



Karta przedmiotu

Nazwa i kod przedmiotu	Hurtownie danych, PG_00044140						
Kierunek studiów	Matematyka						
Data rozpoczęcia studiów	październik 2022 r.	Rok akademicki realizacji przedmiotu			2023/2024		
Poziom kształcenia	II stopnia	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	3	Liczba punktów ECTS			4.0		
Profil kształcenia	ogólnoakademicki	Forma zaliczenia			zaliczenie		
Jednostka prowadząca	Wydział Fizyki Technicznej i Matematyki Stosowanej						
Imię i nazwisko wykładowcy (wykładowców)	Odpowiedzialny za przedmiot	dr hab. Paweł Pilarczyk					
	Prowadzący zajęcia z przedmiotu	dr hab. Paweł Pilarczyk mgr inż. Michał Krzemiński					
Formy zajęć i metody nauczania	Forma zajęć	Wykład	Ćwiczenia	Laboratorium	Projekt	Seminarium	RAZEM
	Liczba godzin zajęć	30.0	0.0	30.0	0.0	0.0	60
	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	60	5.0		35.0		100
Cel przedmiotu	Teoretyczne i praktyczne wprowadzenie do tematyki analizy i eksploracji danych oraz hurtowni danych.						

Efekty uczenia się przedmiotu	Efekt kierunkowy	Efekt z przedmiotu	Sposób weryfikacji i oceny efektu
	[K7_U13] rozumie matematyczne podstawy analizy algorytmów i procesów obliczeniowych, potrafi konstruować algorytmy o dobrych własnościach numerycznych, służące do rozwiązywania typowych i nietypowych problemów matematycznych	Student zna matematyczne podstawy poznanych algorytmów analizy danych, w tym uczenia maszynowego, i umie je właściwie zastosować w konkretnych zadaniach.	[SU4] Ocena umiejętności korzystania z metod i narzędzi [SU1] Ocena realizacji zadania
	[K7_U08] zna podstawowe rozkłady probabilistyczne i ich własności; potrafi je stosować w zagadnieniach praktycznych, orientuje się w podstawach statystyki (zagadnienia estymacji i testowanie hipotez) oraz w podstawach statystycznej obróbki danych	Student potrafi wykonać podstawową analizę statystyczną danych numerycznych ilościowych i jakościowych.	[SU3] Ocena umiejętności wykorzystania wiedzy uzyskanej w ramach przedmiotu [SU2] Ocena umiejętności analizy informacji
	[K7_W10] zna metody numeryczne stosowane do znajdowania przybliżonych rozwiązań zagadnień matematycznych (na przykład równań różniczkowych) stawianych przez dziedziny stosowane (np. technologie przemysłowe, zarządzanie itp.)	Student potrafi dobrać metody statystyczne oraz metody uczenia maszynowego, a także środowisko programistyczne odpowiednie do modelowania rzeczywistych problemów.	[SW3] Ocena wiedzy zawartej w opracowaniu tekstowym i projektowym
	[K7_K02] potrafi precyzyjnie formułować pytania, służące pogłębieniu własnego zrozumienia danego tematu lub odnalezieniu brakujących elementów rozumowania, rozumie potrzebę popularnego przedstawiania laikom wybranych osiągnięć matematyki wyższej	Student umie zaplanować projekt eksploracji danych, a następnie omówić jego wyniki.	[SK4] Ocena umiejętności komunikacji, w tym poprawności językowej
[K7_W08] zna zaawansowane techniki obliczeniowe, wspomagające pracę matematyka i rozumie ich ograniczenia	Student zna metody matematyczne i narzędzia programistyczne stosowane w analizie danych.	[SW1] Ocena wiedzy faktograficznej	
Treści przedmiotu	Wykład: wprowadzenie do tematyki eksploracji danych (data mining) i odkrywania wiedzy w bazach danych. Przygotowanie i czyszczenie danych, wstępna eksploracja danych, proces CRISP-DM. Statystyczna analiza danych i uczenie maszynowe. Metody klasyfikacji i grupowania danych, odkrywanie zależności w danych. Hurtownie danych, modelowanie wielowymiarowe, OLAP. Laboratorium: praktyczna analiza i eksploracja danych przy wykorzystaniu Pythona lub języka R.		
Wymagania wstępne i dodatkowe	Podstawowa umiejętność programowania w R i w Pythonie. Znajomość podstawowych zagadnień i metod statystyki.		
Sposoby i kryteria oceniania osiągnięć uczenia się	Sposób oceniania (składowe)	Próg zaliczeniowy	Składowa oceny końcowej
	Projekty grupowe (laboratorium)	60.0%	50.0%
	Kartkówki (wykład)	60.0%	50.0%
Zalecana lista lektur	Podstawowa lista lektur	Daniel T. Larose, Chantal D. Larose. Discovering Knowledge in Data. An Introduction to Data Mining, 2nd Ed., 2014. Daniel T. Larose. Data Mining. Metody i modele eksploracji danych. Warszawa, Wydawnictwo Naukowe PWN, 2012. https://www.ibuk.pl/fiszka/67154/metody-i-modele-eksploracji-danych.html Jiawei Han, Micheline Kamber, Jian Pei. Data Mining. Concepts and Techniques. 3rd Ed. Elsevier, 2011.	

	<p>Uzupełniająca lista lektur</p>	<p>Wes McKinney. Python w analizie danych. Przetwarzanie danych za pomocą pakietów Pandas i NumPy oraz środowiska IPython. Gliwice, Helion, ©2018.</p> <p>Joel Grus. Data science od podstaw. Analiza danych w Pythonie. Gliwice, Helion, ©2018.</p> <p>Jacek Rumiński. Wprowadzenie do hurtowni i eksploracji danych. Gdańsk, Wydawnictwo Politechniki Gdańskiej, 2015.</p>
	<p>Adresy eZasobów</p>	<p>Adresy na platformie eNauczanie: Hurtownie danych 2023 - Moodle ID: 29955 https://enauczanie.pg.edu.pl/moodle/course/view.php?id=29955</p>
<p>Przykładowe zagadnienia/ przykładowe pytania/ realizowane zadania</p>	<p>Metody przetwarzania wstępnego danych.</p> <p>Czym różni się hurtownia danych od bazy danych?</p> <p>Metody nadzorowanej klasyfikacji danych.</p> <p>Sieci Kohonena i ich związek z sieciami neuronowymi.</p>	
<p>Praktyki zawodowe w ramach przedmiotu</p>	<p>Nie dotyczy</p>	