

## Opis przedmiotu

**Nazwa przedmiotu: Hurtownie danych**

**2. Kod przedmiotu: 11.3 SSK.HDN.05**

**3. Język wykładowy:** polski

**4. Kierunek:** Informatyka

**5. Specjalność:** Systemy i sieci komputerowe

**6. Rok:** 3 **Semestr:** 5

**7. Tytuł/stopień oraz imię i nazwisko prowadzącego przedmiot:**

Prof. dr hab. inż. Ngoc Thanh Nguyen

**8. Tytuły/stopnie oraz imiona i nazwiska pozostałych członków zespołu:**

mgr inż. Michał Malski

**9. Formy zajęć wchodzące w skład przedmiotu, wymiar godzinowy, forma zaliczenia:**

Forma zajęć	Wykład	Ćwiczenia/ Ćwiczenia tablicowe	Laboratorium/ Ćwiczenia praktyczne	Projekt	Seminarium
Liczba godzin w semestrze	30			15	
Forma zaliczenia	Egzamin			Zaliczenie	

**10. Liczba punktów ECTS:** 4

**11. Poziom :** zaawansowany

**12. Wymagania wstępne:**

Zaliczenie kursów: Podstawy baz danych, Nowoczesne systemy baz danych

**13. Cele kształcenia:**

Celem jest zapoznanie z problemami, metodami i rozwiązaniami informatycznych systemów hurtowni i eksploracji danych. Tematyka obejmuje m.in.: hurtownię danych jako podstawę systemów wspomaganie decyzji, technologię MOLAP i ROLAP, narzędzia dostępu do danych, podstawy projektowania hurtowni danych, wybrane metody i techniki eksploracji danych, architekturę systemów eksploracji danych.

**14. Opis treści kształcenia w ramach poszczególnych form zajęć:**

**14.1. Wykład:**

Pojęcia podstawowe. Hurtownie danych a OLTP. Powody stosowania hurtowni danych. Ich cechy charakterystyczne. Proces budowania hurtowni. 2- i 3-warstwowe hurtownie danych. Hurtownie a targowiska danych. Obszary zastosowań hurtowni. Proces decyzyjny.

Architektury i modele hurtowni danych. Tworzenie i utrzymywanie hurtowni. Problem aktualizacji i integracji. Model relacyjny a model wielowymiarowy. Model płatka gwiazdy, płatka śniegu i konstelacji faktów. Kostka wielowymiarowa. MOLAP, ROLAP, HOLAP. Operacje OLAP. Problemy wydzielenia, transformacji i ładowania. Główne problemy hurtowni danych. Procesy KDD. Selekcja. Oczyszczanie. Wzbogacanie. Kodowanie (transformacja). Raportowanie Eksploracja danych. Metody i techniki eksploracji. Obszary zastosowań. Analiza połączeń. Reguły związków. Algorytm a-priori. Wzory sekwencji. Podobne sekwencje czasowe. Grupowanie - segmentacja. Klasyfikacja i predykcja. Algorytm k-means. Algorytmy skupiające HACM. Drzewa decyzyjne. Metody statystyczne. Problemy i kierunki rozwoju. Narzędzia i zastosowania.

#### 14.2. Ćwiczenia/Ćwiczenia tablicowe:

--

#### 15. Literatura podstawowa:

1. Poe V., Klauer P., Brobst S.: Tworzenie hurtowni danych: wspomaganie podejmowania decyzji. WNT, 2000
2. Giovinazzo W.: Object-Oriented Data Warehouse Design: Building a Star Schema. Prentice Hall, 2000
3. Adriaans P., Zantinge D.: Data mining. Addison-Wesley, 1998
4. Groth R.: Data mining: A Hands-On Approach for Business Professionals. Prentice Hall, 1998
5. Sturm J.: Hurtownie danych Microsoft SQL Server 7.0. Przewodnik Techniczny. APN PROMISE, Warszawa 2000

#### 16. Literatura towarzysząca:

1. Barquin R.C., Edelstein H.A. (eds): Planning and Designing the Data Warehouse. Prentice Hall, 1997
2. Bischoff J., Alexander T.: Data Warehouse: Practical Advice from the Experts. Prentice Hall, 1997
3. Fayyad U.M., Piatetsky-Saphiro G., Smyth P., Uthurusamy R.: Advances in Knowledge Discovery and Data Mining. AAAI Press/The MIT Press, 1996
4. Cabena P., i inni: Discovering Data Mining: From Concept to Implementation. Prentice Hall, 1998