WYDZIAŁ INFORMATYKI I TELEKOMUNIKACJI/KATEDRA PODSTAW INFORMATYKI

KARTA PRZEDMIOTU

Nazwa przedmiotu w języku polskim : **Big Data**
Nazwa przedmiotu w języku angielskim : **Big Data**
Kierunek studiów : **Informatyka algorytmiczna**
Specjalność (jeśli dotyczy) : **—**
Poziom i forma studiów : **II stopień, stacjonarna**
Rodzaj przedmiotu : **wybieralny**
Język wykładowy : **polski**
Cykl kształcenia od : **2023/2024**
Kod przedmiotu : **W04INA-SM0126G**
Grupa zajęć : **TAK**

	Wykład	Ćwiczenia	Laboratorium	Projekt	Seminarium
Liczba godzin zajęć zorganizowanych w Uczelni (ZZU)	30	15	15		
Liczba godzin całkowitego nakładu pracy studenta (CNPS)	50	35	40		
Forma zaliczenia	zaliczenie na ocenę				
Dla grupy kursów zaznaczyć kurs końcowy	X				
Liczba punktów ECTS	2	1.4	1.6		
w tym liczba punktów odpowiadająca zajęciom o charakterze praktycznym (P)		3			
w tym liczba punktów ECTS odpowiadająca zajęciom wymagającym bezpośredniego udziału nauczycieli lub innych osób prowadzących zajęcia (BU)	2.72				

WYMAGANIA WSTĘPNE W ZAKRESIE WIEDZY, UMIEJĘTNOŚCI I KOMPETENCJI SPOŁECZNYCH

Znajomość podstaw rachunku prawdopodobieństwa oraz podstaw programowania w języku Python

CELE PRZEDMIOTU

C1 Zapoznanie się z podstawowymi metodami oraz algorytmami przetwarzania dużych kolekcji danych

C2 Pogłębienie zrozumienia zagadnień teoretycznych omawianych na wykładzie

C3 Implementacja wybranych algorytmów dla Big Data

PRZEDMIOTOWE EFEKTY UCZENIA SIĘ

Z zakresu wiedzy studenta:

W1 Ma opanowany aparat teoretyczny potrzebny do zrozumienia omawianych na wykładzie algorytmów

W2 Zna podstawowe metody i algorytmy stosowane do analizy dużych zasobów danych

W3 Zna podstawowe składniki systemów informatycznych przetwarzających duże zasoby danych

Z zakresu umiejętności studenta:

U1 Potrafi badać eksperymentalnie algorytmy wyspecjalizowane w przetwarzaniu dużych zasobów danych

U2 Umie dobrać właściwy algorytm do budowanej aplikacji

U3 Umie korzystać z aktualnych osiągnięć w dziedzinie Bug Data

Z zakresu kompetencji społecznych studenta:

K1 Jest wyczulony na pojęcie skalowalności i efektywności rozwiązań algorytmicznych

K2 Potrafi korzystać z najnowszych osiągnięć teoretycznych z obszaru Big Data

K3 Potrafi stosować szkice i losowe próbki danych

TREŚCI PROGRAMOWE

Forma zajęć - wykład		
Wy1	Funkcje haszujące i filtry Blooma	4h
Wy2	Przetwarzanie strumieni danych	4h
Wy3	Często występujące elementy	4h
Wy4	Zliczanie i algorytm HyperLogLog	4h
Wy5	Odległość i podobieństwo	4h
Wy6	Locally sensitive hashing	4h
Wy7	Model obliczeń Map-Reduce	4h
Wy8	Google Page Rank	2h
	Suma godzin	30h
Forma zajęć - ćwiczenia		
Ćw1	Elementy probabilistyki (k-niezależność)	2h
Ćw2	Konstruowanie funkcji haszujących	2h
Ćw3	Algorytmy strumieniowe	4h
Ćw4	Statystyki pozycyjne i zliczanie	4h
Ćw5	Własności i ograniczenia modelu MapReduce	2h
Ćw6	Przestrzenie metryczne i Odległość Jaccarda	1h
	Suma godzin	15h
Forma zajęć - laboratorium		
Lab1	Konstruowanie i testowanie funkcji haszujących	3h
Lab2	Generowanie losowych próbek ze strumienia i z ograniczonego okna	3h
Lab3	Zliczanie elementów: HyperLogLog, algorytmy oparte o statystyki pozycyjne	3h
Lab4	Porównywanie tekstów	3h
Lab5	Zliczenie słów w modelu MapReduce	3h
	Suma godzin	15h

STOSOWANE NARZĘDZIA DYDAKTYCZNE

1. Wykład tradycyjny
2. Wykład multimedialny
3. Rozwiązywanie zadań i problemów
4. Rozwiązywanie zadań programistycznych
5. Konsultacje
6. Praca własna studentów

OCENA OSIĄGNIĘCIA PRZEDMIOTOWYCH EFEKTÓW UCZENIA SIĘ

Oceny (F - formatująca (w trakcie semestru), P - podsumowująca (na koniec semestru)	Numer efektu uczenia się	Sposób oceny osiągnięcia efektu uczenia się
F1	W1-W3, K1-K3	Test podsumowujący
F2	U1-U3, K1-K3	Aktywność
F3	U1-U3, K1-K3	Realizacja zadań programistycznych
P=30%*F1+30%*F2+40%*F3		

LITERATURA PODSTAWOWA I UZUPEŁNIAJĄCA

1. J. Leskovec, A. Rajaraman, J. D. Ullman, Mining of Massive Datasets, online, 2016
2. DZEJLA MEDJEDOVIC, EMIN TAHIROVIC, Algorithms and Data Structures for Massive Datasets, Manning,2022
3. Michael Mitzenmacher, Eli Upfal, Probability and Computing, Cambridge University Press, 2017

NAUCZYCIEL AKADEMICKI ODPOWIEDZIALNY ZA PRZEDMIOT

prof. Jacek Cichoń

MACIERZ POWIĄZANIA EFEKTÓW UCZENIA SIĘ DLA PRZEDMIOTU
Big Data
Z EFEKTAMI UCZENIA SIĘ NA KIERUNKU INFORMATYKA ALGORYTMICZNA

Przedmiotowy efekt uczenia się	Odniesienie przedmiotowego efektu do efektów uczenia się zdefiniowanych dla kierunku studiów	Cele przedmiotu**	Treści programowe**	Numer narzędzia dydaktycznego**
W1	K2_W01 K2_W02 K2_W03	C1	Wy1-Wy8	1 2 5 6
W2	K2_W04 K2_W05	C1	Wy1-Wy8	1 2 5 6
W3	K2_W07	C1	Wy1-Wy8	1 2 5 6
U1	K2_U01 K2_U02	C2 C3	Ćw1-Ćw6 Lab1-Lab5	3 4 5 6
U2	K2_U03 K2_U04 K2_U05	C2 C3	Ćw1-Ćw6 Lab1-Lab5	3 4 5 6
U3	K2_U06 K2_U11	C2 C3	Ćw1-Ćw6 Lab1-Lab5	3 4 5 6
K1	K2_K01 K2_K10	C1 C2 C3	Wy1-Wy8 Ćw1-Ćw6 Lab1-Lab5	1 2 3 4 5 6
K2	K2_K08	C1 C2 C3	Wy1-Wy8 Ćw1-Ćw6 Lab1-Lab5	1 2 3 4 5 6
K3	K2_K10 K2_K11	C1 C2 C3	Wy1-Wy8 Ćw1-Ćw6 Lab1-Lab5	1 2 3 4 5 6