



KARTA OPISU PRZEDMIOTU - SYLABUS

Nazwa przedmiotu

Eksploracyjna analiza danych

Przedmiot

Kierunek studiów

Automatyka i robotyka

Studia w zakresie (specjalność)

RISA

Poziom studiów

drugiego stopnia

Forma studiów

stacjonarne

Rok/semestr

1/2

Profil studiów

ogólnoakademicki

Język oferowanego przedmiotu

polski

Wymagalność

obieralny

Liczba godzin

Wykład

15

Ćwiczenia

Laboratoria

30

Projekty/seminaria

Inne (np. online)

Liczba punktów

3

Wykładowcy

Odpowiedzialny za przedmiot/wykładowca:

dr inż. Piotr Kaczmarek

Odpowiedzialny za przedmiot/wykładowca:

Wymagania wstępne

Student rozpoczynający ten przedmiot powinien posiadać podstawową wiedzę z programowania w języku Python, narzędzi i metod uczenia maszynowego oraz podstaw statystyki matematycznej z zakresu I stopnia studiów inżynierskich oraz 1 semestru studiów II stopnia.

Cel przedmiotu

Cel modułu kształcenia:

1. Umiejętność zastosowanie metody graficznej analizy i eksploracji danych
2. Umiejętność testowania hipotez, oraz wyboru właściwej metody testowania



3. Umiejętność przetworzenia danych w tym: wykrywania elementów odstających, zastosowania nienadzorowanych metod analizy skupień, redukcja wymiarowości

Przedmiotowe efekty uczenia się

Wiedza

1. Ma uporządkowaną wiedzę w zakresie metod analizy danych oraz narzędzi statystycznych pozwalających na ocenę istotności wyciąganych wniosków.
2. Ma wiedzę z zakresu statystyki w tym sposobów testowania hipotez oraz metod planowania eksperymentu (DoE)
3. Ma wiedzę z zakresu metod analizy skupień

Umiejętności

1. Potrafi wczytać oraz scalić dane z różnych źródeł oraz przeanalizować je pod kątem spójności kompletności jak również zastosować odpowiednią strategię ich czyszczenia
2. Potrafi przeprowadzić analizę problemu oraz dobrać właściwe metody testowania hipotez
3. Potrafi przeprowadzić analizę danych pod kątem ich podobieństwa, skupień i zastosować metody pozwalające na redukcję ich wymiarowości
4. Potrafi analizować i interpretować projektową dokumentację techniczną oraz wykorzystywać literaturę naukową związaną z danym problemem, a także dostrzegać możliwość wykorzystania nowych technik i technologii. Potrafi innowacyjnie wykonywać zadania w nieprzewidywalnych warunkach.

Kompetencje społeczne

1. Posiada świadomość konieczności profesjonalnego podejścia do zagadnień technicznych, skrupulatnego zapoznania się z dokumentacją oraz warunkami środowiskowymi, w których urządzenia i ich elementy mogą funkcjonować. Jest gotów do przestrzegania zasad etyki zawodowej i wymagania tego od innych, poszanowania różnorodności poglądów i kultur.

Metody weryfikacji efektów uczenia się i kryteria oceny

Efekty uczenia się przedstawione wyżej weryfikowane są w następujący sposób:

Wykład: egzamin w formie pisemnej z zakresu wykładu

Laboratorium: realizacja dwóch zadań projektowych obejmujących praktyczne zastosowanie poznawanych metod oraz ocena pracy na zajęciach

Treści programowe

Pogram wykładu i zajęć laboratoryjnych obejmuje następujące zagadnienia:



- metody graficznej analizy danych – sposoby wizualizacji danych wielowymiarowych, wczytywanie danych (w tym praktyczne wykorzystanie modułu pandas, oraz podstawowa obsługa baz danych (SQL) i zasobów sieciowych (RESTfull API)
- rozkład próbek, metody wyznaczania statystyk rozkładu oraz oceny przedziału ufności (w tym metody klasyczne oparte o twierdzenie graniczne oraz metody typu bagging)
- metody testowania hipotez (prawdopodobieństwo testowe, błędy I i II rodzaju, moc testu, testy A/B, permutacyjne, T-Test, ANOVA, analiza wieloczynnikowa, multi-arm bandit)
- zapoznanie z podstawowymi metodami planowania eksperymentu (design of experiments)
- Metody analizy skupień, wyznaczania zależności między danymi (metryki w zbiorach wielowymiarowych, redukcja wymiarowości, ocena istotności danych, nienadzorowana analiza skupień)

Metody dydaktyczne

1. Wykład: prezentacja multimedialna, ilustrowana przykładami podawanymi na tablicy, oraz programami tworzonymi w trakcie zajęć.
2. Ćwiczenia laboratoryjne: samodzielne ćwiczenie materiału wspomagane materiałami dydaktycznymi umieszczanymi na platformie e-learningowej

Literatura

Podstawowa

1. R Schutt, C. O’Niel „Badanie danych – Raport z pierwszej linii działań” O’Reilly
2. Data mining : concepts and techniques / Jiawei Han, Micheline Kamber, Jian Pei

Uzupełniająca

3. Bruce, Peter, Andrew Bruce, and Peter Gedeck. Practical Statistics for Data Scientists: 50+ Essential Concepts Using R and Python. O’Reilly Media, 2020. and Data Structures using Python” Luther College 2018 (dostępna online)



Bilans nakładu pracy przeciętnego studenta

	Godzin	ECTS
Łączny nakład pracy	75	3
Zajęcia wymagające bezpośredniego kontaktu z nauczycielem	45	2
Praca własna studenta (studia literaturowe, przygotowanie do zajęć laboratoryjnych, przygotowanie do egzaminu, wykonanie dwóch projektów) ¹	30	1

¹

niepotrzebne skreślić lub dopisać inne czynności

