

Analiza danych

Sylabus zajęć

Informacje podstawowe

Kierunek studiów Analiza i przetwarzanie danych Specjalność - Jednostka organizacyjna Wydział Matematyki i Informatyki Poziom studiów studia drugiego stopnia Forma studiów studia stacjonarne Profil studiów profil ogólnoakademicki		Cykl dydaktyczny 2023/24 Kod zajęć 06APDS.24K.00079.23 Języki wykładowe polski Obligatoryjność Obowiązkowy Blok zajęciowy Przedmioty kierunkowe
Koordynator zajęć	Tomasz Górecki	
Prowadzący zajęcia	Tomasz Górecki	
Okres Semestr 3	Forma zajęć / liczba godzin / forma zaliczenia - Wykład: 30, Egzamin; w tym zajęcia zdalne: - Wykład synchroniczny: 30 - Laboratorium: 30, Zaliczenie z oceną	Liczba punktów ECTS 6

Cele kształcenia dla zajęć

Kod	Cel
C1	Prezentowanie problematyki analizy danych.
C2	Prezentowanie ogólnej wiedzy z metod analizy danych.
C3	Omówienie podstawowych bibliotek używanych w języku R do analizy danych.
C4	Omówienie metod uzupełniania braków w danych.
C5	Omówienie metod kodowania danych jakościowych.
C6	Omówienie metod wykrywania obserwacji odstających.
C7	Prezentowanie metod analizy regresji.
C8	Prezentowanie metod analizy przeżycia.
C9	Prezentowanie metodami analizy szeregów czasowych.
C10	Prezentowanie metod redukcji wymiaru.
C11	Prezentowanie metod klasyfikacji.
C12	Prezentowanie metod analizy skupień.

Wymagania wstępne

- Podstawowa znajomość metod statystyki i rachunku prawdopodobieństwa
- Podstawowa znajomość języka programowania R

Efekty uczenia się dla zajęć

Kod	Efekty uczenia się dla zajęć w zakresie	Efekty uczenia się dla kierunku	Metody weryfikacji osiągnięcia efektów uczenia się dla zajęć
Wiedzy - Student/ka:			
W1	zna metody uzupełniania braków w danych.	APD_K2_W01, APD_K2_W03, APD_K2_W04	Egzamin pisemny, Kolokwium pisemne
W2	zna metody wykrywania obserwacji odstających w danych.	APD_K2_W01, APD_K2_W03, APD_K2_W04	Egzamin pisemny, Kolokwium pisemne
W3	zna metody kodowania danych jakościowych.	APD_K2_W01, APD_K2_W03, APD_K2_W04	Egzamin pisemny, Kolokwium pisemne
W4	zna metody regresji oraz analizy przeżycia.	APD_K2_W01, APD_K2_W03, APD_K2_W04	Egzamin pisemny, Kolokwium pisemne
W5	zna metody analizy szeregów czasowych.	APD_K2_W01, APD_K2_W03, APD_K2_W04	Egzamin pisemny, Kolokwium pisemne
W6	zna metody klasyfikacji.	APD_K2_W01, APD_K2_W03, APD_K2_W04	Egzamin pisemny, Kolokwium pisemne

Kod	Efekty uczenia się dla zajęć w zakresie	Efekty uczenia się dla kierunku	Metody weryfikacji osiągnięcia efektów uczenia się dla zajęć
W7	zna metody redukcji wymiaru.	APD_K2_W01, APD_K2_W03, APD_K2_W04	Egzamin pisemny, Kolokwium pisemne
W8	zna metody analizy skupień.	APD_K2_W01, APD_K2_W03, APD_K2_W04	Egzamin pisemny, Kolokwium pisemne
Umiejętności - Student/ka:			
U1	potrafi zastosować zdobyte umiejętności analizy danych do rozwiązywania praktycznych problemów.	APD_K2_U01, APD_K2_U04, APD_K2_U05, APD_K2_U11	Kolokwium pisemne
U2	potrafi zastosować poznane algorytmy w języku programowania R.	APD_K2_U01, APD_K2_U03, APD_K2_U04, APD_K2_U07	Kolokwium pisemne
Kompetencji społecznych - Student/ka:			
K1	jest gotowy/gotowa do zrozumienia potrzeby analizy danych i interpretacji uzyskanych wyników.	APD_K2_K01, APD_K2_K02, APD_K2_K03, APD_K2_K04, APD_K2_K05, APD_K2_K06	Egzamin pisemny, Kolokwium pisemne
K2	jest gotowy/gotowa do pracy indywidualnej i grupowej w zakresie realizacji projektów badawczych i społecznych opartych na analizie danych.	APD_K2_K01, APD_K2_K02, APD_K2_K03, APD_K2_K04, APD_K2_K05, APD_K2_K06	Egzamin pisemny, Kolokwium pisemne

Treści programowe dla zajęć

Lp.	Treści programowe dla zajęć	Efekty uczenia się dla zajęć	Formy zajęć
1.	• Uzupełnianie braków w danych • Kodowanie danych jakościowych • Wykrywanie obserwacji odstających	W1, W2, W3, U1, U2, K1, K2	Wykład, Laboratorium, Wykład synchroniczny
2.	• Analiza regresji • Analiza przeżycia	W4, U1, U2, K1, K2	Wykład, Laboratorium, Wykład synchroniczny
3.	Analiza szeregów czasowych	W5, U1, U2, K1, K2	Wykład, Laboratorium, Wykład synchroniczny
4.	Klasyfikacja (analiza dyskryminacyjna)	W6, U1, U2, K1, K2	Wykład, Laboratorium, Wykład synchroniczny
5.	Redukcja wymiaru	W7, U1, U2, K1, K2	Wykład, Laboratorium, Wykład synchroniczny
6.	Analiza skupień	W8, U1, U2, K1, K2	Wykład, Laboratorium, Wykład synchroniczny

Informacje dodatkowe

Forma zajęć	Metody i formy prowadzenia zajęć
Wykład	Wykład z prezentacją multimedialną wybranych zagadnień, Wykład konwersatoryjny
Laboratorium	Dyskusja, Uczenie problemowe (Problem-based learning), Rozwiązywanie zadań (np.: obliczeniowych, artystycznych, praktycznych), Metoda ćwiczeniowa, Metoda laboratoryjna

Forma zajęć	Warunki zaliczenia zajęć
Wykład	Warunkiem przystąpienia do egzaminu jest wcześniejsze zaliczenie laboratoriów. Egzamin pisemny w postaci testu jednokrotnego wyboru (około 50 pytań), 4 odpowiedzi do wyboru. Skala ocen: • 0%- 40%: 2,0 • 50%-59%: 3,0 • 60%-69%: 3,5 • 70%-79%: 4,0 • 80%-89%:4,5 • 90%-100%: 5,0
Laboratorium	Dwa kolokwia zaliczeniowe. O ocenie decyduje suma punktów uzyskanych z dwóch kolokwiów. Skala ocen: • 0%- 40%: 2,0 • 50%-59%: 3,0 • 60%-69%: 3,5 • 70%-79%: 4,0 • 80%-89%:4,5 • 90%-100%: 5,0 Będzie jedno kolokwium poprawkowe z całości materiału.

Literatura

Obowiązkowa

1. Górecki, T. (2011). Podstawy statystyki z przykładami w R. BTC.
2. Hastie, T., Tibshirani, R., Friedman, J. (2009). The Elements of Statistical Learning. Springer.
3. James, G. Witten, D., Hastie, T., Tibshirani, R. (2017). An Introduction to Statistical Learning with Applications in R. Springer.
4. Krzyśko, M., Wołyński, W., Górecki, T., Skorzybut, M. (2008). Systemy uczące się. WNT.
5. Nwanganga, F., Chapple, M. (2022). Praktyczne uczenie maszynowe w języku R. APN Promise.
6. Szeliga, M. (2017). Data science i uczenie maszynowe. PWN.
7. Wickham, H., Golemund, G. (2018). Język R. Helion.

Dodatkowa

1. Biecek, P. (2016). Odkrywać! Ujawniać! Objaśniać! Wydawnictwo UW.
2. Biecek, P. (2017). Przewodnik po pakiecie R. GiS.
3. Gągolewski, M. (2016). Programowanie w języku R. PWN.
4. Ghatak, A. (2019). Deep Learning with R. Springer.
5. Gillespie, C., Lovelace, R. (2018). Wydajne programowanie w R. O'Reilly Media.
6. Lander, J.P. (2018). Język R dla każdego. APN Promise.
7. Winke, C.O. (2020). Podstawy wizualizacji danych. Helion.

Nakład pracy studenta i punkty ECTS

Rodzaje zajęć studenta	Średnia liczba godzin* przeznaczonych na zrealizowane rodzaje zajęć
------------------------	---

Wykład	30
Laboratorium	30
Przygotowanie do zajęć	10
Czytanie wskazanej literatury	20
Przygotowanie do zaliczenia	30
Przygotowanie do egzaminu	30
Łączny nakład pracy studenta	Liczba godzin 150
Liczba punktów ECTS	ECTS 6

* godzina (lekcyjna) oznacza 45 minut

Efekty uczenia się dla kierunku

Kod	Treść
APD_K2_K01	Absolwent/ka jest gotów/gotowa do zrozumienia wagi i znaczenia matematyki i informatyki w rozwiązywaniu problemów poznawczych
APD_K2_K02	Absolwent/ka jest gotów/gotowa do zrozumienia i pogłębienia świadomości roli informatyki w kształtowaniu życia społecznego
APD_K2_K03	Absolwent/ka jest gotów/gotowa do zrozumienia ograniczenia własnej wiedzy i rozumie potrzebę dalszego kształcenia, w tym zdobywania wiedzy pozadziśnawej
APD_K2_K04	Absolwent/ka jest gotów/gotowa do precyzyjnego formułowania pytań służących pogłębieniu własnego zrozumienia danego tematu lub odnalezieniu brakujących elementów rozumowania
APD_K2_K05	Absolwent/ka jest gotów/gotowa do zrozumienia i docenienia znaczenia uczciwości intelektualnej w działaniach własnych i innych osób; postępuje etycznie
APD_K2_K06	Absolwent/ka jest gotów/gotowa do zrozumienia najważniejszych osiągnięć w swojej dziedzinie i stojących przed nią wyzwań; potrafi je przedstawić laikom w sposób popularny
APD_K2_U01	Absolwent/ka potrafi wykorzystywać zaawansowane metody, narzędzia i technologie informatyczne do rozwiązywania problemów teoretycznych i problemów praktycznych w wybranych obszarach zastosowań
APD_K2_U03	Absolwent/ka potrafi pisać, uruchamiać i testować złożone programy w wybranych środowiskach programistycznym
APD_K2_U04	Absolwent/ka potrafi przeprowadzić zaawansowaną statystyczną analizę danych i wnioskowanie statystyczne z wykorzystaniem rachunku prawdopodobieństwa i pakietów statystycznych
APD_K2_U05	Absolwent/ka potrafi projektować, analizować i stosować zaawansowane algorytmy
APD_K2_U07	Absolwent/ka potrafi posługiwać się różnymi formatami reprezentacji rozmaitych rodzajów danych, w szczególności integrować dane pochodzące z różnych źródeł, transformować je do wspólnej spójnej postaci
APD_K2_U11	Absolwent/ka potrafi w sposób przystępny przedstawić fakty, problemy i ich rozwiązania z zakresu matematyki i informatyki, porozumiewać się przy użyciu różnych technik w środowisku zawodowym oraz w innych środowiskach, w tym w języku angielskim oraz z wykorzystaniem narzędzi informatycznych współczesnych metod prezentacyjnych
APD_K2_W01	Absolwent/ka zna i rozumie w pogłębionym stopniu współczesne metody, narzędzia i technologie informatyczne właściwe dla wybranych obszarów zastosowań niezbędne przy analizie i przetwarzaniu danych pochodzących z różnych źródeł
APD_K2_W03	Absolwent/ka zna i rozumie zasady rozwiązywania problemów z wykorzystaniem zaawansowanych algorytmów i struktur danych
APD_K2_W04	Absolwent/ka zna i rozumie zagadnienia składowania, eksploracji i modelowania danych