

SYLABUS

DOTYCZY CYKLU KSZTAŁCENIA 2022 – 2026

Rok akademicki 2024/2025

1. PODSTAWOWE INFORMACJE O PRZEDMIOCIE

Nazwa przedmiotu	<i>metody eksploracji danych</i>
Kod przedmiotu*	
Nazwa jednostki prowadzącej kierunek	<i>Kolegium Nauk Przyrodniczych</i>
Nazwa jednostki realizującej przedmiot	<i>Kolegium Nauk Przyrodniczych</i>
Kierunek studiów	<i>informatyka</i>
Poziom studiów	<i>studia inżynierskie I-go stopnia</i>
Profil	<i>ogólnoakademicki</i>
Forma studiów	<i>stacjonarne</i>
Rok i semestr/y studiów	<i>rok III, semestr 5</i>
Rodzaj przedmiotu	<i>przedmiot kształcenia specjalnościowego</i>
Język wykładowy	<i>polski</i>
Koordynator	<i>dr Wojciech Rząsa</i>
Imię i nazwisko osoby prowadzącej / osób prowadzących	<i>dr Wojciech Rząsa</i>

* -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	30			30					5

1.2. Sposób realizacji zajęć

zajęcia w formie tradycyjnej

1.3 Forma zaliczenia przedmiotu (z toku)

egzamin

2. WYMAGANIA WSTĘPNE

Znajomość zagadnień realizowanych na przedmiotach: algorytmy i struktury danych oraz sztuczna inteligencja, a także umiejętność programowania w języku Python.

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

3.1 Cele przedmiotu

C1	Poszerzenie wiedzy studentów z zakresu metod uczenia maszynowego.
C2	Zaznajomienie studentów z usystematyzowany sposób z metodami eksploracji danych w stosunku do niewielkich danych oraz z wybranymi metodami dostosowanymi do dużych danych.
C3	Wyćwiczenie umiejętności stosowania narzędzi programistycznych w przykładowych zagadnieniach eksploracji danych.

3.2 Efekty uczenia się dla przedmiotu

EK (efekt uczenia się)	Treść efektu uczenia się zdefiniowanego dla przedmiotu	Odniesienie do efektów kierunkowych ¹
EK_01	Ma usystematyzowaną wiedzę dotyczącą modeli odkrywania wiedzy z danych (ang. KDD).	K_Wo4, K_Wo7
EK_02	Ma wiedzę, podbudowaną teoretycznie, dotyczącą technik preprocesingowych, klasycznych metod eksploracji danych oraz technik postprocesingowych. Ma ogólną wiedzę o metodach eksploracji dużych danych	K_Wo4, K_Wo7
EK_03	Potrafi przeprowadzić „na kartce papieru” symulacje algorytmów eksploracji danych oraz stosować wybrane biblioteki do różnego rodzaju eksploracji „łatwych” zbiorów danych.	K_U11, K_U12, K_U20,
EK_04	Zna historię dziedziny eksploracja danych i rozumie proces dalszego rozwoju technik z tego zakresu dostosowanych do współczesnych problemów i narzędzi	K_Ko1

3.3 Treści programowe

A. Problematyka wykładu

1. Wprowadzenie do zagadnienia odkrywania wiedzy z danych – różne modele KDD.
2. Etap preprocesingu – wybrane rodzaje niedoskonałości danych i techniki radzenia sobie z nimi (określanie skal pomiarowych atrybutów występujących w danych, obliczanie statystyk eksplorowanych zbiorów danych, skalowanie i standaryzacja cech, kodowanie cech symbolicznych, redukcja rozmiaru danych, przetwarzanie danych zawierających wartości brakujące, dyskretyzacja danych).
3. Etap eksploracji danych - różne metody i algorytmy eksploracyjnej analizy danych, klasyfikacji, regresji, odkrywania asocjacji, skupień, elementów nietypowych (odstających), wzorców w danych temporalnych

¹ W przypadku ścieżki kształcenia prowadzącej do uzyskania kwalifikacji nauczycielskich uwzględnić również efekty uczenia się ze standardów kształcenia przygotowującego do wykonywania zawodu nauczyciela.

- | |
|---|
| 4. Etap postprocesingu – ewaluacja, filtracja, generalizacja i interpretacja wygenerowanej informacji (wiedzy), |
| 5. Przegląd wybranych funkcji biblioteki języka Python przeznaczonych do eksploracji danych. |

B. Problematyka laboratoriów

- | |
|--|
| 1. Biblioteki matplotlib, plotnine i scikit-learn |
| 2. Wstępna analiza danych – zrozumienie danych lub dziedziny, z której dane pochodzą |
| 3. Wizualizacja danych, odkrywanie podstawowych informacji zawartych w danych (głównie statystycznych) oraz ich niedoskonałości. |
| 4. Stosowanie wybranych technik preprocesingowych. |
| 5. Stosowanie wybranych technik eksploracji danych. |
| 6. Stosowanie wybranych technik postprocesingowych. |

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 (np.: kolokwium, egzamin ustny, egzamin pisemny, projekt, sprawozdanie, obserwacja w trakcie zajęć)	Forma zajęć dydaktycznych (w, ćw, ...)
EK_01	egzamin	wykład
EK_02	egzamin	wykład
EK_03	dwa kolokwia	lab
EK_04	obserwacja w trakcie zajęć	lab

4.2 Warunki zaliczenia przedmiotu (kryteria oceniania)

Na **egzaminie** weryfikowana jest znajomość wszystkich zagadnień prezentowanych na wykładzie. Warunkiem zaliczenia egzaminu jest uzyskanie co najmniej 50% punktów z każdego modułu spośród: modele KDD, techniki preprocesingowe, eksploracja danych, techniki postprocesingowe.

Zaliczenie laboratoriów wystawiane jest na podstawie wyników z dwóch kolokwiów. Każde z dwóch kolokwiów jest dwuczęściowe. W pierwszej części należy rozwiązać na kartce zadania związane z symulacją algorytmów. W drugiej należy wykazać się umiejętnością stosowania poznanych narzędzi informatycznych do eksploracji danych.

Aby zaliczyć laboratoria niezbędne jest zaliczenie na co najmniej 50% każdego z czterech modułów: modułu „teoretycznego” na kolokwium 1 i 2 oraz modułu „praktycznego” na kolokwium 1 i 2.

Ocena końcowa ustalana jest na podstawie średniej liczby zdobytych punktów z każdego z czterech modułów w następujący sposób:

[0; 50%) – 2.0

[50%; 60%) – 3.0

[60%; 70%) – 3.5

[70%; 80%) – 4.0

[80%; 90%) – 4.5
[90%; 100%] – 5.0.

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

Forma aktywności	Średnia liczba godzin na zrealizowanie aktywności
Godziny kontaktowe wynikające 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.)	65
SUMA GODZIN	130
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. Morzy T.: Eksploracja danych; Metody i algorytmy, PWN, 2013.
2. Hand D., Mannila H., Smyth P.: Eksploracja danych, Warszawa, Wydawnictwa Naukowo-Techniczne 2005
3. Harrison, M.: Uczenie maszynowe w Pythonie, Helion, 2020.
4. Albon, C: Uczenie maszynowe w Pythonie. Receptury, Helion, 2019.

Literatura uzupełniająca:

1. Larose D.T., Odkrywanie wiedzy z danych; wprowadzenie do eksploracji danych, PWN, 2006.
2. Cios K., Pedrycz W., Swiniarski R. Kurgan L.: Data mining : a knowledge discovery approach, Springer 2007