

Karta przedmiotu

Nazwa i kod przedmiotu	Data Mining - eksploracja danych, PG_00156751						
Kierunek studiów	Informatyka i ekonometria (O)						
Data rozpoczęcia studiów	październik 2024 r.		Rok akademicki realizacji przedmiotu		2025/2026		
Poziom kształcenia	I stopnia - licencjackie		Grupa zajęć		Grupa zajęć fakultatywnych Grupa zajęć powiązanych z prowadzonymi badaniami naukowymi w dziedzinie nauki związanej z kierunkiem - profil ogólnoakademicki		
Forma studiów	stacjonarne		Sposób realizacji		na uczelni		
Rok studiów	2		Język wykładowy		polski		
Semestr studiów	4		Liczba punktów ECTS		7.0		
Profil kształcenia	ogólnoakademicki		Forma zaliczenia		egzamin		
Jednostka prowadząca	Rektor -> Wydział Zarządzania -> Katedra Statystyki						
Imię i nazwisko wykładowcy (wykładowców)	Odpowiedzialny za przedmiot		dr hab. Krzysztof Najman				
	Prowadzący zajęcia z przedmiotu						
Formy zajęć	Forma zajęć	Wykład	Ćwiczenia	Laboratorium	Projekt	Seminarium	RAZEM
	Liczba godzin zajęć	30.0	30.0	15.0	0.0	0.0	75
	W tym liczba godzin zajęć na odległość: 0.0						
Aktywność studenta i liczba godzin pracy	Aktywność studenta	Udział w zajęciach dydaktycznych, objętych planem studiów		Udział w konsultacjach		Praca własna studenta	RAZEM
	Liczba godzin pracy studenta	75		30.0		70.0	175
Cel przedmiotu	Celem przedmiotu jest zapoznanie studentów z podstawowymi metodami Eksploracyjnej Analizy Danych i jej praktycznymi zastosowaniami.						

Efekty uczenia się przedmiotu	Efekt kierunkowy	Efekt z przedmiotu	Sposób weryfikacji i oceny efektu
	[liEL3_K02] Potrafi komunikować się z otoczeniem na tematy z zakresu informatyki i ekonometrii w miejscu pracy i poza nim, przekazywać swoją wiedzę oraz dzielić się swoimi umiejętnościami za pomocą różnych środków przekazu.	Potrafi dyskutować, komunikować się i dzielić wiedzą w zakresie teorii i praktyki stosowania modeli Data Mining.	[SK1] oral statement/conversation/discussion [SK2] presentation/project/paper/report [SK4] test/exam - oral or written
	[liEL3_U02] Potrafi pozyskiwać podstawowe informacje o procesach i zjawiskach ekonomicznych w drodze obserwacji bezpośredniej, planowanego eksperymentu lub kwerendy baz danych oraz gromadzić je i przetwarzać za pomocą nowoczesnych narzędzi informatycznych.	Student potrafi zbierać dane statystyczne ze źródeł pierwotnych i wtórnych na potrzeby realizacji badań społecznych z zastosowaniem modeli Data Mining.	[SU2] presentation/project/paper/report
	[liEL3_W05] Ma zaawansowaną wiedzę o źródłach danych społeczno-ekonomicznych, ich bazach oraz sposobie ich tworzenia.	Student zna źródła i metody wyszukiwania danych wtórnych wykorzystywanych w badaniach społecznych analizowanych metodami Data Mining.	[SW4] test/exam - oral or written [SW1] oral statement/conversation/discussion [SW2] presentation/project/paper/report
	[liEL3_W04] Zna zaawansowane metody matematyczne, statystyczne, ekonometryczne oraz informatyczne umożliwiające pozyskiwanie, przetwarzanie i analizę danych odzwierciedlających funkcjonowanie i wzrost gospodarki narodowej i jej składowych oraz zjawisk i procesów zachodzących w ich otoczeniu.	Ma pogłębioną wiedzę o zaawansowanych metodach analizy danych z zastosowaniem wybranych modeli Data Mining.	[SW4] test/exam - oral or written [SW2] presentation/project/paper/report
	[liEL3_K01] Rozumie potrzebę ustawicznego uzupełniania i pogłębiania nabytej wiedzy; jest otwarty na nowe idee i metody nauczania.	Rozumie potrzebę stałego uzupełniania wiedzy w zakresie teorii i zastosowań Data Mining.	[SK1] oral statement/conversation/discussion [SK2] presentation/project/paper/report
Treści przedmiotu	[liEL3_U05] Potrafi dokonać opisu statystycznego zaawansowanych kategorii ekonomicznych, sformułować oraz zweryfikować proste hipotezy odnośnie do ich kształtowania się.	Student potrafi samodzielnie przeprowadzić analizę wybranych zjawisk społecznych z zastosowaniem wybranych modeli Data Mining.	[SU1] oral statement/conversation/discussion [SU2] presentation/project/paper/report
	1. Podstawowe pojęcia, cel i funkcje eksploracji danych. Eksploracja danych (EDA), data mining, odkrywanie wiedzy w bazach danych (KDD), eksploracja danych jako proces, etapy procesu eksploracji danych. 2. Wielowymiarowe zbiory danych, pojęcie Big data. Oprogramowanie analityczne. 3. Zadania eksploracji danych: opis, estymacja, przewidywanie, klasyfikacja, grupowanie, odkrywanie reguł. Przykłady zastosowań eksploracji danych w badaniach społeczno-ekonomicznych. Źródła danych. 4. Wstępna obróbka danych: przygotowanie danych do analizy, problem wyboru zmiennych i przypadków, redukcja wymiaru przestrzeni cech, problem danych brakujących, graficzne metody identyfikacji jednostek oddalonych, analityczne metody identyfikacji jednostek oddalonych. Procedury normalizacji zmiennych i cech. 5. Podstawowe metody tabelaryczne i graficzne EDA. 6. Korelacja cząstkowa i wieloraka. Macierze współzależności. 7. Reguły asocjacji, pojęcie asocjacji, reguły binarne i ilościowe, jedno i wielowymiarowe reguły asocjacji, jedno i wielopoziomowe reguły asocjacji, wsparcie i ufnosć reguły asocjacji, elementy analizy koszykowej, zagadnienie koszyka, analiza sekwencji, analiza skojarzeń, inne zastosowania reguł asocjacji w EDA. Grupowanie, klasyfikacja, analiza skupień. 8. Proces grupowania i klasyfikacji danych. Podobieństwo obiektów i jego pomiar. Pomiar zróżnicowania obiektów według cech ilościowych - miary odległości. Pomiar podobieństwa obiektów według cech ilościowych. Podobieństwo cech i jego pomiar. Pomiar podobieństwa cech ilościowych na podstawie korelacji. 9. Grupowanie hierarchiczne. Istota aglomeracyjnego grupowania hierarchicznego. Metoda najbliższego sąsiada, metoda najdalszego sąsiada, metoda średniej grupowej, metoda centroidalna, metoda mediany, metoda Warda. Schemat Lancea i Williamsa. Wizualizacja struktury grupowej. Ocena grupowania na podstawie dendrogramu. 10. Wybrane metody EDA: drzewa klasyfikacyjne, drzewa regresyjne, algorytm k-najbliższych sąsiadów.		
Wymagania wstępne i dodatkowe	Zaawansowana wiedza z zakresu algebry i statystyki.		
Sposoby i kryteria oceniania osiągniętych efektów uczenia się	Sposób oceniania (składowe)	Próg zaliczeniowy	Składowa oceny końcowej
	projekt semestralny	51.0%	50.0%
	egzamin pisemny	51.0%	50.0%

Zalecana lista lektur	Podstawowa lista lektur	1. M.Lasek, Data mining, Biblioteka Menedżera i Bankowca, Warszawa, 2002 2. D.T.Larose, Okrywanie wiedzy z danych. Wprowadzenie do eksploracji danych, Wydawnictwo Naukowe PWN, Warszawa, 2006 3. E.Gatnar, Nieparametryczna metoda dyskryminacji i regresji, Wydawnictwo Naukowe PWN, Warszawa 2001 4. A. Balicki, Statystyczna analiza wielowymiarowa i jej zastosowania społeczno-ekonomiczne, Wydawnictwo UG, Gdańsk, 2009
	Uzupełniająca lista lektur	Nong Ye, The Handbook of Data Mining, Lawrence Erlbaum Associates, New Jersey, 2003
	Adresy eZasobów	
Przykładowe zagadnienia/ przykładowe pytania/ realizowane zadania		
Praktyki zawodowe w ramach przedmiotu	Nie dotyczy	

Dokument wygenerowany elektronicznie. Nie wymaga pieczęci ani podpisu.