

-SYLABUS

DOTYCZY CYKLU KSZTAŁCENIA 2023-2025

Rok akademicki 2024/2025

1. PODSTAWOWE INFORMACJE O PRZEDMIOCIE

Nazwa przedmiotu	<i>przedmiot obieralny 1: statystyczne systemy uczące</i>
Kod przedmiotu	
Nazwa jednostki prowadzącej kierunek	<i>Instytut Informatyki, Kolegium Nauk Przyrodniczych</i>
Nazwa jednostki realizującej przedmiot	<i>Instytut Informatyki, Kolegium Nauk Przyrodniczych</i>
Kierunek studiów	<i>informatyka</i>
Poziom studiów	<i>studia II stopnia</i>
Profil	<i>ogólnoakademicki</i>
Forma studiów	<i>stacjonarne</i>
Rok i semestr/y studiów	<i>rok II, semestr 2</i>
Rodzaj przedmiotu	<i>przedmiot kierunkowy do wyboru</i>
Język wykładowy	<i>język polski</i>
Koordynator	<i>dr Lech Zaręba</i>
Imię i nazwisko osoby prowadzącej / osób prowadzących	<i>dr Lech Zaręba</i>

1.1. Formy zajęć dydaktycznych, wymiar godzin i punktów ECTS

Semestr (nr)	Wykł.	Ćw.	Konw.	Lab.	Sem.	ZP	Prakt.	Projekt	Liczba pkt. ECTS
2	15			30					4

1.2. Sposób realizacji zajęć

zajęcia w formie tradycyjnej

1.3 Forma zaliczenia przedmiotu (z toku)

zaliczenie z oceną

2. WYMAGANIA WSTĘPNE

Znajomość podstawowych definicji i twierdzeń z zakresu rachunku prawdopodobieństwa i statystyki na poziomie studiów I-go stopnia oraz metod statystyki matematycznej z I semestru. Podstawy programu R.

3. CELE, EFEKTY UCZENIA SIĘ, TREŚCI PROGRAMOWE I STOSOWANE METODY DYDAKTYCZNE**3.1 Cele przedmiotu**

C ₁	Zapoznanie studentów z podstawowymi zagadnieniami i pojęciami wykorzystywanymi do statystycznych systemów uczących
C ₂	Wykształcenie u studentów umiejętności posługiwania się programami służącymi analizie, modelowaniu i weryfikacji modeli tworzonych z wykorzystaniem statystycznych systemów uczących
C ₃	Wykształcenie u studentów umiejętności wyboru odpowiedniej metody analizy i modelowania opartej na statystycznych systemach uczących do analizy danego rodzaju danych

3.2 Efekty uczenia się dla przedmiotu

EK (efekt uczenia się)	Treść efektu uczenia się zdefiniowanego dla przedmiotu Student	Odniesienie do efektów kierunkowych
EK_01	Zna w pogłębionym stopniu zagadnienia związane z wykorzystaniem statystycznych systemów uczących używanych w uczeniu maszynowym i eksploracji danych oraz zna podstawowe programy informatyczne do modelowania zjawisk opartych na modelowaniu matematycznym i statystycznym.	K_Wo3
EK_02	Potrafi właściwie zaplanować oraz wykonać pomiary i symulacje komputerowe z wykorzystaniem statystycznych systemów uczących. Dobrać właściwe systemy do rodzaju analizowanych danych i modelowanych zjawisk. Potrafi dokonać interpretacji uzyskanych rezultatów, poprawnie wyciągnąć wnioski, formułować i weryfikować hipotezy. Potrafi również modyfikować i sterować algorytmami wykorzystywanymi w statystycznych systemach uczących.	K_Uo2
EK_03	Uznaje znaczenie matematyki i statystyki w rozwiązywaniu problemów związanych z modelowaniem różnych zjawisk i procesów rzeczywistych, umie korzystać z opinii ekspertów w przypadku trudności z samodzielnym rozwiązaniem problemu.	K_Ko1

3.3 Treści programowe**A. Problematyka wykładu**

Liniowe metody klasyfikacji (Klasyfikacja pod nadzorem, Fisherowska dyskryminacja liniowa, dyskryminacja kwadratowa, metoda FDA, MDA, RDA)
Metody klasyfikacji oparte na rozkładach prawdopodobieństwa (klasyfikator bayesowski i metoda największej wiarygodności, optymalność reguły bayesowskiej, praktyczna konstrukcja klasyfikatorów)
Metody klasyfikacji oparte na nieparametrycznej estymacji (metoda najbliższych sąsiadów)

Metody klasyfikacji SVM i wykorzystanie sieci neuronowych w statystycznych systemach uczących
Drzewa klasyfikacyjne i rodziny klasyfikatorów (algorytmy bagging i boosting, lasy losowe) - CART
Analiza regresji (globalne modele parametryczne, regresja nieparametryczna)
Metody rzutowania, wykrywania zmiennych ukrytych: systemy uczące się bez nadzoru (analiza składowych głównych, analiza czynnikowa)

B. Problematyka ćwiczeń laboratoryjnych

Wykorzystanie programu R do dyskryminacji metodami LDA, ODA, FDA, MDA, RDA
Wykorzystanie programu R do praktycznej konstrukcji klasyfikatorów Bayesowskich.
Wykorzystanie programu R do analizy i modelowania metodą KNN
Wykorzystanie programu R do analizy i modelowania metodą SVM
Wykorzystanie programu R do analizy i modelowania z wykorzystaniem sieci neuronowych
Wykorzystanie programu R do tworzenia drzew klasyfikacyjnych i regresyjnych
Algorytmy bagging i boosting, lasy losowe.
Wykorzystanie programu R do tworzenia globalnych modeli regresji parametrycznej i nieparametrycznej
Wykorzystanie programu R do analizy PCA i analizy czynnikowej oraz rola tych analiz w systemach uczących

3.4 Metody dydaktyczne

Wykład: wykład z prezentacją multimedialną

Ćwiczenia laboratoryjne: indywidualne i grupowe 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	Obserwacja w trakcie zajęć, projekt	wykład, lab
EK_02	Obserwacja w trakcie zajęć, projekt	wykład, lab
EK_03	Obserwacja w trakcie zajęć, projekt	wykład, lab

4.2 Warunki zaliczenia przedmiotu (kryteria oceniania)

Warunkiem zaliczenia przedmiotu jest pozytywna ocena z laboratorium obejmująca wiedzę z wykładu oraz umiejętności praktycznych.

Warunkiem zaliczenia laboratorium jest uzyskanie pozytywnej oceny z projektu praktycznego polegającego na przeprowadzeniu pełnej analizy statystycznej z wykorzystaniem statystycznych systemów uczących wybranych danych. Weryfikacja wiedzy jest realizowana podczas „obrony” projektu.

Projekt będzie oceniany na punkty przy czym: (ocena pozytywna >50% punktów), dst 51-59%, dst plus 60-69%, db 70-79%, db plus 80-89%, bdb 90-100%.

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	45
Inne z udziałem nauczyciela akademickiego (udział w konsultacjach, egzaminie)	10
Godziny niekontaktowe – praca własna studenta (przygotowanie do zajęć, egzaminu, napisanie referatu itp.)	45
SUMA GODZIN	100
SUMARYCZNA LICZBA PUNKTÓW ECTS	4

6. PRAKTYKI ZAWODOWE W RAMACH PRZEDMIOTU

wymiar godzinowy	-
zasady i formy odbywania praktyk	-

7. LITERATURA**LITERATURA PODSTAWOWA:**

1. Jacek Koronacki, Jan Ćwik „Statystyczne systemy uczące się”, Wydawnictwo: Wydawnictwa Naukowo Techniczne W-wa 2005
2. Marek Walesiak, Eugeniusz Gantar „Statystyczna Analiza Danych w wykorzystaniem programu R, PWN, W-wa 2012
3. Katarzyna Kopaczewska, Tomasz Kopaczewski, Piotr Wójcik „Metody Ilościowe w R Aplikacje Ekonomiczne i Finansowe, CEDEWU, W-wa 2017
4. Domański C., Pruska K. „Nieklasyczne metody statystyczne”, PWE, Warszawa 2000.
5. Gajek L., Kałuszka M. „Wnioskowanie statystyczne”, WN-T, Warszawa 2000.
6. Przemysław Biecek „Analiza danych z programem R Modele liniowe z efektami stałymi, losowymi i mieszanymi” . PWN, W-wa 2013
7. Przemysław Biecek „Przewodnik po pakiecie R” . (dostępna on-line - (PDF) Przewodnik po pakiecie R | Przemysław Biecek - Academia.edu)

- | |
|--|
| 8. Przemysław Biecek „Na przełaj przez Data Mining z pakietem R” (dostępna on-line - naPrzelajPrzezDM.pdf (biecek.pl)) |
|--|

LITERATURA UZUPEŁNIAJĄCA:

- | |
|---|
| <ol style="list-style-type: none">1. W. Krysiński, J. Bartos, W. Dyczka, K. Królikowska, M. Wasilewski „ <i>Rachunek prawdopodobieństwa i statystyka matematyczna w zadaniach</i>” Tom 1, 2, PWN, Warszawa 1997.2. Krzyśko M., „<i>Statystyka matematyczna</i>”, WN UAM , Poznań 1996. |
|---|