SYLABUS

DOTYCZY CYKLU KSZTAŁCENIA 2026/2027 - 2029/2030

Rok akademicki 2028/2029

1. PODSTAWOWE INFORMACJE O PRZEDMIOCIE

Nazwa przedmiotu	*eksploracja danych*
Kod przedmiotu*	
Nazwa jednostki prowadzącej kierunek	*Instytut Informatyki, Wydział Nauk Ścisłych i Technicznych*
Nazwa jednostki realizującej przedmiot	*Instytut Informatyki, Wydział Nauk Ścisłych i Technicznych*
Kierunek studiów	*sztuczna inteligencja*
Poziom studiów	*studia I stopnia*
Profil	*ogólnoakademicki*
Forma studiów	*stacjonarne*
Rok i semestr/y studiów	*rok III, semestr 5,6*
Rodzaj przedmiotu	*przedmiot kierunkowy*
Język wykładowy	*polski*
Koordynator	*dr hab. Jan G. Bazan, prof. UR*
Imię i nazwisko osoby prowadzącej / osób prowadzących	*dr hab. Jan G. Bazan, prof. UR*

***** - *opcjonalnie, zgodnie z ustaleniami w Jednostce*

1.1.Formy zajęć dydaktycznych, wymiar godzin i punktów ECTS

Semestr (nr)	Wykł.	Ćw.	Konw.	Lab.	Sem.	ZP	Prakt.	Inne (jakie?)	**Liczba pkt. ECTS**
5	15			15					2
6	15			15					3

1.2. Sposób realizacji zajęć

zajęcia w formie tradycyjnej

1.3 Forma zaliczenia przedmiotu (z toku)

Semestr 5: wykład – zaliczenie bez oceny, laboratoria – zaliczenie z oceną
Semestr 6: wykład – egzamin, laboratoria – zaliczenie z oceną

2.WYMAGANIA WSTĘPNE

Znajomość zagadnień realizowanych na przedmiotach: algorytmy i struktury danych, bazy danych, programowanie obiektowe, programowanie w języku Python, Podstawy sztucznej inteligencji.

3. CELE, EFEKTY UCZENIA SIĘ, TREŚCI PROGRAMOWE I STOSOWANE METODY DYDAKTYCZNE

3.1 Cele przedmiotu

C1	Celem przedmiotu jest zaznajomienie studentów z podstawowymi metodami eksploracji danych zarówno w stosunku do małych danych, jak i dużych danych.

3.2 Efekty uczenia się dla przedmiotu

EK (efekt uczenia się)	Treść efektu uczenia się zdefiniowanego dla przedmiotu	Odniesienie do efektów kierunkowych
EK_01	Ma podstawową wiedzę dotyczącą klasycznych metod eksploracji danych.	K_W04
EK_02	Ma podstawową wiedzę dotyczącą metod eksploracji danych przeznaczonych do eksploracji dużych danych.	K_W04
EK_03	Potrafi stosować wybrane narzędzia i biblioteki informatyczne do eksploracji małych danych, w tym do realizacji w nich podstawowych zadań eksploracji	K_U04, K_U05
EK_04	Potrafi stosować wybrane narzędzia i biblioteki informatyczne do eksploracji dużych danych, w tym do realizacji w nich podstawowych zadań eksploracji.	K_U04, K_U05
EK_05	Potrafi realizować projekty informatyczne związane z budową systemów eksploracji dużych danych.	K_U07, K_U05

3.3 Treści programowe

A. Problematyka wykładu

Część 1 (semestr 5)
1. Wprowadzenie do komputerowych metod eksploracji danych.
2. Przegląd najbardziej znanych systemów i bibliotek programistycznych do eksploracji danych.
3. Środowisko programowania w języku Python jako nowoczesne środowisko do eksploracji danych.
4. Przygotowanie zbiorów danych do eksploracji (obliczanie statystyk dotyczących zbiorów danych, radzenie sobie z pustymi miejscami w danych, skalowanie i standaryzacja cech, kodowanie cech symbolicznych).
5. Klasyfikatory.
6. Grupowanie danych.
7. Reguły asocjacyjne i sekwencyjne.
8. Filtracja cech.
9. Odkrywanie nietypowych przypadków.
10. Eksploracja danych z wykorzystaniem języka SQL.

11. Kolokwium z teorii dotyczącej eksploracji danych.
Część 2 (semestr 6)
1. Duże dane i problemy z ich eksploracją.
2. Zestawienie współczesnych podejść do eksploracji dużych zbiorów danych.
3. Strumieniowa eksploracja danych.
4. Eksploracja danych z wykorzystaniem równoległości obliczeń.
5. Podejście do obliczeń oparte na paradygmacie Map-Reduce z wykorzystaniem Apache Spark.
6. Biblioteka ML- lib dla Apache Spark
7. Koncepcja jeziora danych i jego realizacja za pomocą Apache Spark.
8. Kolokwium z teorii dotyczącej eksploracji dużych danych.

B. Problematyka laboratoriów

Część 1 (semestr 5)
1. Zadania na przygotowanie danych do eksploracji.
2. Zadania na konstruowanie klasyfikatorów.
3. Zadania na grupowanie danych.
4. Zadania na wyliczanie reguł asocjacyjnych i sekwencyjnych.
5. Zadania na filtrację cech i wykrywanie nietypowych przypadków.
6. Zadania na eksploracja danych z wykorzystaniem języka SQL.
7. Kolokwium przy komputerze z eksploracji danych.
Część 2 (semestr 6)
1. Zadania na strumieniową eksplorację danych.
2. Zadania na realizację eksploracji danych z wykorzystaniem równoległości obliczeń.
3. Zadania na zastosowanie paradygmatu Map-Reduce do obliczeń.
4. Zadania na wykorzystanie biblioteki ML-lib dla Apache Spark.
5. Zadania związane z konstrukcją jeziora danych.
6. Zaliczenie projektu z eksploracji dużych danych.

3.4 Metody dydaktyczne

Wykład: wykład z prezentacją multimedialną
Laboratorium: rozwiązywanie zadań z użyciem komputera.

4. METODY I KRYTERIA OCENY

4.1 Sposoby weryfikacji efektów uczenia się

Symbol efektu	Metody oceny efektów uczenia się	Forma zajęć dydaktycznych

	(np.: kolokwium, egzamin ustny, egzamin pisemny, projekt, sprawozdanie, obserwacja w trakcie zajęć)	(w, ćw, ...)
CZĘŚĆ 1 (SEMESTR 5)		
EK_ 01	kolokwium	wykład
EK_03	kolokwium	lab
CZĘŚĆ 2 (SEMESTR 6)		
EK_ 02	egzamin	wykład
EK_04, EK_05	projekt	lab

4.2 Warunki zaliczenia przedmiotu (kryteria oceniania)

Zadania i pytania weryfikujące poszczególne efekty kształcenia formułowane są w taki sposób, że:

- na ocenę dst dotyczą znajomości lub umiejętności zastosowania metod i technologii,
- na ocenę db dotyczą porównywania pojawiających się metod i technologii,
- na ocenę bdb dotyczą optymalizowania jakości i efektywności metod i technologii dla konkretnych zagadnień.

Zasady uzyskania oceny końcowej:

Zaliczenie laboratorium następuje na podstawie zaliczenia wszystkich efektów weryfikowanych przez planowane w danym okresie metody weryfikacji. Przy tym zakłada się, że każda metoda weryfikacji dostarcza osobne oceny dla każdego z weryfikowanych przez nią efektów uczenia się. Jeśli dany efekt jest weryfikowany przez więcej niż jedną metodę, to ocena weryfikująca osiągnięcie tego efektu jest obliczana jako średnia arytmetyczna ocen uzyskanych w poszczególnych metodach weryfikowania tego efektu.

Student otrzymuje z zaliczenia ocenę **niedostateczny,** gdy metody weryfikacji wykażą, iż co najmniej jeden z efektów nie został osiągnięty (średnia ocena dla tego efektu jest niższa niż 3.0);

Student otrzymuje ocenę **dostateczny,** gdy przeciętnie każdy z efektów zostanie osiągnięty na poziomie co najmniej 3.0, ale chociaż jeden z efektów został osiągnięty na poziomie mniejszym od 3.75;

Student otrzymuje ocenę **dobry,** gdy przeciętnie każdy z efektów zostanie osiągnięty na poziomie co najmniej 3.75, ale chociaż jeden z efektów został osiągnięty na poziomie mniejszym od 4.75;

Student otrzymuje ocenę **bardzo dobry,** gdy przeciętnie każdy z efektów zostanie osiągnięty na poziomie co najmniej 4.75

5. CAŁKOWITY NAKŁAD PRACY STUDENTA POTRZEBNY DO OSIĄGNIĘCIA ZAŁOŻONYCH EFEKTÓW W GODZINACH ORAZ PUNKTACH ECTS

Forma aktywności	Średnia liczba godzin na zrealizowanie aktywności
Godziny z harmonogramu studiów	60
Inne z udziałem nauczyciela akademickiego (udział w konsultacjach, egzaminie)	5
Godziny niekontaktowe – praca własna studenta (przygotowanie do zajęć, egzaminu, napisanie referatu itp.)	60
SUMA GODZIN	125
SUMARYCZNA LICZBA PUNKTÓW ECTS	**5**

** Należy uwzględnić, że 1 pkt ECTS odpowiada 25-30 godzin całkowitego nakładu pracy studenta.*

6. PRAKTYKI ZAWODOWE W RAMACH PRZEDMIOTU

wymiar godzinowy	-
zasady i formy odbywania praktyk	-

7. LITERATURA

Literatura podstawowa:

1. Prezentacje wykładowe autorstwa Jana Bazana dostępne na MS Teams.
2. Morzy T., Eksploracja danych; Metody i algorytmy, PWN, 2013.
3. Albon, C: Uczenie maszynowe w Pythonie. Receptury, Helion, 2019.
4. Akash Tandom, Sandy Ryza, Uri Laserson, Sean Owen, Josh Wills: Zaawansowana analiza danych w PySpark : metody przetwarzania informacji na szeroką skalę z wykorzystaniem Pythona i systemu Spark, Helion (2023).
5. Zestaw notebooków dotyczących eksploracji danych w języku Python w środowisku Jupyter autorstwa Jana Bazana - dostępne na Teams.
6. Dokumentacja pakietu Scikit-learn: https://scikit-learn.org

Literatura uzupełniająca:

1. Larose D.T., Odkrywanie wiedzy z danych; wprowadzenie do eksploracji danych, PWN, 2006.
2. Beazley, D.: Python, zwięzłe kompendium dla programisty, Helion, 2022.
3. Moroney, L: Sztuczna inteligencja i uczenie maszynowe dla programistów, Helion, 2021.
4. Raschka, S: Python, Uczenie maszynowe, Helion, 2019.