

Karta przedmiotu

Nazwa i kod przedmiotu	Narzędzia inżynierii danych , PG_00045768						
Kierunek studiów	Fizyka Techniczna						
Data rozpoczęcia studiów	luty 2025 r.		Rok akademicki realizacji przedmiotu		2025/2026		
Poziom kształcenia	II stopnia		Grupa zajęć		Grupa zajęć specjalnościowych Grupa zajęć powiązanych z prowadzonymi badaniami naukowymi w dziedzinie nauki związanej z kierunkiem - profil ogólnoakademicki		
Forma studiów	stacjonarne		Sposób realizacji		na uczelni		
Rok studiów	1		Język wykładowy		polski		
Semestr studiów	2		Liczba punktów ECTS		3.0		
Profil kształcenia	ogólnoakademicki		Forma zaliczenia		zaliczenie		
Jednostka prowadząca	Wydziały Politechniki Gdańskiej -> Wydział Fizyki Technicznej i Matematyki Stosowanej -> Instytut Fizyki i Informatyki Stosowanej						
Imię i nazwisko wykładowcy (wykładowców)	Odpowiedzialny za przedmiot		dr inż. Patryk Jasik				
	Prowadzący zajęcia z przedmiotu		dr inż. Patryk Jasik				
Formy zajęć	Forma zajęć	Wykład	Ćwiczenia	Laboratorium	Projekt	Seminarium	RAZEM
	Liczba godzin zajęć	15.0	0.0	30.0	0.0	0.0	45
	W tym liczba godzin zajęć na odległość: 0.0						
	Adresy kursu na platformie eNauczanie: Moodle ID: 2379 Narzędzia inżynierii danych 2025 https://enauczanie.pg.edu.pl/2025/course/view.php?id=2379						
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	45		5.0		25.0	75
Cel przedmiotu	Zapoznanie studentów z wybranymi systemami informatycznymi służącymi do przetwarzania i analizy danych, w szczególności dużych wolumenów danych. Zdobywanie wiedzy i umiejętności związanych z wykorzystywaniem narzędzi inżynierii danych w praktyce.						
Efekty uczenia się przedmiotu	Efekt kierunkowy		Efekt z przedmiotu		Sposób weryfikacji i oceny efektu		
	[K7_U02] posiada pogłębioną umiejętność programowania w wybranym języku oraz stosowania pakietów oprogramowania		Student posiada pogłębioną umiejętność programowania w języku Python z wykorzystaniem jego pakietów w ramach ekosystemu stworzonego przy użyciu wybranych narzędzi inżynierii danych.		[SU4] Assessment of ability to use methods and tools		
	[K7_K03] potrafi komunikować się, zaprezentować efekty swojej pracy, przekazać informacje w sposób powszechnie zrozumiały		Student potrafi zaprezentować wyniki swojej pracy uzyskane przy użyciu wybranych narzędzi inżynierii danych.		[SK5] Assessment of ability to solve problems that arise in practice		
	[K7_U04] potrafi formułować i testować hipotezy związane z problemami badawczymi		Student potrafi formułować hipotezy związane z problemami badawczymi oraz weryfikować je przy użyciu wybranych narzędzi inżynierii danych.		[SU4] Assessment of ability to use methods and tools		
	[K7_W04] posiada pogłębioną znajomość metod matematycznych, numerycznych i symulacyjnych stosowanych przy opisie i modelowaniu zjawisk fizycznych		Student potrafi wykorzystywać wybrane narzędzia inżynierii danych do opisu i modelowania zjawisk fizycznych bazując na posiadanej pogłębionej znajomości metod matematycznych, numerycznych i symulacyjnych.		[SW3] Assessment of knowledge contained in written work and projects		

Treści przedmiotu	Treści przedmiotu - wykład Przegląd dostępnych narzędzi inżynierii danych. Zalety i wady wybranych narzędzi inżynierii danych. Apache Hadoop jako narzędzie do rozproszonego składowania i przetwarzania wielkich zbiorów danych przy pomocy klastrów komputerowych. Podstawy teoretyczne działania Apache Spark. Podstawy działania sieci neuronowych typu PINN (Physics Informed Neural Networks). Treści przedmiotu - laboratoria Trenowanie modeli uczenia maszynowego z wykorzystaniem Apache Spark. Rozwiązywanie wybranych zagadnień fizycznych przy użyciu sieci neuronowych typu PINN. Przykładowe możliwości wykorzystania oprogramowania H2O w Big Data. Wprowadzenie do MLflow. Wprowadzenie do Apache Airflow. Wprowadzenie do AutoGluon.		
Wymagania wstępne i dodatkowe	Znajomość metod eksploracji danych. Umiejętność programowania w języku Python.		
Sposoby i kryteria oceniania osiągniętych efektów uczenia się	Sposób oceniania (składowe)	Próg zaliczeniowy	Składowa oceny końcowej
	Ukończenie wybranych kursów specjalistycznych na portalu DataCamp	60.0%	70.0%
	Obecność na zajęciach	80.0%	30.0%
Zalecana lista lektur	Podstawowa lista lektur	Viktor Mayer-Schonberger, Kenneth Cukier, Big Data. Rewolucja, która zmieni nasze myślenie, pracę i życie, MT Biznes Tadeusz Morzy, Eksploracja danych. Metody i algorytmy, PWN Tom White, Hadoop. Kompletny przewodnik, Helion Holden Karau, Andy Konwinski, Patrick Wendell, Matei Zaharia, Poznajemy Sparka. Błyskawiczna analiza danych, PWN Rachel Schutt, Cathy O'Neil, Badanie danych. Raport z pierwszej linii działań, Helion	
	Uzupełniająca lista lektur	Nathan Marz, James Warren, Big Data. Najlepsze praktyki budowania skalowalnych systemów obsługi danych w czasie rzeczywistym, Helion Russell Jurney, Zwinna analiza danych Apache Hadoop dla każdego, Helion Sandy Ryza, Uri Laserson, Sean Owen, Josh Wills, Spark. Zaawansowana analiza danych, Helion	
	Adresy eZasobów		
Przykładowe zagadnienia/ przykładowe pytania/ realizowane zadania	DataCamp - Introduction to PySpark (4h) DataCamp - Introduction to MLflow (4h) DataCamp - Introduction to Apache Airflow in Python (4h)		
Zajęcia praktyczne w ramach przedmiotu	Nie dotyczy		

Dokument wygenerowany elektronicznie. Nie wymaga pieczęci ani podpisu.