

Eksploracja danych - opis przedmiotu

Informacje ogólne	
Nazwa przedmiotu	Eksploracja danych
Kod przedmiotu	04.2-WE-BEP-ED
Wydział	Wydział Informatyki, Elektrotechniki i Automatyki
Kierunek	Biznes elektroniczny
Profil	praktyczny
Rodzaj studiów	pierwszego stopnia z tyt. inżyniera
Semestr rozpoczęcia	semestr zimowy 2018/2019

Informacje o przedmiocie	
Semestr	2
Liczba punktów ECTS do zdobycia	4
Typ przedmiotu	obowiązkowy
Język nauczania	polski
Sylabus opracował	dr hab. inż. Marek Kowal, prof. UZ

Formy zajęć					
Forma zajęć	Liczba godzin w semestrze (stacjonarne)	Liczba godzin w tygodniu (stacjonarne)	Liczba godzin w semestrze (niestacjonarne)	Liczba godzin w tygodniu (niestacjonarne)	Forma zaliczenia
Wykład	15	1	9	0,6	Zaliczenie na ocenę
Laboratorium	30	2	18	1,2	Zaliczenie na ocenę

Cel przedmiotu

Prezentacja oprogramowania wykorzystywanego do eksploracji danych. Zapoznanie słuchaczy z metodami czyszczenia danych. Prezentacja metod klasyfikacji danych. Prezentacja metod odkrywania asocjacji i sekwencji. Prezentacja metod klasteryzacji danych. Wykształcenie praktycznych umiejętności obsługi wybranych systemów eksploracji danych. Wykształcenie umiejętności w zakresie stosowania metod eksploracji danych w e-biznesie (segmentacji klientów, ocenie ryzyka kredytowego, strategii cross-sellingu, wykrywaniu nadużyć).

Wymagania wstępne

Zakres tematyczny

Przegląd i charakterystyka oprogramowania wykorzystywanego do eksploracji danych. Zapoznanie z oprogramowaniem do eksploracji danych (SAS). Struktury danych wykorzystywane w eksploracji danych. Typy i role zmiennych w zadaniach eksploracji danych.

Przygotowanie danych do eksploracji. Profilowanie danych. Czyszczenie danych. Próbkowanie danych. Transformacja zmiennych.

Klasyfikacji danych. Klasyfikacja danych za pomocą drzew klasyfikacyjnych, algorytmu k najbliższych sąsiadów, naiwnego klasyfikatora Bayesa, sieci neuronowych, regresji logistycznej. Miary jakości klasyfikacji. Ćwiczenia praktyczne z wykorzystania poznanych metod klasyfikacji danych.

Odkrywanie wzorców asocjacji i sekwencji. Miary opisujące statystyczną ważności i siłę reguł asocjacji i wzorców sekwencji. Problem analizy koszyka zakupów. Złożoność obliczeniowa problemu odkrywania reguł asocjacji. Omówienie algorytmu Apriori i algorytmu Generalized Sequential Pattern. Ćwiczenia praktyczne z wykorzystaniem poznanych metod odkrywania asocjacji sekwencji.

Definicja zagadnienia klasteryzacji danych. Metody grupowania hierarchicznego. Metody grupowania iteracyjno-optymalizacyjnego. Miary odległości stosowane w klasteryzacji. Wyznaczanie charakterystyk klastrow. Metody szacowania liczby klastrow. Ćwiczenia praktyczne z wykorzystania poznanych metod klasteryzacji danych.

Metody kształcenia

Wykład - wykład konwencjonalny z wykorzystaniem wideoprojektora.
Laboratorium - ćwiczenia praktyczne w laboratorium komputerowym.

Efekty uczenia się i metody weryfikacji osiągnięcia efektów uczenia się

Opis efektu	Symbole efektów	Metody weryfikacji	Forma zajęć
Zna technologie i narzędzia informatyczne stosowane w eksploracji danych	- K_W04 - K_W07	test z pytaniami zamkniętymi i otwartymi	Wykład

Opis efektu	Symbole efektów	Metody weryfikacji	Forma zajęć
Zna typowe zastosowania metod eksploracji danych w e-biznesie	K_W06 K_W18	test z pytaniami zamkniętymi i otwartymi	Wykład
Potrafi posługiwać się algorytmami odkrywania asocjacji i sekwencji	K_U02 K_U16 K_U18 K_U20	- obserwacja i ocena aktywności na zajęciach - obserwacje i ocena umiejętności praktycznych studenta - sprawdzian z progami punktowymi	Laboratorium
Potrafi przygotować dane do eksploracji wykonując ich czyszczenie	K_U16 K_U20	- obserwacja i ocena aktywności na zajęciach - obserwacje i ocena umiejętności praktycznych studenta - sprawdzian z progami punktowymi	Laboratorium
Zna metody eksploracji danych	K_W02 K_W17	test z pytaniami zamkniętymi i otwartymi	Wykład
Potrafi dobrać metodę klasteryzacji danych, przeprowadzić proces grupowania danych i wyznaczyć charakterystyki klastrów	K_U02 K_U16 K_U18 K_U20	- obserwacja i ocena aktywności na zajęciach - obserwacje i ocena umiejętności praktycznych studenta - sprawdzian z progami punktowymi	Laboratorium
Potrafi dobrać metodę klasyfikacji danych, zbudować klasyfikator i przeprowadzić testy dokładności klasyfikacji	K_U02 K_U16 K_U18 K_U20	- obserwacja i ocena aktywności na zajęciach - obserwacje i ocena umiejętności praktycznych studenta - sprawdzian z progami punktowymi	Laboratorium
Potrafi zastosować metody eksploracji danych w innowacyjnych rozwiązaniach dla e-biznesu	K_K03 K_K04 K_K08 K_K09	- aktywność w trakcie zajęć - wykonanie sprawozdań laboratoryjnych	Laboratorium
Rozumie potrzebę ciągłego uzupełniania i poszerzania wiedzy w zakresie intensywnie rozwijających się metod eksploracji danych	K_K01 K_K11	- aktywność w trakcie zajęć - wykonanie sprawozdań laboratoryjnych	Laboratorium

Warunki zaliczenia

Wykład - warunkiem zaliczenia jest uzyskanie pozytywnych ocen z testów przeprowadzonych, co najmniej raz w semestrze.

Laboratorium - warunkiem zaliczenia jest uzyskanie pozytywnych ocen ze wszystkich ćwiczeń laboratoryjnych oraz kolokwiów i sprawdzianów przewidzianych w ramach laboratorium.

Składowe oceny końcowej = wykład: 50% + laboratorium: 50%

Literatura podstawowa

1. Morzy T.: Eksploracja danych. Metody i algorytmy., Warszawa, PWN, 2013.
2. Hand, D., Mannila, H., Smyth, P., Eksploracja danych, Warszawa, WNT, 2005.
3. Larose, D. T., Odkrywanie wiedzy z danych, Warszawa, PWN, 2006.
4. Larose, D. T., Metody i modele eksploracji danych, Warszawa, PWN, 2008.
5. Kwiatkowska, A. M., Systemy wspomagania decyzji: jak korzystać z wiedzy i informacji w praktyce, Warszawa, PWN, 2007.

Literatura uzupełniająca

1. Aggarwal C.C.: Data Mining, Springer, 2015.
2. Alexander M., Decker J., Wehbe B.: Zaawansowane wykorzystanie Excela. Analizy Business Intelligence, Helion, 2015.
3. Słapór, K., Automatyczna klasyfikacja obiektów, EXIT, 2005.
4. Han, J., Kamber, M., Data Mining: Concepts and Techniques, Morgan Kaufmann, 2011.

Uwagi

Zmodyfikowane przez dr hab. inż. Marek Kowal, prof. UZ (ostatnia modyfikacja: 24-04-2018 11:55)