



Databases in data mining

Course description sheet

Basic information

Field of study Social Informatics	Didactic cycle 2026/2027	
Major All	Course code HIFSS.II2.07657.26	
Organisational unit Faculty of Humanities	Lecture languages Polish	
Study level Second-cycle studies	Mandatoriness Obligatory	
Form of study Full-time studies	Block Core Modules	
Profile Practical	Course related to scientific research Yes	
Course coordinator	Bartosz Ziółko	
Lecturer	Bartosz Ziółko, Mirosława Długosz	
Period Semester 2	Method of verification of the learning outcomes Completing the classes	Number of ECTS credits 4
	Activities and hours Laboratory classes: 30 Lectures: 15	

Goals

C1	Learning concepts related to data warehouses, data lakes and their role in the analytical ecosystem.
C2	Learning how to design, implement and conduct data analysis processes.

Course's learning outcomes

Code	Outcomes in terms of	Learning outcomes prescribed to a field of study	Methods of verification
Knowledge - Student knows and understands:			
W1	Student zna zasady projektowania, analizowania i optymalizowania systemów gromadzenia i przetwarzania danych przestrzegając zasad związanych z zapewnieniem bezpieczeństwa danych.	IFS2P_W04	Activity during classes, Execution of exercises, Project
W2	Student dysponuje aktualną wiedzą na temat realizacji projektów bazodanowych.	IFS2P_W02	Activity during classes, Project, Presentation
Skills - Student can:			
U1	Student potrafi samodzielnie opracować projekt bazodanowy, czyli: zdefiniować cele i przeprowadzić analizę wymagań użytkownika, zaprojektować podstawową funkcjonalność realizowaną przez zaproponowane rozwiązanie,	IFS2P_U04	Activity during classes, Project
U2	Student potrafi opracować aplikację realizującą założoną funkcjonalność,	IFS2P_U03	Project, Case study
U3	Student potrafi opracować dokumentację opisującą proces tworzenia aplikacji bazodanowej, dokumentację wdrożeniową i techniczną umożliwiającą rozbudowę opracowanego projektu.	IFS2P_U05	Project, Presentation
Social competences - Student is ready to:			
K1	Student ma kompetencje pozwalające na uczestniczenie w przygotowaniu i wdrażaniu projektów społecznych opartych o wiedzę pochodzącą z baz danych.	IFS2P_K03	Activity during classes, Project

Student workload

Activity form	Average amount of hours* needed to complete each activity form
Laboratory classes	30
Lectures	15
Preparation for classes	30
Contact hours	5
Preparation of project, presentation, essay, report	20
Student workload	Hours 100
Workload involving teacher	Hours 45

* hour means 45 minutes

Program content

No.	Program content	Course's learning outcomes	Activities
1.	<p>Optymalizacja w relacyjnych bazach danych Porównanie możliwości oferowanych przez komercyjne relacyjne bazy danych. Bezpieczeństwo i administracja – realizacja uwierzytelnienia i autoryzacji. Buforowanie danych. Ścieżki dostępu do pojedynczych tabel. Rodzaje złączeń. Przetwarzanie zapytań – analiza, normalizacja, analiza semantyczna, uproszczenie i restrukturyzacja. Algebraiczne reguły transformacji zapytań. Kosztowa optymalizacja zapytań. Statystyki i koncepcja histogramów. Selektywność danych i jej wyznaczanie. Typy drzew zapytań i zagadnienie określania porządku wykonywania operacji połączenia. Narzędzia wspierające optymalizację przetwarzania zapytań.</p> <p>Rozproszone bazy danych Obiektowe bazy danych Bazy danych XML, dane XML w relacyjnych bazach danych Technologia OLAP – przetwarzanie analityczne danych</p>	W1, W2, U1, U2, U3, K1	Laboratory classes
2.	<p>Wprowadzenie do baz danych i modelu relacyjnego, rys historyczny, porównanie możliwości oferowanych przez komercyjne i niekomercyjne relacyjne bazy danych. Optymalizacja w relacyjnych bazach danych. Bezpieczeństwo i administracja – realizacja uwierzytelnienia i autoryzacji. Buforowanie danych. Ścieżki dostępu do pojedynczych tabel. Rodzaje złączeń. Przetwarzanie zapytań – analiza, normalizacja, analiza semantyczna, uproszczenie i restrukturyzacja. Algebraiczne reguły transformacji zapytań. Kosztowa optymalizacja zapytań. Statystyki i koncepcja histogramów. Selektywność danych i jej wyznaczanie. Typy drzew zapytań i zagadnienie określania porządku wykonywania operacji połączenia. Hurtownie danych, technologia OLAP. Rozproszone i obiektowe bazy danych Bazy danych XML, dane XML w relacyjnych bazach danych. Narzędzia wspierające optymalizację przetwarzania zapytań.</p>	W1, W2, U1, U2, U3, K1	Lectures

Extended information/Additional elements

Teaching methods and techniques :

Discussion, Group work, Project Based Learning, Lecture

Activities	Methods of verification	Credit conditions
Lab. classes	Activity during classes, Execution of exercises, Project	Warunkiem zaliczenia jest obecność i wykonanie ćwiczeń na laboratoriach. W zależności od realizacji przedmiotu możliwe jest kolokwium zaliczeniowe ustne lub pisemne.
Lectures	Activity during classes, Case study, Presentation	Udział w dyskusji.

Rules of participation in given classes, indicating whether student presence at the lecture is obligatory

Laboratory classes: Studenci wykonują ćwiczenia laboratoryjne zgodnie z materiałami udostępnionymi przez prowadzącego. Student jest zobowiązany do przygotowania się w przedmiocie wykonywanego ćwiczenia, co może zostać zweryfikowane kolokwium w formie ustnej lub pisemnej. Zaliczenie zajęć odbywa się na podstawie zaprezentowania rozwiązania postawionego problemu. Zaliczenie modułu jest możliwe po zaliczeniu wszystkich zajęć laboratoryjnych.

Literature

Obligatory

1. Han, J., Kamber, M., Pei, J. – Data Mining: Concepts and Techniques
2. Dokumentacja systemów baz danych: MySQL, Snowflake, LanceDB.
3. Kanał Youtube prowadzącego przedmiot: <https://www.youtube.com/bartoszziolko>
4. Repozytoria narzędzi: Weka, RapidMiner, Orange.

Optional

1. Kursy online: Stanford CS345: Data Mining oraz Coursera: Data Mining Specialization.

Scientific research and publications

Publications

1. Zajęcia będzie prowadził Mariusz Mąsior.
2. Informatyk. CTO i współzałożyciel Consonance sp. z o.o. (od 2014). Developer z doświadczeniem w zakresie inżynierii i architektury systemów. Naukowo zajmuje się przetwarzaniem sygnałów cyfrowych. Kierował kilkoma projektami dotyczącymi aplikacji zaawansowanego przetwarzania sygnałów, rozwiązań mobilnych oraz urządzeń elektronicznych.
3. Wybrane publikacje:
4. M. Mąsior, B. Ziółko, D. Skurzok, T. Jadczyk: Baza danych słownika języka polskiego ze statystykami słów dla systemu automatycznego rozpoznawania mowy. *Studia Informatica* 2011 vol. 32 no. 2B, s. 349-357.
5. Acquisition of multimodal data corpus for automatic sign language processing — Rejestracja multimedialnego korpusu danych dla automatycznego przetwarzania języka migowego / Jakub GAŁKA, Przemysław Węgrzynowicz, Mariusz MAŚSIOR // *Studia Informatica* ; ISSN 1642-0489. — Tytuł poprz.: Zeszyty Naukowe Politechniki Śląskiej. Seria Informatyka ; ISSN: 0208-7286. — 2016 vol. 37 no. 1, s. 19-31.
6. Comparative study of SQLite and Berkeley DB implementations of n-gram model of Polish language — Porównawcze studium implementacji modelu n-gramowego języka polskiego w SQLite i Berkeley DB / Dawid SKURZOK, Bartosz ZIÓŁKO, Aleksander Pohl, Tomasz JADCZYK, Mariusz MAŚSIOR // *Studia Informatica* ; ISSN 1642-0489. — Tytuł poprz.: Zeszyty Naukowe Politechniki Śląskiej. Seria Informatyka ; ISSN: 0208-7286. — 2012 vol. 33 no. 2B, s. 153-162.
7. Using components of corpus linguistics and annotation tools in sign language teaching / Katarzyna BARCZEWSKA, Jakub GAŁKA, Dorota SZULC, Mariusz MAŚSIOR, Rafał SAMBORSKI, Tomasz J. WILCZYŃSKI // *International Journal of Modern Education and Computer Science* ; ISSN 2075-0161. — 2016 vol. 8 no. 2, s. 14-21

Learning outcomes prescribed to a field of study

Code	Content
IFS2P_K03	ma pełne kompetencje pozwalające na uczestniczenie w przygotowaniu i wdrażaniu projektów społecznych oraz potrafi w pełni przewidywać wielokierunkowe skutki społeczne swojej działalności
IFS2P_U03	potrafi wykorzystać odpowiednie metody analityczne oraz ocenić ich przydatność i możliwości zastosowania rozwiązań przy projektowaniu rozwiązań złożonych problemów lokujących się na styku technologii i świata społecznego
IFS2P_U04	potrafi formułować i testować hipotezy związane z prostymi problemami wdrożeniowymi; umie zaprojektować zaawansowane rozwiązania techniczne (w formie urządzeń, obiektów, systemów lub procesów) oraz ma umiejętności pozwalające na ich realizację
IFS2P_U05	potrafi prowadząc debatę wykorzystać zaawansowaną wiedzę do szczegółowego opisu i praktycznej analizy procesów związanych z przemianami dokonującymi się w społeczeństwie pod wpływem technologii, potrafi również na zaawansowanym poziomie przetwarzać dane i wykorzystywać odpowiednie metody i narzędzia, w tym zaawansowane techniki informacyjno-komunikacyjne
IFS2P_W02	ma zaawansowaną wiedzę z zakresu metod i technik programowania, konstrukcji i ekstrakcji baz danych oraz grafiki komputerowej oraz projektowania interfejsów, interakcji i doświadczeń użytkownika
IFS2P_W04	ma zaawansowaną wiedzę z zakresu znajomości systemów zajmujących się gromadzeniem i przetwarzaniem danych oraz zna zasady projektowania, analizowania i optymalizowania systemów gromadzenia i przetwarzania danych, przy zachowaniu zasad bezpieczeństwa danych