

KARTA PRZEDMIOTU

Inżynieria danych <i>nazwa przedmiotu</i>
Data engineering <i>nazwa przedmiotu w języku angielskim</i>
polski <i>język wykładowy</i>
przedmioty specjalnościowe <i>kategoria przedmiotu/grupa zajęć</i>

Cykl kształcenia rozpoczynający się od: rok akademicki 2024/25, semestr letni

Jednostka organizacyjna: Wydział Inżynierii Lądowej

Kierunek studiów: Transport - stacjonarne

Specjalność: Logistyka i spedycja

Profil studiów: ogólnoakademicki

Poziom studiów: II stopnia

Forma studiów: stacjonarne

Wymagania wstępne:

1. Zaliczenie z przedmiotu: Narzędzia informatyczne w analizie danych
2. Podstawowa wiedza i umiejętności w zakresie programowania w języku Python.

Cele przedmiotu:

1. Zapoznanie studentów z pojęciem inżynierii danych oraz znaczeniem danych w logistyce i transporcie.
2. Rozwinięcie umiejętności przetwarzania, analizy i klasyfikacji dużych zbiorów danych z wykorzystaniem zaawansowanych metod uczenia maszynowego.
3. Udoskonalenie zdolności wizualizacji danych i interpretacji wyników w kontekście procesów logistycznych.

Efekty uczenia się:

Kod efektu uczenia się	Opis efektu uczenia się	Kod kierunkowego efektu uczenia się
Wiedza Absolwent zna i rozumie:		
EW1	Student(ka) zna podstawowe pojęcia inżynierii danych oraz rozumie rolę i znaczenie danych w transporcie i logistyce, w tym procesy przetwarzania i wzbogacania danych.	T2_W01
EW2	Student(ka) opisuje wybrane metody klasyfikacji danych (np. drzewa decyzyjne, klasyfikatory Bayesowskie) oraz metody klasteryzacji (np. metoda najbliższych sąsiadów, miary odległości), i rozumie, jak są stosowane do analizy dużych zbiorów danych w logistyce.	T2_W03
Umiejętności Absolwent potrafi:		
EU1	Student(ka) analizuje duże zbiory danych związane z procesami logistycznymi, przetwarza je i stosuje odpowiednie metody klasyfikacji oraz klasteryzacji w dedykowanym środowisku programistycznym.	T2_U06
EU2	Student(ka) potrafi opracować raport zawierający wyniki analizy oraz przejrzystą wizualizację danych. Student(ka) zdobywa przygotowanie do realizacji badań naukowych.	T2_U03
Kompetencje społeczne Absolwent jest gotów do:		
EK1	Student(ka) interpretuje, analizuje i ocenia uzyskane rezultaty rozwiązań zadań problemowych w sposób rzetelny i z uwzględnieniem zasad etyki.	T2_K10

Forma zajęć, semestralna liczba godzin:

Semestr	Punkty ECTS	Forma zaliczenia (E/Z)	Wykłady (W)	Ćwiczenia (C)	Laboratoria (L)	Laboratoria komputerowe (LK)	Projekty (P)	Seminaria (S)	Praktyka zawodowa (PZ)
2	3	Z	10	0	0	0	20	0	0

E - egzamin; Z - zaliczenie

Treści programowe:

Lp.	Forma zajęć	Tematyka zajęć	Liczba godzin
1	W	1. Wprowadzenie do inżynierii danych: omówienie kluczowych pojęć z zakresu inżynierii danych, takich jak gromadzenie, przetwarzanie i analiza danych. Zostaną przedstawione etapy cyklu życia danych, od pozyskiwania aż po interpretację wyników. Studenci poznają również różnice między inżynierią danych a analizą danych oraz jak te procesy wspierają zarządzanie informacją w organizacjach logistycznych. 2. Znaczenie i wykorzystanie danych w transporcie/logistyce - zajęcia wprowadzą studentów do roli danych w logistyce i transporcie, omawiając przykłady wykorzystania danych np. w zarządzaniu flotą, optymalizacji tras, monitorowaniu łańcucha dostaw oraz przewidywaniu popytu na usługi transportowe. Poruszona zostanie kwestia, jak zbieranie i analiza danych mogą wpłynąć na zwiększenie efektywności operacyjnej. 3. Przykłady dużych zbiorów danych w transporcie/logistyce. Ocena przydatności danych; wstępne przetwarzanie danych, wzbogacanie danych studenci zapoznają się z typowymi źródłami danych w logistyce. Zostanie omówiony proces oceny jakości danych, metody wstępnego przetwarzania, takie jak czyszczenie danych, usuwanie braków, a także techniki wzbogacania danych poprzez łączenie różnych źródeł informacji.	4
2	W	4. Klasyfikacja danych (drzewa decyzyjne, klasyfikatory Bayesowskie, klasyfikacja binarna i wieloklasowa) zajęcia będą dotyczyły metod klasyfikacji danych, takich jak drzewa decyzyjne i klasyfikatory Bayesowskie. Studenci poznają różnice między klasyfikacją binarną a wieloklasową oraz nauczą się, jak dobierać odpowiednie metody klasyfikacji do analizowanych danych logistycznych, np. do przewidywania opóźnień dostaw lub identyfikacji anomalii w przepływach towarów. 5. Wybrane metody klasteryzacji i analizy podobieństwa (m.in. metody najbliższych sąsiadów, miary odległości do obliczania podobieństwa) - studenci poznają metody klasteryzacji, takie jak metoda najbliższych sąsiadów i miary odległości, które są stosowane w analizie podobieństw między obiektami logistycznymi (np. centrami dystrybucji) oraz grupowaniu podobnych tras lub przesyłek w systemach zarządzania transportem. 6. Wizualizacja danych w transporcie/logistyce - omówienie technik wizualizacji danych, w tym wykresów punktowych, wykresów mapowych, diagramów słupkowych i diagramów przepływów, które pozwalają na prezentację danych logistycznych w czytelny sposób.	6
3	P	Analiza dużego zbioru danych dotyczącego wybranego procesu logistycznego - studenci będą analizować duży zbiór danych związany z wybranym procesem logistycznym, np. dane o transporcie, optymalizacji tras, zarządzaniu zapasami lub czasach dostaw. Zadanie będzie polegało na: 1. Opracowaniu schematu postępowania, który obejmuje wstępne przetwarzanie, czyszczenie, wzbogacanie oraz analizę danych. 2. Implementacji wybranego algorytmu klasyfikacji lub klasteryzacji w środowisku Python (np. za pomocą bibliotek scikit-learn, Pandas, NumPy).	14
4	P	3. Prezentacji wyników w postaci raportu zawierającego wykresy, tabele, wizualizacje oraz analizę statystyczną danych. Studenci będą zobowiązani do omówienia wyników i przedstawienia możliwych usprawnień w analizowanym procesie logistycznym.	5
5	P	Analiza i interpretacja uzyskanych rezultatów rozwiązywanych zadań problemowych w ramach zajęć.	1

Praca własna studenta:

Lp.	Opis pracy własnej	Liczba godzin
1	Przygotowanie się do zajęć, w tym studiowanie wymaganej i zalecanej literatury.	24
2	Konsultacje przedmiotowe (w bezpośrednim kontakcie z prowadzącym).	10
3	Zaliczenia w sesji (w bezpośrednim kontakcie z prowadzącym).	6
4	Przygotowanie sprawozdań, raportów, projektów, prezentacji.	5

Metody dydaktyczne:

konsultacje, projekt, symulacja, wykład, prezentacja

Sposoby weryfikacji i oceny efektów uczenia się:

uczestnictwo w zajęciach, kolokwium, odpowiedź ustna, prezentacja, rozwiązanie zadania problemowego, sprawozdanie z pracy indywidualnej, sprawdzian wiadomości z wykładów, sprawozdanie

Kryteria oceny:

1. Wymaganie na ocenę 5,0: Student(ka) dogłębnie rozumie kluczowe pojęcia inżynierii danych, wtym znaczenie i wykorzystanie danych w logistyce i transporcie.
2. Wymaganie na ocenę 5,0: Student(ka) potrafi precyzyjnie wyjaśnić metody klasyfikacji (drzewa decyzyjne, klasyfikatory Bayesowskie) oraz klasteryzacji (KNN, miary odległości) i ich zastosowanie do analiz danych logistycznych.
3. Wymaganie na ocenę 5,0: Student(ka) samodzielnie analizuje duże zbiory danych logistycznych, stosując zaawansowane metody przetwarzania, klasyfikacji i klasteryzacji.
4. Wymaganie na ocenę 5,0: Student(ka) potrafi implementować algorytmy oraz przygotować kompleksowy raport zawierający szczegółowe wyniki analizy danych, wizualizacje oraz wnioski dotyczące ulepszeń procesów logistycznych.
5. Wymaganie na ocenę 5,0: Student(ka) poprawnie wykonał wszystkie zadania rozwiązywane w ramach zajęć projektowych, przygotował(a) wszystkie sprawozdania z zajęć, interpretację uzyskanych wyników oraz sformułował(a) wnioski. Realizuje zadania zgodnie z zasadami etyki.

Literatura:

obowiązkowa:

1. Geron A. Uczenie maszynowe z użyciem Scikit-Learn, Keras i TensorFlow, Wydanie 3, Wydawnictwo Helion.
2. McKinney W., Python w analizie danych, wyd. 3, Wydawnictwo Helion.

zalecana/fakultatywna:

1. Hastie, T., Tibshirani, R., Friedman, J. (2009). The Elements of Statistical Learning: Data Mining, Inference, and Prediction (2nd Edition). Springer.