

WYDZIAŁ / STUDIUM.....	
KARTA PRZEDMIOTU	
Nazwa przedmiotu w języku polskim Przetwarzanie danych i odkrywanie wiedzy	
Nazwa przedmiotu w języku angielskim Data processing and knowledge discovery	
Kierunek studiów (jeśli dotyczy): Sztuczna inteligencja	
Specjalność (jeśli dotyczy):	
Poziom i forma studiów:	I / II stopień / jednolite studia magisterskie* , stacjonarna / niestacjonarna*
Rodzaj przedmiotu:	obowiązkowy / wybieralny / ogólnouczelniany *
Kod przedmiotu
Grupa kursów	TAK / NIE*

	Wykład	Ćwiczenia	Laboratorium	Projekt	Seminarium
Liczba godzin zajęć zorganizowanych w Uczelni (ZZU)	15		15		
Liczba godzin całkowitego nakładu pracy studenta (CNPS)	60		30		
Forma zaliczenia	Egzamin / zaliczenie na ocenę*	Egzamin / zaliczenie na ocenę*	Egzamin / zaliczenie na ocenę*	Egzamin / zaliczenie na ocenę*	Egzamin / zaliczenie na ocenę*
Dla grupy kursów zaznaczyć kurs końcowy (X)					
Liczba punktów ECTS	2		1		
w tym liczba punktów odpowiadająca zajęciom o charakterze praktycznym (P)			1		
w tym liczba punktów ECTS odpowiadająca zajęciom wymagającym bezpośredniego udziału nauczycieli lub innych osób prowadzących zajęcia (BU)	1.2		0.6		

*niepotrzebne skreślić

WYMAGANIA WSTĘPNE W ZAKRESIE WIEDZY, UMIEJĘTNOŚCI I KOMPETENCJI SPOŁECZNYCH
<ol style="list-style-type: none"> 1. Posiadanie podstawowej wiedzy z zakresu uczenia maszynowego 2. Umiejętność programowania w języku Python 3. Posiadanie podstawowych umiejętności z zakresu wykorzystania systemów z rodziny Unix

CELE PRZEDMIOTU
<p>C1 Zapoznanie studentów z potokiem przetwarzania w zagadnieniach analizy danych i uczenia maszynowego.</p> <p>C2 Nabycie umiejętności podstawowej analizy danych i doboru metod uczenia maszynowego.</p>

C3 Nabycie umiejętności wdrażania prostych aplikacji wykorzystujących modele uczenia maszynowego.

PRZEDMIOTOWE EFEKTY UCZENIA SIĘ

Z zakresu wiedzy:

KSI_W10 - Ma wiedzę na temat metod pozyskiwania wiedzy z danych różnych typów danych, ich transformacji i walidacji oraz potencjalnych zastosowań

Z zakresu umiejętności:

KSI_U03 - Potrafi sformułować problem do rozwiązania, zebrać i oczyścić dane, dobrać metodę oraz przeprowadzić eksperymenty a otrzymane wyniki poddać krytycznej analizie, dokonać ich interpretacji i prezentacji

KSI_U04 - Potrafi dokonać oceny rozwiązania w zakresie pozyskania danych, ich przetwarzania oraz analizy a także ekstrakcji wiedzy a także zaproponować jego ulepszenie.

Z zakresu kompetencji społecznych:

KSI_K01 - Jest gotów do krytycznej oceny odbieranych treści, ma świadomość znaczenia wiedzy w rozwiązywaniu problemów.

TREŚCI PROGRAMOWE

Forma zajęć - wykład		Liczba godzin
Wy1	Wprowadzenie do kursu. Business Analytics (BI), Data Analytics i Data Science. Proces odkrywania wiedzy. Etapy: definicja problemu, pozyskiwanie danych, czyszczenie danych, modelowanie danych, ewaluacja, komunikowanie wyników. Studium przypadku demonstrujące kolejne etapy procesu.	2
Wy2	Data Model Canvas. Data Science Workflow Canvas. Definiowanie problemu odkrywania wiedzy.	2
Wy3	Pozyskiwanie danych. Anotacja danych (miary zgodności). Integracja różnych źródeł danych. Pomiar jakości danych. Transformacja danych (one-hot, kodowanie, standaryzacja, normalizacja). Czyszczenie danych (eliminacja, imputacja, ocena cech).	2
Wy4	Modelowanie danych. Inżynieria cech (ekstrakcja, generowanie, selekcja). Rodzaje cech (grupowe, listy, sekwencje, serie czasowe, sygnały, inne).	2
Wy5	Uczenie i ewaluacja. Podział zbiorów uczących/testujących. Stratyfikacja.	2
Wy6	Metody badania własności modelu. Metody wyszukiwania najlepszych parametrów. Wrażliwość modelu.	2
Wy7	Kolokwium zaliczeniowe	1
Wy8	Wybór modeli uczenia maszynowego - ewaluacja, testy statystyczne. Interpretowalność i rozliczalność procesów odkrywania wiedzy. Komunikowanie i wizualizacja	2
	Suma godzin	15

Forma zajęć - ćwiczenia	Liczba godzin
-------------------------	---------------

Ćw1		
Ćw2		
Ćw3		
Ćw4		
..		
	Suma godzin	

Forma zajęć - laboratorium		Liczba godzin
La1	Zapoznanie z programem laboratorium, sposobami oceny, szkolenie BHP.	1
La2	Podstawowe biblioteki do odkrywania wiedzy w języku Python (Numpy, Pandas, Matplotlib, Scikit-learn, Statsmodels)	2
La3	Analiza statystyczna zbioru danych. Zastosowanie prostych metod uczenia maszynowego w analizie danych.	2
La4	Ekstrakcja, transformacja i selekcja cech w oparciu o różnorodne zbiory danych. Dogenerowywanie danych (augmentacja). Ewaluacja modeli uczenia maszynowego. Analiza eksploracyjna danych.	2
La5	Narzędzia zarządzania eksperymentami. Wersjonowanie, przeprowadzanie oraz śledzenie eksperymentów.	2
La6	Produktyzacja procesów odkrywania wiedzy. Udostępnianie metod uczenia maszynowego w środowisku produkcyjnym. Metody monitorowania rozwiązań w środowisku produkcyjnym.	2
La7-8	Rozwiązanie realnego problemu przetwarzania danych i odkrywania wiedzy	4
	Suma godzin	15

Forma zajęć - projekt		Liczba godzin
Pr1		
Pr2		
Pr3		
Pr4		
...		
	Suma godzin	

Forma zajęć - seminarium		Liczba godzin
Se1		
Se2		
Se3		
...		
	Suma godzin	

STOSOWANE NARZĘDZIA DYDAKTYCZNE
N1. Wykład - prezentacja multimedialna
N2. Środowisko deweloperskie Python
N3. Biblioteki deweloperskie dla języka Python

OCENA OSIĄGNIĘCIA PRZEDMIOTOWYCH EFEKTÓW UCZENIA SIĘ

Oceny (F – formująca (w trakcie semestru), P – podsumowująca (na koniec semestru))	Numer efektu uczenia się	Sposób oceny osiągnięcia efektu uczenia się
F1	KSI_W03 KSI_W10	Kolokwium z pytaniami otwartymi i zamkniętymi. Do każdego pytania przypisana jest liczba punktów. Suma punktów jest przeliczana na końcową ocenę w następujący sposób: <50%, 60%) □ dst <60%, 70%) □ dst+ <70%, 80%) □ db <80%, 90%) □ db+ <90%, □ bdb
F2	KSI_U03 KSI_U04 KSI_K01	Kontrola zrealizowania zadań studentów. Każde zadanie ma ustaloną liczbę punktów. Suma punktów przeliczana na ocenę w następujący sposób: <50%, 60%) □ dst <60%, 70%) □ dst+ <70%, 80%) □ db <80%, 90%) □ db+ <90%, □ bdb
P - Ocena z wykładu będzie wystawiana na podstawie oceny z kolokwium - Ocena końcowa z laboratorium będzie wystawiana na podstawie uzyskanych ocen z realizacji poszczególnych zadań.		

LITERATURA PODSTAWOWA I UZUPEŁNIAJĄCA
<p><u>LITERATURA PODSTAWOWA:</u></p> <p>[1] Provost F., Fawcett T., “Data Science for Business: What You Need to Know about Data Mining and Data-Analytic Thinking”, O'Reilly Media, 2013.</p> <p>[2] VanderPlas J., “Python Data Science Handbook: Essential Tools for Working with Data”, O'Reilly Media, 2016.</p> <p><u>LITERATURA UZUPEŁNIAJĄCA:</u></p> <p>[1] William McKinney, “Python for Data Analysis”, O'Reilly Media, 2012.</p> <p>[2] Emmanuel Ameisen, “Building Machine Learning Powered Applications - Going from Idea to Product”, O'Reilly Media, 2020.</p>
OPIEKUN PRZEDMIOTU (IMIE, NAZWISKO, ADRES E-MAIL)

Tomasz Kajdanowicz, tomasz.kajdanowicz@pwr.edu.pl