

POLITECHNIKA KRAKOWSKA
IM. TADEUSZA KOŚCIUSZKI

KARTA PRZEDMIOTU

obowiązuje studentów rozpoczynających studia w roku akademickim 2023/2024

Wydział Informatyki i Telekomunikacji

Kierunek studiów: Matematyka Stosowana

Profil: Praktyczny

Forma sudiów: stacjonarne

Kod kierunku: MS

Stopień studiów: I

Specjalności: Analityka Danych

1 INFORMACJE O PRZEDMIOCIE

NAZWA PRZEDMIOTU	Przetwarzanie dużych zbiorów danych
NAZWA PRZEDMIOTU W JĘZYKU ANGIELSKIM	Processing Large Data Sets
KOD PRZEDMIOTU	WiT MS pIS D13 23/24
KATEGORIA PRZEDMIOTU	Przedmioty specjalnościowe
LICZBA PUNKTÓW ECTS	4.00
SEMESTRY	6

2 RODZAJ ZAJĘĆ, LICZBA GODZIN W PLANIE STUDIÓW

SEMESTR	WYKŁAD	ĆWICZENIA	LABORATORIUM	LABORATORIUM KOMPUTERO- WE	SEMINARIUM	PROJEKT
6	30	0	0	30	0	0

3 CELE PRZEDMIOTU

Cel 1 Wprowadzenie podstawowych pojęć związanych z przetwarzaniem dużych zbiorów danych.

Cel 2 Zapoznanie z ważniejszymi algorytmami i metodami stosowanymi do przechowywania, przetwarzania, analizy, modelowania i wizualizacji dużych zbiorów danych.

Cel 3 Zapoznanie z systemami zarządzania bazami danych typu NoSQL.

4 WYMAGANIA WSTĘPNE W ZAKRESIE WIEDZY, UMIEJĘTNOŚCI I INNYCH KOMPETENCJI

- 1 Znajomość języka SQL.
- 2 Znajomość języka programowania Python.

5 EFEKTY KSZTAŁCENIA

EK1 Wiedza Student rozumie i potrafi wyjaśnić pojęcia związane z pozyskiwaniem i przetwarzaniem dużych zbiorów danych

EK2 Umiejętności Student wykorzystuje funkcjonalności systemów zarządzania bazami danych do przetwarzania i analizowania zbiorów danych.

EK3 Umiejętności Student potrafi połączyć zaprojektowaną i zbudowaną bazę danych z prostym programem/skrypcem wizualizującym dane.

EK4 Kompetencje społeczne Student jest gotów do rozwiązywania problemów z zakresu analizy danych zarówno w ramach pracy indywidualnej jak i grupowej, a także poszukiwania niezbędnej w tym celu wiedzy

6 TREŚCI PROGRAMOWE

LABORATORIUM KOMPUTEROWE		
LP	TEMATYKA ZAJĘĆ OPIS SZCZEGÓŁOWY BLOKÓW TEMATYCZNYCH	LICZBA GODZIN
K1	Instalacja i konfiguracja środowiska PyCharm. Wczytywanie i formatowanie danych z plików (pliki tekstowe, pliki csv, pliki multimedialne).	2
K2	Funkcje lambda. List Comprehension.	2
K3	Instalacja biblioteki Pandas. Operacje na ramkach danych DataFrame. Wykonywanie zapytań SQL na ramkach danych DataFrame.	2
K4	Instalacja biblioteki NumPy. Operacje na danych wielowymiarowych w dedykowanych tablicach biblioteki NumPy.	2
K5	Wizualizacja danych w oparciu o bibliotekę Matplotlib.	2
K6	Wizualizacja danych w oparciu o bibliotekę Seaborn i plotnine.	2
K9	Instalacja i konfiguracja systemu zarządzania bazą danych MongoDB.	2
K10	Zapis, przechowywanie, pobieranie i manipulacja danymi przechowywanymi w ramach MongoDB.	4
K11	Instalacja i konfiguracja systemu zarządzania bazą danych Redis.	2
K12	Zapis, przechowywanie, pobieranie i manipulacja danymi przechowywanymi w ramach Redis.	4
K13	Instalacja i konfiguracja systemu zarządzania bazą danych Cassandra.	2

LABORATORIUM KOMPUTEROWE		
LP	TEMATYKA ZAJĘĆ OPIS SZCZEGÓŁOWY BLOKÓW TEMATYCZNYCH	LICZBA GODZIN
K14	Zapis, przechowywanie, pobieranie i manipulacja danymi przechowywanymi w ramach Redis.	4

WYKŁAD		
LP	TEMATYKA ZAJĘĆ OPIS SZCZEGÓŁOWY BLOKÓW TEMATYCZNYCH	LICZBA GODZIN
W1	Wprowadzenie w tematykę przetwarzania danych - czwarta rewolucja przemysłowa. Wprowadzenie podstawowych pojęć związanych z przetwarzaniem dużych zbiorów danych.	2
W2	Omówienie pojęcia ramki danych (tj. DataFrame) oraz ważniejszych operacji na dużych zbiorach danych. Omówienie bibliotek Pandas i Numpy. Przedstawienie i omówienie przykładowych skryptów w języku Python.	4
W3	Charakterystyka zbiorów danych typu Big Data - model 3V (i jego rozszerzenia). Praktyczne aspekty przetwarzania dużych zbiorów danych.	2
W4	Społeczeństwo informacyjne - model DIKW (pojęcia: dane, informacje, wiedza, mądrość).	2
W5	Metody wizualizacji danych. Omówienie bibliotek Matplotlib, Seaborn, plotnine. Przedstawienie i omówienie przykładowych skryptów w języku Python.	4
W6	Wprowadzenie do pojęć z zakresu baz danych NoSQL. Charakterystyki, wady i zalety baz NoSQL. Porównanie z relacyjnymi bazami danych.	2
W7	Modele danych w bazach NoSQL: klucz-wartość, rodziny kolumn, grafy, dokumenty. Przechowywanie baz w pamięci.	2
W8	Zapoznanie z systemami: MongoDB, Redis, Cassandra. Przedstawienie i omówienie przykładowych skryptów w języku Python.	6
W9	Techniki DataScrapping. Przedstawienie i omówienie przykładowych skryptów w języku Python.	4
W10	Odkrywanie wiedzy w bazach danych (knowledge discovery in databases KDD). Model ETL (extract, transform and load).	2

7 NARZĘDZIA DYDAKTYCZNE

- N1** Wykłady (w przypadku realizacji zajęć w trybie zdalnym z wykorzystaniem stosownych narzędzi teleinformatycznych, np. MS TEAMS)
- N2** Prezentacje multimedialne (w przypadku realizacji zajęć w trybie zdalnym z wykorzystaniem stosownych narzędzi teleinformatycznych, np. MS TEAMS)

- N3** Ćwiczenia laboratoryjne (w przypadku realizacji zajęć w trybie zdalnym z wykorzystaniem stosownych narzędzi teleinformatycznych, np. MS TEAMS)
- N4** Konsultacje (w przypadku realizacji zajęć w trybie zdalnym z wykorzystaniem stosownych narzędzi teleinformatycznych, np. MS TEAMS)
- N5** Dyskusja (w przypadku realizacji zajęć w trybie zdalnym z wykorzystaniem stosownych narzędzi teleinformatycznych, np. MS TEAMS)

8 OBCIĄŻENIE PRACĄ STUDENTA

FORMA AKTYWNOŚCI	ŚREDNIA LICZBA GODZIN NA ZREALIZOWANIE AKTYWNOŚCI
Godziny kontaktowe z nauczycielem akademickim, w tym:	
Godziny wynikające z planu studiów	60
Konsultacje przedmiotowe	10
Egzaminy i zaliczenia w sesji	2
Godziny bez udziału nauczyciela akademickiego wynikające z nakładu pracy studenta, w tym:	
Przygotowanie się do zajęć, w tym studiowanie zalecanej literatury	10
Opracowanie wyników	10
Przygotowanie raportu, projektu, prezentacji, dyskusji	28
SUMARYCZNA LICZBA GODZIN DLA PRZEDMIOTU WYNIKAJĄCA Z CAŁEGO NAKŁADU PRACY STUDENTA	120
SUMARYCZNA LICZBA PUNKTÓW ECTS DLA PRZEDMIOTU	4.00

9 SPOSOBY OCENY

OCENA FORMUJĄCA

- F1** Ćwiczenia praktyczne (waga 1/3)
- F2** Odpowiedzi ustne (waga 1/3)
- F3** Sprawozdania z umiejętności wykorzystania wybranych bibliotek (waga 1/3)

OCENA PODSUMOWUJĄCA

- P1** Średnia ważona ocen formujących (waga 1/2)
- P2** Egzamin pisemny (waga 1/2)

WARUNKI ZALICZENIA PRZEDMIOTU

- W1** Warunkiem zaliczenia przedmiotu jest uzyskanie pozytywnej oceny z każdego efektu kształcenia, pozytywnej oceny z egzaminu oraz pozytywnych wszystkich ocen cząstkowych. Ocena końcowa to średnia arytmetyczna ocen egzaminu oraz średniej ważonej ocen formujących z laboratoriów komputerowych.

W2 Spełnienie warunku obecności na obowiązkowych formach zajęć (dopuszczalna jedna nieobecność nieusprawiedliwiona na każdej z obowiązkowych form zajęć).

KRYTERIA OCENY

EFEKT KSZTAŁCENIA 1	
NA OCENĘ 2.0	Student nie spełnia warunków określonych dla oceny 3.0
NA OCENĘ 3.0	Uzyskanie co najmniej 51% punktów w ramach sposobów weryfikacji nabytych kompetencji.
NA OCENĘ 3.5	Uzyskanie co najmniej 61% punktów w ramach sposobów weryfikacji nabytych kompetencji.
NA OCENĘ 4.0	Uzyskanie co najmniej 71% punktów w ramach sposobów weryfikacji nabytych kompetencji.
NA OCENĘ 4.5	Uzyskanie co najmniej 81% punktów w ramach sposobów weryfikacji nabytych kompetencji.
NA OCENĘ 5.0	Uzyskanie co najmniej 91% punktów w ramach sposobów weryfikacji nabytych kompetencji.
EFEKT KSZTAŁCENIA 2	
NA OCENĘ 2.0	Student nie spełnia warunków określonych dla oceny 3.0
NA OCENĘ 3.0	Uzyskanie co najmniej 51% punktów w ramach sposobów weryfikacji nabytych kompetencji.
NA OCENĘ 3.5	Uzyskanie co najmniej 61% punktów w ramach sposobów weryfikacji nabytych kompetencji.
NA OCENĘ 4.0	Uzyskanie co najmniej 71% punktów w ramach sposobów weryfikacji nabytych kompetencji.
NA OCENĘ 4.5	Uzyskanie co najmniej 81% punktów w ramach sposobów weryfikacji nabytych kompetencji.
NA OCENĘ 5.0	Uzyskanie co najmniej 91% punktów w ramach sposobów weryfikacji nabytych kompetencji.
EFEKT KSZTAŁCENIA 3	
NA OCENĘ 2.0	Student nie spełnia warunków określonych dla oceny 3.0.
NA OCENĘ 3.0	Uzyskanie co najmniej 51% punktów w ramach sposobów weryfikacji nabytych kompetencji.
NA OCENĘ 3.5	Uzyskanie co najmniej 61% punktów w ramach sposobów weryfikacji nabytych kompetencji.
NA OCENĘ 4.0	Uzyskanie co najmniej 71% punktów w ramach sposobów weryfikacji nabytych kompetencji.

NA OCENĘ 4.5	Uzyskanie co najmniej 81% punktów w ramach sposobów weryfikacji nabytych kompetencji.
NA OCENĘ 5.0	Uzyskanie co najmniej 91% punktów w ramach sposobów weryfikacji nabytych kompetencji.
EFEKT KSZTAŁCENIA 4	
NA OCENĘ 2.0	Student nie spełnia warunków określonych dla oceny 3.0
NA OCENĘ 3.0	Uzyskanie co najmniej 51% punktów w ramach sposobów weryfikacji nabytych kompetencji.
NA OCENĘ 3.5	Uzyskanie co najmniej 61% punktów w ramach sposobów weryfikacji nabytych kompetencji.
NA OCENĘ 4.0	Uzyskanie co najmniej 71% punktów w ramach sposobów weryfikacji nabytych kompetencji.
NA OCENĘ 4.5	Uzyskanie co najmniej 81% punktów w ramach sposobów weryfikacji nabytych kompetencji.
NA OCENĘ 5.0	Uzyskanie co najmniej 91% punktów w ramach sposobów weryfikacji nabytych kompetencji.

10 MACIERZ REALIZACJI PRZEDMIOTU

EFEKT KSZTAŁCENIA	ODNIESIENIE DANEGO EFEKTU DO SZCZEGÓŁOWYCH EFEKTÓW ZDEFINIOWANYCH DLA PROGRAMU	CELE PRZEDMIOTU	TREŚCI PROGRAMOWE	NARZĘDZIA DYDAKTYCZNE	SPOSOBY OCENY
EK1	K_W06 K_W13 K_W17 K_W20 K_W22 K_W23 K_W24	Cel 1 Cel 3	W1 W2 W3 W4 W5 W6 W7 W8 W9 W10	N1 N2 N4 N5	F2 F3 P1 P2
EK2	K_U07 K_U10 K_U13 K_U17 K_U19 K_U20 K_U22 K_U25	Cel 2 Cel 3	K1 K2 K3 K4 K5 K6 K9 K10 K11 K12 K13 K14 W2 W5 W8 W9	N1 N2 N3 N4 N5	F1 F2 F3 P1 P2
EK3	K_U17 K_U18 K_U19 K_U20 K_U21 K_U22 K_U23 K_U24 K_U25	Cel 2	K9 K10 K11 K12 K13 K14 W5 W8 W9	N1 N2 N3 N4 N5	F1 F2 F3 P1 P2

EFEKT KSZTAŁCENIA	ODNIESIENIE DANEGO EFEKTU DO SZCZEGÓŁOWYCH EFEKTÓW ZDEFINIOWANYCH DLA PROGRAMU	CELE PRZEDMIOTU	TREŚCI PROGRAMOWE	NARZĘDZIA DYDAKTYCZNE	SPOSOBY OCENY
EK4	K_K01 K_K02 K_K06 K_K08	Cel 1 Cel 2 Cel 3	K1 K2 K3 K4 K5 K6 K9 K10 K11 K12 K13 K14 W1 W2 W3 W4 W5 W6 W7 W8 W9 W10	N3 N4 N5	F2 F3 P1

11 WYKAZ LITERATURY

LITERATURA PODSTAWOWA

- [1] | **Robert Johansson** — *Matematyczny Python : obliczenia naukowe i analiza danych z użyciem NumPy, SciPy i Matplotlib*, Gliwice, 2021, Helion
- [2] | **Pramod J. Sadalage, Martin Fowler** — *NoSQL : kompendium wiedzy*, Gliwice, 2015, Helion
- [3] | **Shannon Bradshaw, Eoin Brazil, Kristina Chodorow** — *Przewodnik po MongoDB : wydajna i skalowalna baza danych*, Gliwice, 2021, Helion

12 INFORMACJE O NAUCZYCIELACH AKADEMICKICH

OSOBA ODPOWIEDZIALNA ZA KARTĘ

dr inż. Daniel Grzonka (kontakt: daniel.grzonka@pk.edu.pl)

OSOBY PROWADZĄCE PRZEDMIOT

- 1 dr Barbara Borowik (kontakt: barbara.borowik@pk.edu.pl)
- 2 dr inż. Daniel Grzonka (kontakt: daniel.grzonka@pk.edu.pl)
- 3 dr inż. Anna Plichta (kontakt: anna.plichta@pk.edu.pl)
- 4 mgr inż. Piotr Biskup (kontakt: piotr.biskup@pk.edu.pl)

13 ZATWIERDZENIE KARTY PRZEDMIOTU DO REALIZACJI

(miejscowość, data)

(odpowiedzialny za przedmiot)

(dziekan)



PRZYJMUJĘ DO REALIZACJI (data i podpisy osób prowadzących przedmiot)

.....
.....
.....
.....