



Bazy danych w data mining Karta opisu przedmiotu

Informacje podstawowe

Kierunek studiów Informatyka Społeczna	Cykl dydaktyczny 2026/2027	
Specjalność Wszystkie	Kod przedmiotu HIFSS.II2.07657.26	
Jednostka organizacyjna Wydział Humanistyczny	Języki wykładowe polski	
Poziom kształcenia Studia magisterskie II stopnia	Obligatoryjność Obowiązkowy	
Forma studiów Stacjonarne	Blok zajęciowy Przedmioty kierunkowe	
Profil studiów Praktyczny	Przedmiot powiązany z badaniami naukowymi Tak	
Koordinator przedmiotu	Bartosz Ziółko	
Prowadzący zajęcia	Bartosz Ziółko, Mirosława Długosz	
Okres Semestr 2	Forma zaliczenia Zaliczenie	Liczba punktów ECTS 4
	Forma prowadzenia i godziny zajęć Ćwiczenia laboratoryjne: 30 Wykład: 15	

Cele kształcenia dla przedmiotu

C1	Nauczenie pojęć związanych z hurtowniami danych, jeziorami danych (Data Lakes) i ich rolą w ekosystemie analitycznym.
C2	Nauczenie projektowania, implementowania i przeprowadzania procesów analizy danych.

Efekty uczenia się dla przedmiotu

Kod	Efekty w zakresie	Kierunkowe efekty uczenia się	Metody weryfikacji
Wiedzy - Student zna i rozumie:			
W1	Student zna zasady projektowania, analizowania i optymalizowania systemów gromadzenia i przetwarzania danych przestrzegając zasad związanych z zapewnieniem bezpieczeństwa danych.	IFS2P_W04	Aktywność na zajęciach, Wykonanie ćwiczeń, Projekt
W2	Student dysponuje aktualną wiedzą na temat realizacji projektów bazodanowych.	IFS2P_W02	Aktywność na zajęciach, Projekt, Prezentacja
Umiejętności - Student potrafi:			
U1	Student potrafi samodzielnie opracować projekt bazodanowy, czyli: zdefiniować cele i przeprowadzić analizę wymagań użytkownika, zaprojektować podstawową funkcjonalność realizowaną przez zaproponowane rozwiązanie,	IFS2P_U04	Aktywność na zajęciach, Projekt
U2	Student potrafi opracować aplikację realizującą założoną funkcjonalność,	IFS2P_U03	Projekt, Studium przypadków
U3	Student potrafi opracować dokumentację opisującą proces tworzenia aplikacji bazodanowej, dokumentację wdrożeniową i techniczną umożliwiającą rozbudowę opracowanego projektu.	IFS2P_U05	Projekt, Prezentacja
Kompetencje społecznych - Student jest gotów do:			
K1	Student ma kompetencje pozwalające na uczestniczenie w przygotowaniu i wdrażaniu projektów społecznych opartych o wiedzę pochodzącą z baz danych.	IFS2P_K03	Aktywność na zajęciach, Projekt

Treści programowe zapewniające uzyskanie efektów uczenia się dla modułu zajęć

Student ma wiedzę i umiejętności z zakresu struktury baz do pracy z big data, zna i stosuje algorytmy i struktury danych oraz wie jakie są metody ich optymalizacji.

Nakład pracy studenta

Rodzaje zajęć studenta	Średnia liczba godzin* przeznaczonych na zrealizowane aktywności
Ćwiczenia laboratoryjne	30
Wykład	15
Przygotowanie do zajęć	30
Dodatkowe godziny kontaktowe	5
Przygotowanie projektu, prezentacji, pracy pisemnej, sprawozdania	20
Łączny nakład pracy studenta	Liczba godzin 100

Liczba godzin kontaktowych	Liczba godzin 45
-----------------------------------	----------------------------

* godzina (lekcyjna) oznacza 45 minut

Treści programowe

Lp.	Treści programowe	Efekty uczenia się dla przedmiotu	Formy prowadzenia zajęć
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	Ćwiczenia laboratoryjne
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	Wykład

Informacje rozszerzone

Metody i techniki kształcenia :

Dyskusja, Praca grupowa, Metoda projektowa (ang. Project Based Learning), Wykład

Rodzaj zajęć	Metody zaliczenia	Warunki zaliczenia przedmiotu
Ćwiczenia laboratoryjne	Aktywność na zajęciach, Wykonanie ćwiczeń, Projekt	Warunkiem zaliczenia jest obecność i wykonanie ćwiczeń na laboratoriach. W zależności od realizacji przedmiotu możliwe jest kolokwium zaliczeniowe ustne lub pisemne.
Wykład	Aktywność na zajęciach, Studium przypadków, Prezentacja	Udział w dyskusji.

Warunki i sposób zaliczenia poszczególnych form zajęć, w tym zasady zaliczeń poprawkowych, a także warunki dopuszczenia do egzaminu

Projekt. Zasady i forma zaliczenia w drugim (w sesji) i trzecim (w sesji poprawkowej) terminie pozostaje bez zmian.

Sposób obliczania oceny końcowej

Ocena z projektu zakresu zastosowania baz danych do diagnozowania i rozwiązywania problemów społecznych 100%.

Sposób i tryb wyrównywania zaległości powstałych wskutek nieobecności studenta na zajęciach

Zaległości student może nadrobić w oparciu o literaturę zaleconą przez wykładowcę. Powstałe zaległości student zalicza w terminie ustalonym z wykładowcą.

Wymagania wstępne i dodatkowe

Obecność na zajęciach jest obowiązkowa. Nieobecności wymagają zaliczenia ćwiczeń i/lub literatury przedmiotu w formie i zakresie ustalonym z osobą prowadzącą zajęcia.

Zasady udziału w poszczególnych zajęciach, ze wskazaniem, czy obecność studenta na zajęciach jest obowiązkowa

Ćwiczenia laboratoryjne: 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.

Literatura

Obowiązkowa

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.

Dodatkowa

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

Badania i publikacje

Publikacje

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 MAŚIOR // *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 MAŚIOR // *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 MAŚIOR, Rafał SAMBORSKI, Tomasz J. WILCZYŃSKI // *International Journal of Modern Education and Computer Science* ; ISSN 2075-0161. — 2016 vol. 8 no. 2, s. 14-21

Kierunkowe efekty uczenia się

Kod	Treść
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