



UNIWERSYTET
IM. ADAMA MICKIEWICZA
W POZNANIU

Gromadzenie i eksploracja danych Sylabus zajęć

Informacje podstawowe

Kierunek studiów Analiza i przetwarzanie danych Specjalność - Jednostka organizacyjna Wydział Matematyki i Informatyki Poziom studiów studia drugiego stopnia Forma studiów studia niestacjonarne Profil studiów profil ogólnoakademicki		Cykl dydaktyczny 2024/25 Kod zajęć 06APDN.24K.00332.24 Języki wykładowe polski Obligatoryjność Obowiązkowy Blok zajęciowy Przedmioty kierunkowe
Koordynator zajęć	Anna Stachowiak	
Prowadzący zajęcia	Anna Stachowiak	
Okres Semestr 3	Forma zajęć / liczba godzin / forma zaliczenia - Wykład: 15, Egzamin; w tym zajęcia zdalne: - Wykład synchroniczny: 15 - Laboratorium: 15, Zaliczenie z oceną	Liczba punktów ECTS 6

Cele kształcenia dla zajęć

Kod	Cel
C1	Celem przedmiotu jest przedstawienie rozwiązań technologicznych klasy "business intelligence" oraz systemów analitycznych (OLAP), w szczególności tworzenie hurtowni danych, import i transformacja danych, stosowanie wybranych metod analizy i prezentacji danych oraz narzędzi raportujących.

Wymagania wstępne

Znajomość relacyjnych baz danych i języka SQL.

Umiejętność czytania schematów baz, tworzenia zapytań SQL.

Efekty uczenia się dla zajęć

Kod	Efekty uczenia się dla zajęć w zakresie	Efekty uczenia się dla kierunku	Metody weryfikacji osiągnięcia efektów uczenia się dla zajęć
Wiedzy - Student/ka:			
W1	zna i rozumie zależności między systemami transakcyjnymi a analitycznymi, zna cele tworzenia oraz modele hurtowni danych.	APD_K2_W01, APD_K2_W04	Egzamin pisemny, Test, realizacja zadań częściowych
W2	zna różne modele danych źródłowych; zna metody ich transformacji i integracji; rozumie potrzebę zapewnienia wysokiej jakości danych.	APD_K2_W01, APD_K2_W04	Egzamin pisemny, Test, realizacja zadań częściowych
W3	zna i rozumie wybrane algorytmy zgłębiania danych i ich zastosowania w systemach analitycznych.	APD_K2_W01, APD_K2_W04	Egzamin pisemny, Test, realizacja zadań częściowych
W4	zna metody i narzędzia prezentacji danych.	APD_K2_W01	Egzamin pisemny, Test, realizacja zadań częściowych
Umiejętności - Student/ka:			
U1	umie projektować hurtownie danych i powiązane z nimi systemy analityczne.	APD_K2_U01, APD_K2_U06, APD_K2_U08	realizacja zadań częściowych
U2	potrafi integrować dane pochodzące z różnych źródeł, transformować je do spójnej postaci oraz wypełniać nimi hurtownie danych.	APD_K2_U01, APD_K2_U07	realizacja zadań częściowych
U3	potrafi wykorzystać wybrane modele i algorytmy zgłębiania danych w systemach analitycznych.	APD_K2_U01, APD_K2_U02, APD_K2_U05, APD_K2_U06, APD_K2_U09	realizacja zadań częściowych
U4	potrafi tworzyć analizy i raporty na podstawie różnych źródeł i modeli danych.	APD_K2_U01, APD_K2_U06	realizacja zadań częściowych

Treści programowe dla zajęć

Lp.	Treści programowe dla zajęć	Efekty uczenia się dla zajęć	Formy zajęć
1.	Systemy transakcyjne (OLTP) a analityczne (OLAP); porównanie; obszary zastosowań.	W1	Wykład, Wykład synchroniczny
2.	Hurtownie danych. Modele, fakty, miary i wymiary.	W1, U1	Wykład, Laboratorium, Wykład synchroniczny
3.	Tworzenie hurtowni danych z wykorzystaniem schematów gwiazdy i płotka śniegu.	W1, U1	Wykład, Laboratorium, Wykład synchroniczny
4.	Rozszerzenia języka SQL wspierające rozwiązania analityczne.	W1, U2, U4	Wykład, Laboratorium, Wykład synchroniczny

Lp.	Treści programowe dla zajęć	Efekty uczenia się dla zajęć	Formy zajęć
5.	Język DAX w projektowaniu systemów analitycznych w modelu tabelarycznym.	W1, U1	Wykład, Laboratorium, Wykład synchroniczny
6.	Integracja i transformacja danych. Jakość danych i metody jej polepszania.	W2, W3, U1, U2	Wykład, Laboratorium, Wykład synchroniczny
7.	Pozyskiwanie i integracja danych z różnych źródeł (baz relacyjnych, wielowymiarowych, usług sieciowych i innych).	W2, U2	Wykład, Laboratorium, Wykład synchroniczny
8.	Automatyzacja procesów ETL.	W2, W3, U1, U2, U3	Wykład, Laboratorium, Wykład synchroniczny
9.	Wybrane narzędzia i metody wizualizacji danych.	W4, U4	Wykład, Laboratorium, Wykład synchroniczny
10.	Analiza i prezentacja danych z wykorzystaniem interaktywnych raportów, pulpitów menadżerskich i wybranych narzędzi BI	W4, U4	Wykład, Laboratorium, Wykład synchroniczny
11.	Tworzenie interaktywnych raportów; raporty lokalne; serwery raportów	W4, U4	Wykład, Laboratorium, Wykład synchroniczny
12.	Zastosowania wbranych modeli i algorytmów eksploracji danych w systemach analitycznych.	W3, U3	Wykład, Laboratorium, Wykład synchroniczny
13.	Kierunki rozwoju systemów analitycznych.	W1	Wykład, Wykład synchroniczny

Informacje dodatkowe

Forma zajęć	Metody i formy prowadzenia zajęć
Wykład	Wykład z prezentacją multimedialną wybranych zagadnień, Metoda analizy przypadków
Laboratorium	Metoda analizy przypadków, Rozwiązywanie zadań (np.: obliczeniowych, artystycznych, praktycznych), Metoda laboratoryjna, Metoda projektu

Forma zajęć	Warunki zaliczenia zajęć
Wykład	Warunkiem koniecznym zaliczenia przedmiotu jest wykazanie się umiejętnościami praktycznymi potwierdzonymi pozytywnym zaliczeniem zajęć laboratoryjnych. Zaliczenie wykładu odbywa się na podstawie egzaminu pisemnego w postaci testu. Skala ocen: 1. bardzo dobry (bdb; 5,0) – od 90% punktów, 2. dobry plus (db plus; 4,5) – od 80% punktów, 3. dobry (db; 4,0) – od 70% punktów, 4. dostateczny plus (dst plus; 3,5) – od 60% punktów, 5. dostateczny (dst; 3,0) – od 50% punktów, 6. niedostateczny (ndst; 2,0) – poniżej 50% punktów.
Laboratorium	Ocena z zajęć laboratoryjnych wyznaczana jest na podstawie liczby punktów uzyskiwanych podczas pracy na zajęciach i pracy własnej przy rozwiązywaniu zadań praktycznych (około 10 zadań ocenianych na 2-5 pkt). Skala ocen: 1. bardzo dobry (bdb; 5,0) – od 90% punktów, 2. dobry plus (db plus; 4,5) – od 80% punktów, 3. dobry (db; 4,0) – od 70% punktów, 4. dostateczny plus (dst plus; 3,5) – od 60% punktów, 5. dostateczny (dst; 3,0) – od 50% punktów, 6. niedostateczny (ndst; 2,0) – poniżej 50% punktów.

Literatura

Obowiązkowa

1. R. Kimball , M. Ross, The Data Warehouse Toolkit , Wiley , 2013
2. Russo Marco , Ferrari Alberto, Microsoft SQL Server 2016. Analysis Services. Modelowanie tabelaryczne, APN Promise 2017
3. Itzik Ben-Gan (2020) Funkcje okna w języku T-SQL dla SQL Server 2019, Promise
4. T. Morzy, Eksploracja danych. Metody i algorytmy, 2013, PWN

Dodatkowa

1. Ferrari Alberto , Russo Marco, Power BI i Power Pivot dla Excela. Analiza danych, Helion, 2021
2. U. Malik, M. Goldwasser, B. Johnston, SQL for Data Analytics: Perform fast and efficient data analysis with the power of SQL, 2019, Packt Publishing
3. A. Chodkowska Gyurics , Hurtownie danych. Teoria i praktyka , PWN, 2017
4. Bill Inmon, Building The Data Lakehouse, 2021, Technics Publications
5. M. Walker, Data Cleaning and Exploration with Machine Learning, 2022, Packt Publishing

Nakład pracy studenta i punkty ECTS

Rodzaje zajęć studenta	Średnia liczba godzin* przeznaczonych na zrealizowane rodzaje zajęć
Wykład	15
Laboratorium	15
Czytanie wskazanej literatury	30
Przygotowanie projektu	60
Przygotowanie do zaliczenia	30
Inne	30
Łączny nakład pracy studenta	Liczba godzin 180
Liczba punktów ECTS	ECTS 6

* godzina (lekcyjna) oznacza 45 minut

Efekty uczenia się dla kierunku

Kod	Treść
APD_K2_U01	Absolwent/ka potrafi wykorzystywać zaawansowane metody, narzędzia i technologie informatyczne do rozwiązywania problemów teoretycznych i problemów praktycznych w wybranych obszarach zastosowań
APD_K2_U02	Absolwent/ka potrafi zastosować zaawansowaną wiedzę matematyczną do formułowania, analizowania i rozwiązywania złożonych zadań związanych z analizą i przetwarzaniem danych
APD_K2_U05	Absolwent/ka potrafi projektować, analizować i stosować zaawansowane algorytmy
APD_K2_U06	Absolwent/ka potrafi posługiwać się wybranymi bazami danych w zakresie gromadzenia i analizowania danych oraz wyszukiwania informacji z danych
APD_K2_U07	Absolwent/ka potrafi posługiwać się różnymi formatami reprezentacji rozmaitych rodzajów danych, w szczególności integrować dane pochodzące z różnych źródeł, transformować je do wspólnej spójnej postaci
APD_K2_U08	Absolwent/ka potrafi projektować i implementować zaawansowane struktury danych
APD_K2_U09	Absolwent/ka potrafi pozyskiwać informacje z literatury, baz wiedzy, Internetu oraz innych wiarygodnych źródeł, integrować je, dokonywać ich interpretacji oraz wyciągać wnioski i formułować opinie
APD_K2_W01	Absolwent/ka zna i rozumie w pogłębionym stopniu współczesne metody, narzędzia i technologie informatyczne właściwe dla wybranych obszarów zastosowań niezbędne przy analizie i przetwarzaniu danych pochodzących z różnych źródeł
APD_K2_W04	Absolwent/ka zna i rozumie zagadnienia składowania, eksploracji i modelowania danych