Politechnika Krakowska im. Tadeusza Kościuszki

Karta przedmiotu

obowiązuje studentów rozpoczynających studia w roku akademickim 2019/2020

Wydział Mechaniczny

Kierunek studiów: Informatyka Stosowana

Profil: Ogólnoakademicki

Forma sudiów: stacjonarne

Kod kierunku: S

Stopień studiów: II

Specjalności: Bez specjalności

1 Informacje o przedmiocie

Nazwa przedmiotu	Uczenie maszynowe i big data
Nazwa przedmiotu w języku angielskim	Machine learning and big data
Kod przedmiotu	WM INFST oIIS B6 19/20
Kategoria przedmiotu	Przedmioty kierunkowe
Liczba punktów ECTS	3.00
Semestry	2

2 Rodzaj zajęć, liczba godzin w planie studiów

Semestr	Wykład	Ćwiczenia	Laboratorium	Laboratorium komputerowe	Projekt	Seminarium
2	15	0	0	15	15	0

3 Cele przedmiotu

Cel 1 Przekazanie wiedzy i umiejętności z zakresu uczenia maszynowego i analiz danych typu BigData

4 Wymagania wstępne w zakresie wiedzy, umiejętności i innych kompetencji

5 Efekty kształcenia

EK1 Wiedza Student zna i rozumie składnię i semantykę języka R.

EK2 Wiedza Student zna i rozumie metody uczenia nienadzorowanego.

EK3 Wiedza Student zna i rozumie metody składowania danych w chmurze i rozproszonego analizowania danych.

EK4 Umiejętności Student potrafi zastosować poznane metody uczenia maszynowego i analiz BigData.

6 Treści programowe

Wykład		
Lp	Tematyka zajęć Opis szczegółowy bloków tematycznych	Liczba godzin
W1	Wprowadzenie do uczenia maszynowego. Elementy języka R. Klasyfikacja binarna. Klasyfikatory bayesowskie. Typy predefiniowane. Modele regresji. Regularyzacja. Optymalizacja. Uczenie nienadzorowane. Analiza składowych głównych. Skalowanie wielowymiarowe. Grupowanie i analiza skupień. Systemy rekomendacyjne. Analiza sieci społecznych. Wprowadzenie do big data. Czyszczenie i integracja danych: dostępne narzędzia i protokoły do integracji danych, problemy z czyszczeniem, przekształcaniem i konsolidacją danych. Hostowane platformy danych i chmura: jaki wpływ na wydajność, skalowalność i modele kosztów mają hostowane platformy danych w chmurze, wewnętrzne i zewnętrzne platformy do przechowywania danych. Przechowywanie dużych danych: rozwiązania do zarządzania danymi ankietowymi na dzisiejszym rynku, w tym tradycyjne RDBMS, NoSQL, NewSQL i Hadoop. Rozproszone platformy obliczeniowe: równoległe systemy obliczeniowe umożliwiające rozproszone przetwarzanie danych w klastrach, w tym MapReduce, Dryad, Spark. Analizy BigData. Redukcja wymiarowości.	15

Laboratorium komputerowe		
Lp	Tematyka zajęć Opis szczegółowy bloków tematycznych	Liczba godzin
K1	Modelowanie nieparametryczne. Redukcja wymiarowości. Analiza skupień. Gromadzenie dużych danych. Przechowywanie dużych danych.	15

Projekt		
Lp	Tematyka zajęć Opis szczegółowy bloków tematycznych	Liczba godzin
P1	Konfiguracja środowiska pracy. Realizacja projektu uczenia maszynowego. Realizacja projektu rozproszonej analizy danych. Konsultacje specjalistyczne.	15

7 Narzędzia dydaktyczne

N1 Wykłady

N2 Ćwiczenia laboratoryjne

N3 Ćwiczenia projektowe

8 Obciążenie pracą studenta

Forma aktywności	Średnia liczba godzin na zrealizowanie aktywności
Godziny kontaktowe z nauczycielem akademickim, w tym:	
Godziny wynikające z planu studiów	45
Konsultacje przedmiotowe	4
Egzaminy i zaliczenia w sesji	0
Godziny bez udziału nauczyciela akademickiego wynikające z nakładu pracy studenta, w tym:	
Przygotowanie się do zajęć, w tym studiowanie zalecanej literatury	21
Opracowanie wyników	0
Przygotowanie raportu, projektu, prezentacji, dyskusji	20
Sumaryczna liczba godzin dla przedmiotu wynikająca z całego nakładu pracy studenta	**90**
Sumaryczna liczba punktów ECTS dla przedmiotu	3.00

9 Sposoby oceny

Ocena formująca

F1 Test z wykładu

F2 Ćwiczenie praktyczne

F3 Projekt indywidualny

Ocena podsumowująca

P1 Średnia ważona ocen formujących

Warunki zaliczenia przedmiotu

W1 Pozytywna ocena z wykładu

W2 Pozytywne oceny z laboratoriów

W3 Pozytywne oceny z projektów

W4 Obecność studenta na min. 75% zajęć laboratoryjnych

Kryteria oceny

Efekt kształcenia 1	
Na ocenę 3.0	Student zna i rozumie w podstawowym zakresie składnię i semantykę języka R.
Efekt kształcenia 2	
Na ocenę 3.0	Student zna i rozumie w podstawowym zakresie metody uczenia nienadzorowanego.
Efekt kształcenia 3	
Na ocenę 3.0	Student zna i rozumie w podstawowym zakresie metody składowania danych w chmurze i rozproszonego analizowania danych.
Efekt kształcenia 4	
Na ocenę 3.0	Student potrafi zastosować w podstawowym zakresie poznane metody uczenia maszynowego i analiz BigData.

10 Macierz realizacji przedmiotu

Efekt kształcenia	Odniesienie danego efektu do szczegółowych efektów zdefiniowanych dla programu	Cele przedmiotu	Treści programowe	Narzędzia dydaktyczne	Sposoby oceny
EK1	K2_W17 K2_W18 K2_W20 K2_U21 K2_U22 K2_U24 M2_U08	Cel 1	W1 K1 P1	N1 N2 N3	F1 F2 F3 P1
EK2	K2_W17 K2_W18 K2_W20 K2_U21 K2_U22 K2_U24 M2_U08	Cel 1	W1 K1 P1	N1 N2 N3	F1 F2 F3 P1

Efekt kształcenia	Odniesienie danego efektu do szczegółowych efektów zdefiniowanych dla programu	Cele przedmiotu	Treści programowe	Narzędzia dydaktyczne	Sposoby oceny
EK3	K2_W17 K2_W18 K2_W20 K2_U21 K2_U22 K2_U24 M2_U08	Cel 1	W1 K1 P1	N1 N2 N3	F1 F2 F3 P1
EK4	K2_W17 K2_W18 K2_W20 K2_U21 K2_U22 K2_U24 M2_U08	Cel 1	W1 K1 P1	N1 N2 N3	F1 F2 F3 P1

11 Wykaz literatury

12 Informacje o nauczycielach akademickich

Osoba odpowiedzialna za kartę

dr hab. inż., prof. PK Jacek Pietraszek (kontakt: jacek.pietraszek@pk.edu.pl)

Osoby prowadzące przedmiot

1 pracownicy Instytutu Informatyki Stosowanej (kontakt:)

13 Zatwierdzenie karty przedmiotu do realizacji

(miejscowość, data) (odpowiedzialny za przedmiot) (dziekan)

Przyjmuję do realizacji (data i podpisy osób prowadzących przedmiot)

. . . . . .