., co odpowiada 2 pkt. ECTS; • nakład pracy związany z zajęciami o charakterze praktycznym: 18+36+3+28+20 = 105 godz., co odpowiada 4 pkt ECTS		
Status przedmiotu • specjalnościowy/do wyboru		Język wykładowy Polski (możliwość realizacji w języku angielskim)		
Metody dydaktyczne • wykład / wykład problemowy / wykład z prezentacją multimedialną • ćwiczenia audytoryjne: dyskusja / rozwiązywanie zadań		Forma i sposób zaliczenia oraz podst. kryteria oceny lub wymagania egzaminacyjne <i>Na ogólnych zasadach określonych w programie kształcenia, a w szczególności</i> A. Sposób zaliczenia • zaliczenie na ocenę (W) • zaliczenie z oceną (K) B. Formy zaliczenia • (W) ustalenie oceny na podstawie referatu/raportu/projektu zaliczeniowego; • (K) ustalenie oceny zaliczeniowej na podstawie ocen cząstkowych otrzymywanych w trakcie trwania semestru za wystąpienia ustne i za prace pisemne; C. Podstawowe kryteria • (W) uzyskanie pozytywnej oceny; • (K) uzyskanie pozytywnej oceny końcowej.		
Określenie przedmiotów wprowadzających wraz z wymogami wstępnymi Należy określić: A. Wymagania formalne: zaliczenie kursu Prawdopodobieństwo i statystyka B. Wymagania wstępne:				
Cele przedmiotu <i>Omówienie podstawowych metod statystyki matematycznej stosowanych w eksploracji danych (ang. data mining)</i>				

Treści programowe

A. Problematyka wykładu/B. Problematyka konwersatorium:

Systemy uczące się pod nadzorem. Liniowe metody klasyfikacji, dyskryminacja oparta na regresji liniowej i logistycznej. Metody klasyfikacji oparte na rozkładach prawdopodobieństwa. Klasyfikator bayesowski i metoda największej wiarygodności. Metody klasyfikacji oparte na nieparametrycznej estymacji rozkładów prawdopodobieństwa. Drzewa klasyfikatorów.

Systemy uczące się bez nadzoru. Analiza składowych głównych. Analiza czynnikowa. Podobieństwo, odmienność i odległość między obiektami. Skalowanie wielowymiarowe. Analiza skupień.

Wykaz literatury

A. Literatura wymagana

A.1. wykorzystywana podczas zajęć/A.2. studiowana samodzielnie przez studenta

1. J. Koronacki, J. Ćwik, Statystyczne systemy uczące się.
2. D.T. Larose, Metody i modele eksploracji danych.

B. Literatura uzupełniająca

1. M. Krzyśko, W. Wołyński, T. Górecki i M. Skorzybut, Systemy uczące się. Rozpoznawanie wzorców, analiza skupień i redukcja wymiarowości.
2. D. Hand, H. Mannila, P. Smyth, Eksploracja danych.

Efekty kształcenia	Wiedza			
	Symb.	Efekt	Metoda weryfikacji	Odniesienie
	W01	Znajomość podstawowych metod statystycznych stosowanych w eksploracji danych: metody klasyfikacji, analiza regresji, metody rzutowania, analiza skupień.	referat/ wystąpienia ustne/ sprawdziany pi- semne	K_W04,05
	Umiejętności:			
	Symb.	Efekt	Metoda weryfikacji	Odniesienie
	U01	Umiejętność przeprowadzenia klasyfikacji algorytmami LDA i QDA.	referat/ wystąpienia ustne/ sprawdziany pi- semne	K_U19-m4
	U02	Umiejętność stosowania podstawowych metod uczenia się bez nadzoru.	sprawdziany pi- semne	K_U19-m4
	Kompetencje społeczne (postawy)			
	Symb.	Efekt	Metoda weryfikacji	Odniesienie
	K01	zna ograniczenia własnej wiedzy i rozumie potrzebę dalszego kształcenia	konwersacja	K_K01
K02	potrafi precyzyjnie formułować pytania, zarówno werbalnie w trakcie zajęć jak i na potrzeby agregatów wyszukujących i naukowych baz danych, służące pogłębieniu własnego zrozumienia danego tematu lub odnalezieniu brakujących elementów rozumowania	referat	K_K02,06	

Kontakt:

Wykaz numerów telefonicznych i adresów mailowych pracowników znajduje się na stronie Instytutu Matematyki i Informatyki:
www.math.uni.opole.pl