



Karta przedmiotu

Nazwa i kod przedmiotu	Business Data Analytics, PG_00053096						
Kierunek studiów	Inżynieria danych						
Data rozpoczęcia studiów	październik 2024 r.	Rok akademicki realizacji przedmiotu			2026/2027		
Poziom kształcenia	I stopnia - inżynierskie	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	3	Język wykładowy			angielski		
Semestr studiów	5	Liczba punktów ECTS			4.0		
Profil kształcenia	ogólnoakademicki	Forma zaliczenia			zaliczenie		
Jednostka prowadząca	Wydziały Politechniki Gdańskiej -> Wydział Zarządzania i Ekonomii -> Katedra Informatyki w Zarządzaniu						
Imię i nazwisko wykładowcy (wykładowców)	Odpowiedzialny za przedmiot	dr Nina Rizun					
	Prowadzący zajęcia z przedmiotu	dr Nina Rizun mgr Jaromir Durkiewicz					
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						
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	50.0	100	
Cel przedmiotu	Celem przedmiotu jest zapoznanie studentów z podstawami modelowania procesów biznesowych z wykorzystaniem standardowych notacji, ich analiza oraz optymalizacja.						

Efekty uczenia się przedmiotu	Efekt kierunkowy	Efekt z przedmiotu	Sposób weryfikacji i oceny efektu
	[K6_U06] zdobywa nową wiedzę, planując własny rozwój sprzyjający osiągnięciu wyznaczonych celów	Student samodzielnie poszerza wiedzę i kompetencje w zakresie analityki danych, wyszukując specjalistyczne materiały, dokumentację narzędzi analitycznych oraz przykłady zastosowań biznesowych, a następnie wykorzystuje je do planowania własnego rozwoju w obszarze analizy danych i podejmowania decyzji opartych na danych	[SU1] Ocena realizacji zadania [SU2] Ocena umiejętności analizy informacji [SU3] Ocena umiejętności wykorzystania wiedzy uzyskanej w ramach przedmiotu [SU4] Ocena umiejętności korzystania z metod i narzędzi [SU5] Ocena umiejętności zaprezentowania wyników realizacji zadania
	[K6_K03] wykazuje się zdolnością do krytycznego i analitycznego myślenia oraz integruje wiedzę z wielu dyscyplin w celu podejmowania efektywnych decyzji	Student wykazuje się zdolnością do krytycznej interpretacji wyników analiz danych, integruje wiedzę z zakresu analityki danych, ekonomii oraz zarządzania i wykorzystuje ją do formułowania wniosków oraz wspierania procesu podejmowania decyzji biznesowych	[SK1] Ocena umiejętności pracy w grupie [SK2] Ocena postępów pracy [SK3] Ocena umiejętności organizacji pracy [SK4] Ocena umiejętności komunikacji, w tym poprawności językowej [SK5] Ocena umiejętności rozwiązywania problemów występujących w praktyce
	[K6_W03] identyfikuje wiarygodne źródła informacji istotne dla analizowanych zagadnień	Student identyfikuje, pozyskuje oraz krytycznie ocenia dane pochodzące z różnych źródeł (np. bazy danych, raporty biznesowe, otwarte zbiory danych), a następnie wykorzystuje je do analizy i interpretacji problemów biznesowych z zastosowaniem podstawowych narzędzi analityki danych	[SW1] Ocena wiedzy faktograficznej [SW2] Ocena wiedzy zawartej w prezentacji [SW3] Ocena wiedzy zawartej w opracowaniu tekstowym i projektowym
Treści przedmiotu	<p>Treści przedmiotu - wykład</p> <ol style="list-style-type: none"> <li>1. Czym jest Business Data Analytics? CRISP-DM. Przygotowanie danych. Metody eksploracji danych</li> <li>2. Czym jest Text Analytics (BTA)? Text Analytics &amp; NLP. Zadania Text Analytics. Miary oceny wyszukiwania. Wyszukiwanie w oparciu o słowa kluczowe. Wyszukiwanie w oparciu o reprezentację wektorową. Techniki wstępnego przetwarzania dokumentów tekstowych. Prawo Zipfa</li> <li>3. Korpus .Własności reprezentacji wektorowej. Wyznaczanie macierzy częstości (Obliczanie wag).Reprezentacja binarna. Odwrotna częstość słowa (inverse-document frequency, IDF). Schemat wagowy TFIDF.Wektorowa reprezentacja dokumentu. Pomiar odległości/podobieństwa. Metody obliczania odległości.Metody obliczania podobieństwa. Analiza skupień grupowanie. Metody hierarchiczne aglomeracyjne. Przykład Metody hierarchicznej. Metody k- optymalizacyjne.</li> <li>4. Podobieństwo Kosinusowe. Odległość Euklidesowa vs. Podobieństwo Kosinusowe. Macierz sąsiedztwa i wizualizacja. Podobieństwa za pomocą grafów. Macierz sąsiedztwa. Stopień wężła grafu. Struktura społeczności sieci.</li> <li>5. Skalowanie wielowymiarowe (MDS): Motywacja. Cele. Formalny model. Stress (Goodness-of-fit). Analiza ukrytych grup semantycznych (LSA): Motywacja. Model wektorowy: Ograniczenia. Rozkład według wartości osobliwych (Singular Value Decomposition). LSA zmniejszenie wymiaru. Podobieństwo dokumentów.Podobieństwo wyrazów. Automatyczne modelowanie tematyki danych tekstowych. Latent Dirichlet Allocation (LDA).</li> <li>6. Wprowadzenie do Sentiment Analysis</li> <li>7. Structural Topic Modeling.</li> <li>8. SNA - wprowadzenie. Historia teorii analizy sieci społecznej. Teoria małego świata. Eksperyment Milgrama. Relacje jako sieć. Graf skierowany. Graf nieskierowany. Atrybuty krawędzi. Gęstość grafu. Osiągalność i odległość. Miary centralności</li> </ol> <p>Treści przedmiotu - laboratoria</p> <p>Podstawowa analiza tekstu w R Transformacja TF-IDF Analiza skupień (cluster analysis) w R Podobieństwo cosinusowe (cosine similarity) Grafy sieciowe w R Modelowanie tematów (Topic Modeling) Wielowymiarowe skalowanie (Multidimensional Scaling, MDS) oraz analiza semantyczna LSA (Latent Semantic Analysis) Analiza opinii i analiza sentymentu (Opinion Mining and Sentiment Analysis) Strukturalne modelowanie tematów (Structural Topic Modeling) Analiza sieci społecznych (Social Network Analysis) Podstawowe miary sieci (Basic Network Measures)</p>		

Wymagania wstępne i dodatkowe			
Sposoby i kryteria oceniania osiągniętych efektów uczenia się	Sposób oceniania (składowe)	Próg zaliczeniowy	Składowa oceny końcowej
	Projekt	60.0%	40.0%
	Kolokwium końcowe	60.0%	40.0%
	Ćwiczenia na zajęciach	60.0%	20.0%
Zalecana lista lektur	Podstawowa lista lektur	1. Provost, Foster, Fawcett, Tom. Data Science for Business: What you need to know about data mining and data-analytic thinking. " O'Reilly Media, Inc.", 2019 2. Sharda, R., Delen, D., & Turban, E. Business intelligence: a managerial perspective on analytics. Pearson. 2018 3. Siegel, E. Predictive analytics: The power to predict who will click, buy, lie, or die. John Wiley & Sons, 2016. 338 pp. 4. James, G., Witten, D., Hastie, T., Tibshirani, R. An introduction to statistical learning with applications in R. Springer, 2018.	
	Uzupełniająca lista lektur	1. Han, J., Kamber, M., Pei, J. Data mining concepts and techniques. Morgan Kaufmann, 2021. 2. Hastie, T., Tibshirani, R., Friedman, J. The Elements of Statistical Learning, 2nd edition. Springer, 2022. 3. James, G. et al. An introduction to statistical learning. Springer, 2013. 426 pp. 4. Murphy, K. Machine Learning: A Probabilistic Perspective. MIT Press, 2019	
	Adresy eZasobów		
Przykładowe zagadnienia/ przykładowe pytania/ realizowane zadania	Wybierz główne właściwości odległości euklidesowej Wybierz odpowiednią wartość miary podobieństwa cosinusa między dokumentem 1 a dokumentem 2 Algorytm K-średnich pozwala ... Jaka jest różnica między odległością euklidesową a miarą podobieństwa cosinusów? Osoba o wysokiej centralności wektora własnego to ... Proszę obliczyć IDF dla warunków z następującego Korpusu Zbuduj wykres prawa Zipfa dla następującego Korpusu Podaj ogólną charakterystykę aktorów następującej sieci społecznościowej		
Zajęcia praktyczne w ramach przedmiotu	Nie dotyczy		

Dokument wygenerowany elektronicznie. Nie wymaga pieczęci ani podpisu.