

Wprowadzenie do hurtowni danych i baz wiedzy - opis przedmiotu

Informacje ogólne	
Nazwa przedmiotu	Wprowadzenie do hurtowni danych i baz wiedzy
Kod przedmiotu	11.3-WI-INFP-WHDiBW
Wydział	Wydział Informatyki, Elektrotechniki i Automatyki
Kierunek	Informatyka
Profil	ogólnoakademicki
Rodzaj studiów	pierwszego stopnia z tyt. inżyniera
Semestr rozpoczęcia	semestr zimowy 2020/2021

Informacje o przedmiocie	
Semestr	6
Liczba punktów ECTS do zdobycia	5
Typ przedmiotu	obieralny
Język nauczania	polski
Sylabus opracował	- dr inż. Jacek Tkacz - dr hab. inż. Artur Gramacki, prof. UZ - dr hab. inż. Wiesław Miczulski, 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
Projekt	15	1	9	0,6	Zaliczenie na ocenę

Cel przedmiotu

1. Zapoznanie studenta z zasadami działania i możliwościami współczesnych systemów hurtowni danych oraz systemów eksploracji danych.
2. Ukształtowanie umiejętności w zakresie wstępnego przygotowywania danych oraz zasilania w dane systemów hurtowni danych i systemów eksploracji danych.
3. Ukształtowanie umiejętności w zakresie efektywnego dostępu do danych w kontekście hurtowni danych i systemów eksploracji danych.
4. Ukształtowanie umiejętności w zakresie przeprowadzania typowych zadań eksploracyjnych na posiadanych danych.

Wymagania wstępne

Bazy danych, Algorytmy i struktury danych

Zakres tematyczny

Hurtownie danych (ang. *data warehouse*). Architektura hurtowni danych (star, snow flake). Zasilanie hurtowni danych (procesy ETL) oraz odświeżanie hurtowni danych. Modelowanie (struktury danych typu ROLAP, MOLAP) Wsparcie dla hurtowni ze strony języka SQL - funkcje analityczne. Indeksowanie oraz optymalizacja zapytań. Indeksy kolumnowe, perspektywy zmaterializowane.

Eksploracja danych (ang. *data mining*). Omówienie podstawowych zagadnień związanych z dziedziną eksploracji danych. Związek eksploracja danych z hurtowniami danych oraz bazami o charakterze transakcyjnym. Wstępne przetwarzanie danych (czyszczenie, integracja, transformacja, redukcja liczebności i wymiarowości, określanie ważności atrybutów). Budowanie modeli do celów szacowania (estymacji) i przewidywania (predykcji). Regresja. Odkrywanie osobiowości. Klasyfikacja danych. Grupowanie danych. Reguły asocjacyjne (analiza koszyka sklepowego, zbiory częste). Odkrywanie wzorców sekwencji. Ważność atrybutów. Wykrywanie anomalii. Wydobywanie z danych cech istotnych. Eksploracja danych potokowych i szeregów czasowych. Eksploracja struktur grafowych oraz sieci Web. Eksploracja tekstu. Wybrane środowiska informatyczne związane z tworzeniem i obsługą hurtowni danych oraz systemów eksploracji danych.

Metody kształcenia

wykład: wykład konwencjonalny

laboratorium: ćwiczenia laboratoryjne

projekt: praca w grupach, metoda projektu

Efekty uczenia się i metody weryfikacji osiągnięcia efektów uczenia się

Opis efektu	Symbole efektów	Metody weryfikacji	Forma zajęć
Eksploruje dane tekstowe i sieć Web.	K_U25	- bieżąca kontrola na zajęciach - sprawdzian	Laboratorium

Opis efektu	Symbol e efektów	Metody weryfikacji	Forma zajęć
Student wymienia i tłumaczy podstawowe pojęcia związane z hurtowniami danych i eksploracją danych.	• K_W12	• kolokwium • test końcowy	• Wykład
Potrafi pobrać (lub uaktualnić) dane do systemów hurtowni danych oraz systemów eksploracji danych z różnych źródeł zewnętrznych.	• K_U25	• przygotowanie projektu	• Projekt
Indeksuje oraz optymalizuje zapytania analityczne.	• K_U25	• bieżąca kontrola na zajęciach • sprawdzian	• Laboratorium
Potrafi zrealizować proces ETL	• K_U25	• bieżąca kontrola na zajęciach • sprawdzian	• Laboratorium
Potrafi zdefiniować oraz przeprowadzić typowe zadania eksploracyjne na danych.	• K_U25	• przygotowanie projektu	• Projekt
Tworzy elementy hurtowni danych	• K_U25	• bieżąca kontrola na zajęciach • sprawdzian	• Laboratorium

Warunki zaliczenia

Wykład - warunkiem zaliczenia jest uzyskanie pozytywnych ocen z kolokwiów pisemnych lub ustnych przeprowadzonych co najmniej raz w semestrze.

Laboratorium - warunkiem zaliczenia jest uzyskanie pozytywnych ocen ze wszystkich ćwiczeń realizowanych w semestrze.

Projekt - warunkiem zaliczenia jest uzyskanie pozytywnych ocen ze wszystkich projektów realizowanych w semestrze.

Składowe oceny końcowej = wykład: 30% + laboratorium: 30% + projekt: 40%

Literatura podstawowa

- Larose D.: Odkrywanie wiedzy z danych, Wydawnictwo Naukowe PWN, Karwasza, 2006.
- Larose D.: Metody I modele eksploracji danych, Wydawnictwo Naukowe PWN, Warszawa, 2007
- Larose D., Markov Z.: Eksploracja zasobów internetowych, Analiza struktury, zawartości i użytkowania sieci WWW, Wydawnictwo Naukowe PWN, Warszawa, 2010
- Hand D., Mannila H., Smyth P.: *Principles of Data Mining*. Massachusetts Institute of Technology, 2001.
- Jarke M., Lenzerini M., Vassiliou Y., Vassiliadis P.: *Fundamentals of Data Warehouses*. Springer-Verlag, Berlin, 2002.
- Rutkowski L.: Computational Intelligence. Methods and Techniques. Springer-Verlag, Berlin, 2008.

Literatura uzupełniająca

- Cichosz P.: Systemy uczące się, WNT, 2000.
- Beynon-Davies P.: Systemy baz danych, Wydawnictwa Naukowo-Techniczne, 2005.
- Todman Ch.: Projektowanie hurtowni danych. Zarządzanie kontaktami z klientami (CRM), Wydawnictwa Naukowo-Techniczne, 2005.
- Data Mining: Concepts and Techniques, J. Han, M. Kamber, Morgan Kaufman, 2000.
- Data Mining: Practical Machine Learning Tools and Techniques with Java Implementations, I. H. Witten, E. Frank, Morgan Kaufman, 2000.
- Principles of Data Mining, J. Hand, H. Mannila, P. Smyth, MIT Press, 2001.

Uwagi

Zmodyfikowane przez dr inż. Jacek Tkacz (ostatnia modyfikacja: 22-04-2020 16:46)

Wygenerowano automatycznie z systemu SylabUZ