

KARTA PRZEDMIOTU

Kod przedmiotu	0719-2ID-F48-AED	
Nazwa przedmiotu w języku	polskim	Algorytmy eksploracji danych <i>Data Exploration Algorithms</i>
	angielskim	

1. USYTUOWANIE PRZEDMIOTU W SYSTEMIE STUDIÓW

1.1. Kierunek studiów	Inżynieria danych
1.2. Forma studiów	stacjonarne
1.3. Poziom studiów	studia I-stopnia
1.4. Profil studiów	ogólnoakademicki
1.5. Osoba przygotowująca kartę przedmiotu	Mgr Dariusz Pasieka
1.6. Kontakt	dariusz.pasieka@ujk.edu.pl

2. OGÓLNA CHARAKTERYSTYKA PRZEDMIOTU

2.1. Język wykładowy	polski
2.2. Wymagania wstępne	Matematyka dyskretna Statystyka Rachunek prawdopodobieństwa

3. SZCZEGÓŁOWA CHARAKTERYSTYKA PRZEDMIOTU

3.1. Forma zajęć	wykłady, ćwiczenia laboratoryjne, projekt	
3.2. Miejsce realizacji zajęć	zajęcia w pomieszczeniu dydaktycznym UJK	
3.3. Forma zaliczenia zajęć	wykłady – zaliczenie z oceną ćwiczenia laboratoryjne – zaliczenie z oceną projekt – zaliczenie	
3.4. Metody dydaktyczne	Słowne, oglądowe, praktyczne	
3.5. Wykaz literatury	podstawowa	1. T. Morzy, Eksploracja danych, Wydawnictwo Naukowe PWN, Warszawa, 2013
	uzupełniająca	1. J. Koronacki, J. Ćwik, Statystyczne systemy uczące się, Akademicka Oficyna Wydawnicza EXIT Andrzej Lang, Warszawa, 2015 2. T. Hastie, R. Tibshirani, J. Friedman, The Elements of Statistical Learning: Data Mining, Inference, and Prediction, Springer, 2009 3. A. Géron, Uczenie maszynowe z użyciem Scikit-Learn i TensorFlow. Wydanie III, Helion, Gliwice, 2023

4. CELE, TREŚCI I EFEKTY UCZENIA SIĘ

4.1. Cele przedmiotu

Wykład:

- C1. Zapoznanie studentów z podstawowymi pojęciami, definicjami i celami eksploracji danych oraz wyzwaniem związanym z analizą dużych zbiorów danych.
- C2. Przedstawienie rodzajów zadań eksploracji danych, takich jak klasyfikacja, regresja, grupowanie, wykrywanie reguł i analiza semantyczna.
- C3. Omówienie zasad budowy, oceny jakości i walidacji modeli eksploracyjnych, w tym metod testowania i kroswalidacji.
- C4. Zapoznanie studentów z wybranymi algorytmami i technikami eksploracji danych, w szczególności klasyfikatorami bayesowskimi, drzewami decyzyjnymi i algorytmami wykrywania reguł asocjacyjnych.

Laboratorium:

- C1. Nabycie praktycznych umiejętności pracy z danymi w języku Python przy użyciu bibliotek wspierających analizę danych.
- C2. Opanowanie metod wstępnego przetwarzania i transformacji danych, takich jak kodowanie, normalizacja, standaryzacja czy redukcja wymiarów.
- C3. Rozwinięcie umiejętności implementacji i testowania klasyfikatorów oraz metod regresyjnych, a także oceny ich jakości.
- C4. Pozyskanie doświadczenia w stosowaniu metod grupowania, wykrywania reguł asocjacyjnych i analizy

semantycznej w praktycznych przykładach eksploracji danych.

Projekt:

1. Rozwijanie umiejętności samodzielnej analizy i interpretacji rzeczywistych danych poprzez realizację projektu zespołowego z wykorzystaniem metod eksploracji danych.

4.2. Treści programowe

Wykład:

1. Wprowadzenie do eksploracji danych: definicje, cele, wyzwania,
2. Rodzaje zadań eksploracji danych: identyfikacja rozkładu, przetwarzanie i transformacja danych. klasyfikacja, regresja, wykrywanie reguł i wzorców,
3. Grupowanie i segmentacja danych,
4. Ocena jakości maszyn klasyfikujących i regresyjnych, testowanie, krosvalidacja,
5. Klasyfikator Bayesa, założenie naiwne, optymalny klasyfikator bayesowski,
6. Drzewa klasyfikujące, budowa drzewa, miary zanieczyszczenia, przycinanie,
7. Wykrywanie reguł asocjacyjnych, wsparcie, zaufanie, algorytm a priori
8. Analiza semantyczna, przetwarzanie języka naturalnego.

Laboratorium:

1. Narzędzia analizy danych, analiza danych przy pomocy języka programowania Python i bibliotek scikit-learn i pandas, wczytywanie danych, publiczne repozytoria danych
2. Wstępne przetwarzanie danych: kodowanie gorąco-jedynkowe, normalizacja, standaryzacja,
3. Naiwny klasyfikator bayesowski
4. Testowanie klasyfikatora: metoda krossvalidacji, głosowania większościowego, macierz pomyłek
5. Klasyfikatory: drzewa decyzyjne, losowe lasy, leniwy algorytm k-sąsiadów
6. Analiza składowych głównych, liniowa analiza dyskryminacyjna - zastosowanie do wizualizacji danych
7. Przewidywanie ciągłych zmiennych docelowych za pomocą analizy regresywnej
8. Algorytm wyszukiwania reguł asocjacyjnych (implementacja) warianty a priori
9. Przetwarzanie danych tekstowych za pomocą analizy semantycznej: alokacja ukrytej zmiennej Dirichleta
10. Praca z nieoznakowanymi danymi: analiza skupień

Projekt:

1. Studenci w zespołach tworzą własną analizę danych z wykorzystaniem metod eksploracji danych oraz ich interpretacja.

4.3. Przedmiotowe efekty uczenia się

Efekt	Student, który zaliczył przedmiot	Odniesienie do kierunkowych efektów uczenia się
w zakresie WIEDZY:		
W01	rozdzieli podstawowe pojęcia, cele i zadania eksploracji danych oraz zna wybrane technologie informatyczne stosowane w analizie danych	ID1A_W06
W02	zna wybrane algorytmy analizy i eksploracji danych, w tym klasyfikacji, regresji, grupowania i wykrywania reguł	ID1A_W08
w zakresie UMIEJĘTNOŚCI:		
U01	potrafi wykorzystywać oprogramowanie i narzędzia informatyczne w celu gromadzenia, przetwarzania, wizualizacji i interpretacji danych	ID1A_U06 ID1A_U07
U02	potrafi rozpoznać problemy analityczne możliwe do rozwiązania algorytmicznie oraz dobrać odpowiednie algorytmy i techniki eksploracji danych	ID1A_U07 ID1A_U08
U03	potrafi implementować i testować wybrane algorytmy eksploracji danych z wykorzystaniem języka programowania wysokiego poziomu	ID1A_U07
U04	potrafi analizować i rozwiązywać złożone problemy związane z eksploracją danych oraz dobrać metody w warunkach nie w pełni przewidywalnych	ID1A_U08
w zakresie KOMPETENCJI SPOŁECZNYCH:		
K01	jest w stanie wymagania biznesowe i społeczne zastosować w realizowanym projekcie	ID1A_K02 ID1A_K03
K02	jest świadomy konieczności permanentnej modyfikacji wdrożonego projektu w aspektach merytorycznych i społecznych	ID1A_K04

4.4. Sposoby weryfikacji osiągnięcia przedmiotowych efektów uczenia się																		
Efekty przedmiotowe (symbol)	Sposób weryfikacji (+/-)																	
	Egzamin pisemny			Projekt			Aktywność na zajęciach			Praca własna			Praca w grupie					
	Forma zajęć			Forma zajęć			Forma zajęć			Forma zajęć			Forma zajęć					
	W	C	P	W	C	P	W	C	P	W	C	P	W	C	P			
W01							+	+										
W02							+	+										
U01						+					+				+			
U02						+					+				+			
U03						+					+				+			
U04						+					+				+			
K01						+									+			
K02						+									+			

4.5. Kryteria oceny stopnia osiągnięcia efektów uczenia się		
Forma zajęć	Ocena	Kryterium oceny
wykład (W)	3	osiągnięcie <50-60) % wymogów stosowanych w metodach oceny
	3,5	osiągnięcie <61-70) % wymogów stosowanych w metodach oceny
	4	osiągnięcie <71-80) % wymogów stosowanych w metodach oceny
	4,5	osiągnięcie <81-90) % wymogów stosowanych w metodach oceny
	5	osiągnięcie <91-100> % wymogów stosowanych w metodach oceny
laboratorium (L)	3	osiągnięcie <50-60) % wymogów stosowanych w metodach oceny
	3,5	osiągnięcie <61-70) % wymogów stosowanych w metodach oceny
	4	osiągnięcie <71-80) % wymogów stosowanych w metodach oceny
	4,5	osiągnięcie <81-90) % wymogów stosowanych w metodach oceny
	5	osiągnięcie <91-100> % wymogów stosowanych w metodach oceny
Projekt (P)	3	osiągnięcie <50-60) % wymogów stosowanych w metodach oceny
	3,5	osiągnięcie <61-70) % wymogów stosowanych w metodach oceny
	4	osiągnięcie <71-80) % wymogów stosowanych w metodach oceny
	4,5	osiągnięcie <81-90) % wymogów stosowanych w metodach oceny
	5	osiągnięcie <91-100> % wymogów stosowanych w metodach oceny

5. BILANS PUNKTÓW ECTS – NAKŁAD PRACY STUDENTA

Kategoria	Obciążenie studenta	
	Studia stacjonarne	Studia niestacjonarne
LICZBA GODZIN REALIZOWANYCH PRZY BEZPOŚREDNIM UDZIALE NAUCZYCIELA /GODZINY KONTAKTOWE/	90	
Udział w wykładach*	30	
Udział w ćwiczeniach, konwersatoriach, laboratoriach*	30	
Projekt	30	
SAMODZIELNA PRACA STUDENTA /GODZINY NIEKONTAKTOWE/	60	
Przygotowanie do wykładu*	10	
Przygotowanie do ćwiczeń, konwersatorium, laboratorium*	20	
Przygotowanie do egzaminu/kolokwium*	15	
Zebranie materiałów do projektu, kwerenda internetowa*	15	
ŁĄCZNA LICZBA GODZIN	150	
PUNKTY ECTS za przedmiot	6	

*niepotrzebne usunąć

Przyjmuję do realizacji (data i czytelne podpisy osób prowadzących przedmiot w danym roku akademickim)

.....