

KARTA PRZEDMIOTU

Nazwa w języku polskim: **Wprowadzenie do analizy danych II**

Nazwa w języku angielskim: **Introduction to Data Science II**

Kierunek studiów (jeśli dotyczy): **Transport**

Specjalność (jeśli dotyczy): **Organizacja i Projektowanie Systemów Transportowych**

Poziom i forma studiów: **II stopień, stacjonarne**

Rodzaj przedmiotu: **wybieralny**

Kod przedmiotu: **TRM041111**

Grupa kursów: **nie**

	Wykład	Ćwiczenia	Laboratorium	Projekt	Seminarium
Liczba godzin zajęć zorganizowanych w Uczelni (ZZU)	15			15	
Liczba godzin całkowitego nakładu pracy studenta (CNPS)	30			30	
Forma zaliczenia	Zaliczenie na ocenę			Zaliczenie na ocenę	
Grupa kursów					
Liczba punktów ECTS	1			1	
w tym liczba punktów odpowiadająca zajęciom o charakterze praktycznym (P)				1	
w tym liczba punktów ECTS odpowiadająca zajęciom wymagającym bezpośredniego kontaktu (BK)	0.6			0.7	

WYMAGANIA WSTĘPNE W ZAKRESIE WIEDZY, UMIEJĘTNOŚCI I KOMPETENCJI SPOŁECZNYCH

CELE PRZEDMIOTU

- C1. Utrwalenie wiedzy z zakresu sposobów pozyskiwania, przechowywania i eksploracji danych.
- C2. Poznanie technologii i narzędzi stosowanych w analizie danych
- C3. Rozwijanie umiejętności analitycznego rozwiązywania problemów z użyciem technologii komputerowej

PRZEDMIOTOWE EFEKTY UCZENIA SIĘ

I. Z zakresu wiedzy:

PEK_W01 - Student ma pogłębioną wiedzę o technologiach i technikach stosowanych w analizie danych.

PEK_W02 - Student rozumie i zna aspekty wydajnościowe związane z przetwarzaniem danych

II. Z zakresu umiejętności:

PEK_U01 - Student potrafi przeprowadzić poszerzoną eksplorację danych.

PEK_U02 - Student potrafi jednoznacznie zdefiniować problem i użyć narzędzia programistyczne do jego rozwiązania.

III. Z zakresu kompetencji społecznych:

PEK_K01 - Student potrafi w grupie zaplanować i przeprowadzać wnioskowanie rozwiązując złożone problemy decyzyjne.

TREŚCI PROGRAMOWE

Forma zajęć – Wykład		Liczba godzin
Wy1	Wprowadzenie do eksploracji i ekstrakcji danych z zastosowaniem języka Python.	1
Wy2	Algebra liniowa, rachunek prawdopodobieństwa i statystyka z wykorzystaniem bibliotek języka Python.	2
Wy3	Analiza szeregów czasowych.	1
Wy4	Pozyskiwanie, przechowywanie i przetwarzanie danych.	2
Wy5	Wydajność obliczeniowa w analizie danych - techniki przetwarzania równoległego danych.	1
Wy6	Czyszczenie, przygotowanie i agregacja danych.	2
Wy7	Potok danych i Big Data.	1
Wy8	Wprowadzenie do uczenia maszynowego.	1
Wy9	Uczenie maszynowe i głębokie uczenie z wykorzystaniem języka Python.	2
Wy10	Zastosowanie języków R i Julia w analizie danych.	1
Wy11	Sprawdzian	1
		Suma: 15
Forma zajęć – Projekt		Liczba godzin
Proj1	Techniki gromadzenia danych wejściowych	1
Proj2	Przetwarzanie wsadowe - pisanie skryptów w języku Python	1
Proj3	Ekstrakcja danych z internetu.	2
Proj4	Normalizacja danych	2
Proj5	Czyszczenie, przygotowanie i agregacja zgromadzonych danych.	2
Proj6	Określenie założeń i celu analizy	1

Proj7	Identyfikacja i pozyskanie danych	3
Proj8	Przeprowadzenie analizy danych i przygotowanie raportu	3
		Suma: 15

STOSOWANE NARZĘDZIA DYDAKTYCZNE

N1. wykład tradycyjny z wykorzystaniem transparencji i slajdów
N2. ćwiczenia problemowe
N3. przygotowanie sprawozdania

OCENA OSIĄGNIĘCIA PRZEDMIOTOWYCH EFEKTÓW UCZENIA SIĘ (Wykład)

Oceny (F – formująca (w trakcie semestru), P – podsumowująca (na koniec semestru))	Numer efektu uczenia się	Sposób oceny osiągnięcia efektu uczenia się
F1	PEK_W01, PEK_W02, PEK_K01	Sprawdzian pisemny
P =		

OCENA OSIĄGNIĘCIA PRZEDMIOTOWYCH EFEKTÓW UCZENIA SIĘ (Projekt)

Oceny (F – formująca (w trakcie semestru), P – podsumowująca (na koniec semestru))	Numer efektu uczenia się	Sposób oceny osiągnięcia efektu uczenia się
F1	PEK_U01, PEK_U02, PEK_K01	Sprawozdanie
P =		

LITERATURA PODSTAWOWA I UZUPEŁNIAJĄCA

LITERATURA PODSTAWOWA

1. Gągolewski M., Bartoszek M., Cena A., Przetwarzanie i analiza danych w języku Python, PWN, Warszawa 2016
2. Szeliga M., Data science i uczenie maszynowe, PWN, Warszawa 2017
3. Boschetti A., Massaron L., Python. Podstawy nauki o danych, Wydanie II, Helion, 2017
4. Grus J., Data science od podstaw. Analiza danych w Pythonie, Helion 2018
5. McKinney W., Python w analizie danych, Helion 2018

LITERATURA UZUPEŁNIAJĄCA

1. Mitchell R., Ekstrakcja danych z językiem Python, Helion, 2019
2. Lott S. F., Python. Programowanie funkcyjne, Helion, 2019
3. Rashka S., Mirjalili V., Python. Uczenie maszynowe, Helion 2019
4. Zocca V., Spacagna G., Slater D., Foelants P., Deep Learning, Uczenie głębokie z językiem Python, Helion, 2018
5. Biecek P., Analiza danych z programem R, Wydawnictwo Naukowe PWN, Warszawa 2011
6. Zagdański A., Suchwałko A., Analiza i prognozowanie szeregów czasowych. Praktyczne wprowadzenie na podstawie środowiska R, PWN, Warszawa 2016

OPIEKUN PRZEDMIOTU

dr inż. Piotr Czaja tel.: 320-26-64 email: piotr.czaja@pwr.edu.pl

SUBJECT CARD

Name in Polish: **Wprowadzenie do analizy danych II**

Name in English: **Introduction to Data Science II**

Main field of study (if applicable): **Transport**

Specialization (if applicable): **Transportation Systems Management and Designing**

Level and form of studies: **II level, full-time**

Kind of subject: **optional**

Subject code: **TRM041111**

Group of courses: **no**

	Lecture	Classes	Laboratory	Project	Seminar
Number of hours of organized classes in University (ZZU)	15			15	
Number of hours of total student workload (CNPS)	30			30	
Form of crediting	Crediting with grade			Crediting with grade	
Group of courses					
Number of ECTS points	1			1	
including number of ECTS points for practical (P) classes				1	
including number of ECTS points for direct teacher-student contact (BK) classes	0.6			0.7	

PREREQUISITES RELATING TO KNOWLEDGE, SKILLS AND OTHER COMPETENCES

SUBJECT OBJECTIVES

- C1. Consolidation of knowledge in the field of data acquisition, storage and data mining.
- C2. Understanding the technologies and tools used in data analysis
- C3. Developing skills in analytical problem solving using computer technology

SUBJECT LEARNING OUTCOMES

I. Relating to knowledge:

PEK_W01 - The student has in-depth knowledge of technologies and techniques used in data analysis.

PEK_W02 - The student understands and knows the performance aspects related to data processing

II. Relating to skills:

PEK_U01 - The student is able to carry out extended data mining.

PEK_U02 - The student is able to clearly define the problem and use programming tools to solve it.

III. Relating to social competences:

PEK_K01 - Student is able to plan and carry out conclusions in a group solving complex decision problems.

PROGRAM CONTENT

Form of classes – Lecture		Number of hours
Lec1	Introduction to data mining in Python	1
Lec2	Linear algebra, probability theory and statistics using Python libraries.	2
Lec3	Time series analysis.	1
Lec4	Collecting, storage and processing of data.	2
Lec5	Computational performance in data analysis - parallel data processing techniques.	1
Lec6	Cleaning, preparation and aggregation of data.	2
Lec7	Data streaming and Big Data.	1
Lec8	Introduction to machine learning.	1
Lec9	Machine learning and deep learning in Python.	2
Lec10	Application of R and Julia languages in data analysis.	1
Lec11	Test	1
		Total hours: 15
Form of classes – Project		Number of hours
Proj1	Input data collection techniques	1
Proj2	Batch processing - writing scripts in Python	1
Proj3	Web scraping.	2
Proj4	Data normalization	2
Proj5	Cleaning, preparation and aggregation of collected data.	2
Proj6	Defining the assumptions and purpose of the analysis	1
Proj7	Identification and data acquisition	3

Proj8	Performing data analysis and preparing the final report	3
		Total hours: 15

TEACHING TOOLS USED		
N1. traditional lecture with the use of transparencies and slides N2. problem exercises N3. report preparation		

EVALUATION OF SUBJECT LEARNING OUTCOMES ACHIEVEMENT (Lecture)		
Evaluation (F – forming (during semester), P – concluding (at semester end))	Learning outcomes number	Way of evaluating learning outcomes achievement
F1	PEK_W01, PEK_W02, PEK_K01	Written test
P =		

EVALUATION OF SUBJECT LEARNING OUTCOMES ACHIEVEMENT (Project)		
Evaluation (F – forming (during semester), P – concluding (at semester end))	Learning outcomes number	Way of evaluating learning outcomes achievement
F1	PEK_U01, PEK_U02, PEK_K01	Report
P =		

PRIMARY AND SECONDARY LITERATURE

PRIMARY LITERATURE

1. Gągolewski M., Bartoszek M., Cena A., Przetwarzanie i analiza danych w języku Python, PWN, Warszawa 2016
2. Szeliga M., Data science i uczenie maszynowe, PWN, Warszawa 2017
3. Boschetti A., Massaron L., Python Data Science Essentials, Second Edition, Packt Publishing, 2016
4. Grus J., Data Science from Scratch: First Principles with Python, O'Reilly, 2015
5. McKinney W., Python for Data Analysis: Data Wrangling with Pandas, NumPy and Ipython, O'Reilly, 2018

SECONDARY LITERATURE

1. Mitchell R., Web Scraping with Python: Collecting More Data from the Modern Web, O'Reilly, 2018
2. Lott S. F., Functional Python Programming, Packt Publishing 2018
3. Rashka S., Mirjalili V., Python Machine Learning. Machine Learning and Deep Learning with Python, scikit-learn and TensorFlow, Packt Publishing 2017
4. Zocca V., Spacagna G., Slater D., Foelants P., Python Deep Learning, Packt Publishing 2017
5. Biecek P., Analiza danych z programem R, Wydawnictwo Naukowe PWN, Warszawa 2011
6. Zagdański A., Suchwałko A., Analiza i prognozowanie szeregów czasowych. Praktyczne wprowadzenie na podstawie środowiska R, PWN, Warszawa 2016

SUBJECT SUPERVISOR

dr inż. Piotr Czaja tel.: 320-26-64 email: piotr.czaja@pwr.edu.pl