

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 2020/2021

Informacje o przedmiocie	
Semestr	2
Liczba punktów ECTS do zdobycia	3
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 skale zmiennych w zadaniach eksploracji danych.

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

Klasyfikacji danych. Klasyfikacja danych za pomocą drzew klasyfikacyjnych, 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
Potrafi wymienić i opisać technologie i narzędzia informatyczne stosowane do implementacji zadań z obszaru eksploracji danych	K_W04	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_W20	• 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 • wykonanie sprawozdań laboratoryjnych	• Laboratorium
Potrafi przygotować dane do eksploracji danych wykonując ich czyszczenie oraz selekcję zmiennych	• K_U16 • K_U20	• obserwacja i ocena aktywności na zajęciach • wykonanie sprawozdań laboratoryjnych	• Laboratorium
Rozumie jak funkcjonują metody klasyfikacji danych, klasteryzacji danych, odkrywania reguł asocjacji i sekwencji	• K_W02 • K_W17	• test z pytaniami zamkniętymi i otwartymi	• Wykład
Potrafi posługiwać się metodami klasteryzacji danych do eksploracji danych	• K_U02 • K_U16 • K_U18 • K_U20	• obserwacja i ocena aktywności na zajęciach • wykonanie sprawozdań laboratoryjnych	• Laboratorium
Potrafi posługiwać się drzewami decyzyjnymi, regresją logistyczną oraz sieciami neuronowymi na potrzeby klasyfikacji danych	• K_U02 • K_U16 • K_U18 • K_U20	• obserwacja i ocena aktywności na zajęciach • wykonanie sprawozdań laboratoryjnych	• Laboratorium
Potrafi opracować scenariusz wykorzystania wybranej metody eksploracji danych na potrzeby wybranego e-biznesu	• K_K03 • K_K04 • K_K08 • K_K09	• referat	• Laboratorium
Potrafi samodzielnie odszukać materiały źródłowe opisujące nowatorskie wykorzystanie metod eksploracji danych w zastosowaniach w e-biznesie i przygotować krótkie podsumowanie w formie referatu	• K_K01 • K_K11	• referat	• 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. Foreman J.W., Mistrz analizy danych Od danych do wiedzy, Helion, 2017
2. Lasek M., Pęczkowski M., Enterprise Miner. Wykorzystywanie narzędzi Data Mining w systemie SAS, Wydawnictwa Uniwersytetu Warszawskiego, 2017.
3. Morzy T.: Eksploracja danych. Metody i algorytmy., Warszawa, PWN, 2013.
4. Osowski S., Metody i narzędzia eksploracji danych, BTC, 2013.
5. Provost F., Fawcett T., Analiza danych w biznesie. Sztuka podejmowania skutecznych decyzji, Onepress, 2014.
6. Larose D. T., Metody i modele eksploracji danych, Warszawa, PWN, 2008.
7. Kwiatkowska A. M., Systemy wspomagania decyzji: jak korzystać z wiedzy i informacji w praktyce, Warszawa, PWN, 2007.
8. Hand D., Mannila H., Smyth P., Eksploracja danych, Warszawa, WNT, 2005.
9. Larose D. T., Odkrywanie wiedzy z danych, Warszawa, PWN, 2006.

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. Stąpor K., Automatyczna klasyfikacja obiektów, EXIT, 2005.
4. Han J., Kamber, M., Data Mining: Concepts and Techniques, Morgan Kaufmann, 2011.
5. Hand D., Mannila H., Smyth P., Eksploracja danych, Warszawa, WNT, 2005.
6. Larose D. T., Odkrywanie wiedzy z danych, Warszawa, PWN, 2006.

Uwagi

