

OPIS PRZEDMIOTU KSZTAŁCENIA

Nazwa przedmiotu Eksploracja danych					
Nazwa jednostki prowadzącej przedmiot Instytut Matematyki					
kierunek	specjalność	specjalizacja	semestr/y	poziom kształcenia/ profil kształcenia	forma studiów
Informatyka	Programowanie	-	5	SPS/ praktyczny	Stacjonarne/niestacjonarne
Nazwisko osoby prowadzącej (osób prowadzących) dr Ryszard Motyka					
Formy zajęć	Liczba godzin				Liczba punktów ECTS
	N (nauczyciel)		S (student)		
	studia stacjonarne	studia niestacjonarne	studia stacjonarne	studia niestacjonarne	
(CL) ćwiczenia laboratoryjne	30	18	45	57	3
Przygotowanie do zajęć (rozwiązywanie zadań)			15	27	
Poszukiwanie dodatkowych materiałów z różnych źródeł			10	10	
Przygotowanie do kolokwium			20	20	
Razem	30	18	45	57	3
Metody dydaktyczne <ul style="list-style-type: none"> (CL) ćwiczenia laboratoryjne: ćwiczenia praktyczne - rozwiązywanie zadań, metoda problemowa, praca w grupach 					
Określenie przedmiotów wprowadzających wraz z wymogami wstępnymi <p>A. Wymagania formalne: Wstęp do gromadzenia i przetwarzania danych</p> <p>B. Wymagania wstępne: wiadomości z podstaw gromadzenia i przetwarzania danych, umiejętności korzystania ze środowisk obliczeniowych</p>					
Cele przedmiotu <ul style="list-style-type: none"> Zapoznanie z podstawowymi pojęciami, metodami i algorytmami eksploracji danych Przedstawienie głównych kierunków rozwoju oraz metod eksploracji danych, jak również możliwości zastosowań w świecie rzeczywistym Poznanie podstawowych procedur pracy z dużymi zbiorami danych Nabycie umiejętności wyboru właściwych technik eksploracji danych Nabycie umiejętności obsługi wybranego programu do analizy danych do celów związanych z eksploracją danych (np. RapidMiner, Orange, Statistica, MS Excel, R) Nabycie umiejętności dokonania analizy eksploracyjnej na wybranych zbiorach danych 					
<p>Pojęcie eksploracji danych i data mining. Wstępne przetwarzanie danych (czyszczenie danych, obsługa brakujących danych, graficzne metody identyfikacji punktów oddalonych, przekształcanie danych, normalizacja, standaryzacja).</p> <p>Eksploracyjna analiza danych (EDA). Klasyfikacja probabilistyczna - metoda naiwna bayesowska Klasyfikacja metodą kNN. Drzewa klasyfikacyjne i regresyjne (bagging, losowy las, drzewa ze wzmacnianiem).</p> <p>Wybrane metody grupowania. Analiza skupień uogólnioną metodą k-średnich (skale pomiarowe, miary odległości i podobieństwa obiektów, algorytmy segmentacji) .</p> <p>Analiza koszykowa (cel analizy koszykowej, postać danych, reguły asocjacyjne, miary jakości reguł)</p>					

<p>Efekty kształcenia</p> <p>Wiedza W_01 Zna podstawowe metody eksploracyjnej analizy danych z zakresu data mining. W_02 Zna zakres możliwości obliczeniowych wybranego środowiska komputerowego do przeprowadzania eksploracyjnej analizy danych. W_03 Zna co najmniej jedno oprogramowanie służące do przeprowadzania analizy danych.</p> <p>Umiejętności U_01 Umie wybrać i zastosować właściwe metody data mining i statystyki do przeprowadzenia eksploracyjnej analizy danych. U_02 Umie przeprowadzić eksploracyjną analizę danych. U_03 Umie posługiwać się wybranym oprogramowaniem służącym do analizy danych. U_04 Umie napisać raport z rozwiązania problemu posługując się poprawnym i zrozumiałym językiem matematycznym.</p> <p>Kompetencje społeczne K_01 Wykazuje kreatywność przy rozwiązywaniu problemów</p>	<p>Sposób zaliczenia oraz formy i podstawowe kryteria oceny/wymagania egzaminacyjne</p> <p>A. Sposób zaliczenia CL – zaliczenie z oceną</p> <p>B. Sposoby weryfikacji i oceny efektów (CL) Ćwiczenia laboratoryjne - kolokwia pisemne – pytania otwarte - efekty: W_01, W_02, W_03, U_01, U_02, U_03, U_04K_01</p> <p>Maksymalna liczba punktów to a. Ocena K z kolokwium jest wyliczona według zasady: K ∈ [0% a, 50% a) niedostateczna K ∈ [50% a, 60% a) dostateczna K ∈ [60% a, 70% a) dostateczna plus K ∈ [70% a, 80% a) dobra K ∈ [80% a, 90% a) db plus K ∈ [90% a, 100% a] bardzo dobra</p> <p>Ocena zaliczenia ćwiczeń jest obliczona jako średnia arytmetyczna ocen z kolokwiów pisemnych.</p>
---	--

Matryca efektów kształcenia dla przedmiotu

Numer (symbol) efektu kształcenia	Odniesienie do efektów kształcenia dla programu	Odniesienie do charakterystyki drugiego stopnia PRK dla obszaru/ obszarów
W_01	K1_W17	P6S_WG
W_02	K1_W17	P6S_WG
W_03	K1_W08	P6S_WG
U_01	K1_U33	P6S_UW
U_02	K1_U33	P6S_UW
U_03	K1_U33	P6S_UW
U_04	K1_U39	P6S_UW
K_01	K1_K02	P6S_KO

Wykaz literatury

A. Literatura wymagana do ostatecznego zaliczenia zajęć (zdania egzaminu):

1. Larose T. D., *Odkrywanie wiedzy z danych. Wprowadzenie do eksploracji danych*, PWN, Warszawa 2004
2. Koronacki J., Ćwik J.: *Statystyczne systemy uczące się*. WNT, Warszawa 2005

A. Literatura uzupełniająca

1. *Zastosowania statystyki i data mining w badaniach naukowych - materiały z seminarium StatSoft Polska*, 2010
2. *Zastosowania statystyki i data mining w badaniach naukowych - materiały z seminarium StatSoft Polska*, 2007
3. *Praktyczne wykorzystanie analizy danych i data mining - materiały z seminarium StatSoft Polska*, 2008
4. Artykuły z poświęcone tematyce Data Mining znajdujące się na stronie www.statsoft.pl w dziale Czytelnia

Kontakt

dr Ryszard Motyka
ryszard.motyka@apsl.edu.pl

