

Karta przedmiotu

Nazwa i kod przedmiotu	BIG DATA, PG_00068679						
Kierunek studiów	Analityka gospodarcza						
Data rozpoczęcia studiów	październik 2025 r.		Rok akademicki realizacji przedmiotu		2026/2027		
Poziom kształcenia	II stopnia		Grupa zajęć		Grupa zajęć fakultatywnych Grupa zajęć specjalnościowych Grupa zajęć powiązanych z prowadzonymi badaniami naukowymi w dziedzinie nauki związanej z kierunkiem - profil ogólnoakademicki		
Forma studiów	niestacjonarne		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	Wydziały Politechniki Gdańskiej -> Wydział Zarządzania i Ekonomii -> Katedra Statystyki i Ekonometrii						
Imię i nazwisko wykładowcy (wykładowców)	Odpowiedzialny za przedmiot						
	Prowadzący zajęcia z przedmiotu						
Formy zajęć i metody nauczania	Forma zajęć	Wykład	Ćwiczenia	Laboratorium	Projekt	Seminarium	RAZEM
	Liczba godzin zajęć	8.0	0.0	24.0	0.0	0.0	32
	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	32		4.0		64.0	100
Cel przedmiotu	Demonstruje w pogłębionym stopniu wszechstronne przygotowanie do analizy dużych zbiorów danych						
Efekty uczenia się przedmiotu	Efekt kierunkowy		Efekt z przedmiotu		Sposób weryfikacji i oceny efektu		
	[K7_W06] zna i rozumie zasady oceny wiarygodności wykorzystywanych danych, wykorzystując pogłębioną wiedzę specjalistyczną z zakresu analityki gospodarczej.		zna zaawansowane technologie do przetwarzania dużych zbiorów danych, oceniając ich wiarygodność i jakość na podstawie pogłębionej wiedzy z zakresu analityki gospodarczej w celu przygotowania danych do rozwiązywania złożonych problemów decyzyjnych		[SW1] Ocena wiedzy faktograficznej		
	[K7_U01] potrafi tworzyć innowacyjne rozwiązania złożonych i nieustrukturyzowanych procesów uwzględniając nieprzewidywalne warunki otoczenia poprzez syntezę informacji pochodzących z różnych źródeł.		potrafi formułować innowacyjne rozwiązania złożonych problemów na podstawie analizy dużych i nieustrukturyzowanych zbiorów danych, integrując informacje z różnych źródeł oraz stosując zaawansowane algorytmy uczenia maszynowego w warunkach niepewności.		[SU3] Ocena umiejętności wykorzystania wiedzy uzyskanej w ramach przedmiotu		
Treści przedmiotu	Przegląd Big Data. Rodzaje danych cyfrowych, wprowadzenie do Big Data Narzędzia do programowania Big Data (np. Hadoop, MongoDB, Spark itp.). Używanie Spark z R Ekstrakcja i integracja dużych zbiorów danych Przechowywanie dużych zbiorów danych; technologie obsługi dużych zbiorów danych - wprowadzenie do Hadoop HDFS (rozproszony system plików Hadoop) Dogłębne zrozumienie podstaw MapReduce i HBase Hadoop MapReduce w R; Integracja Hadoop i R - RHIPE; RHadoop Analiza danych z R i Hadoop - wstępne przetwarzanie danych, wizualizacja danych Analiza dużych zbiorów danych i uczenie maszynowe - nadzorowane i nienadzorowane algorytmy uczenia maszynowego. Spark Machine Learning z R Importowanie i eksportowanie danych z różnych baz danych (RMySQL, RSQLite, RHive, RHBase). Używanie SparkSQL z R Analiza dużych zbiorów danych z BigR Algorytmy głębokiego uczenia z R i H2O						

Wymagania wstępne i dodatkowe			
Sposoby i kryteria oceniania osiągniętych efektów uczenia się	Sposób oceniania (składowe)	Próg zaliczeniowy	Składowa oceny końcowej
	Egzamin	60.0%	60.0%
	Kolokwium	60.0%	40.0%
Zalecana lista lektur	Podstawowa lista lektur	Hamstra, M., & Zaharia, M. (2013). Learning Spark: lightning-fast big data analytics. O'Reilly & Associates Densmore, J. (2021). Data pipelines pocket reference. O'Reilly Media Drabas, T., & Lee, D. (2017). Learning PySpark. Packt Publishing Ltd Haines, S. (2022). Modern Data Engineering with Apache Spark: A Hands-on Guide for Building Mission-critical Streaming Applications. Apress	
	Uzupełniająca lista lektur	Warren, J., & Marz, N. (2015). Big Data: Principles and best practices of scalable realtime data systems. Simon and Schuster Ilijason, R. (2020). Beginning Apache Spark Using Azure Databricks: Unleashing Large Cluster Analytics in the Cloud. Apress	
	Adresy eZasobów	Adresy na platformie eNauczanie:	
Przykładowe zagadnienia/ przykładowe pytania/ realizowane zadania			
Praktyki zawodowe w ramach przedmiotu	Nie dotyczy		

Dokument wygenerowany elektronicznie. Nie wymaga pieczęci ani podpisu.