

| STUDY MODULE DESCRIPTION FORM | | |
|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-----------------------------------------------------------------|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Name of the module/subject Data integration and data mining | | Code 1010332421010335195 |
| Field of study Information Engineering | Profile of study (general academic, practical) (brak) | Year /Semester 1 / 2 |
| Elective path/specialty Safety of Computer Systems | Subject offered in: polish | Course (compulsory, elective) obligatory |
| Cycle of study: Second-cycle studies | Form of study (full-time, part-time) full-time | |
| No. of hours Lecture: 2 Classes: - Laboratory: 1 Project/seminars: 1 | | No. of credits 5 |
| Status of the course in the study program (Basic, major, other) (brak) | | (university-wide, from another field) (brak) |
| Education areas and fields of science and art technical sciences | | ECTS distribution (number and %) 5 100% |
| Responsible for subject / lecturer: dr hab. Tadeusz Pankowski email: tadeusz.pankowski@put.poznan.pl tel. 607-033-007 Wydział Elektryczny ul. Piotrowo 3A 60-965 Poznań | | |
| Prerequisites in terms of knowledge, skills and social competencies: | | |
| 1 | Knowledge | Ma uporządkowaną i podbudowaną teoretycznie wiedzę w zakresie baz danych. |
| 2 | Skills | Potrafi pozyskiwać informacje z literatury, baz danych i innych źródeł; potrafi integrować uzyskane informacje, dokonywać ich interpretacji, a także wyciągać wnioski oraz formułować i uzasadniać opinie. |
| 3 | Social competencies | Ma kompetencje społeczne wynikające z programu szkoły średniej. |
| Assumptions and objectives of the course: Poznanie metod i wybranych narzędzi przeznaczonych do integracji, analizy i eksploracji danych. | | |
| Study outcomes and reference to the educational results for a field of study | | |
| Knowledge: 1. Ma uporządkowaną i podbudowaną teoretycznie wiedzę w zakresie baz danych, hurtowni danych, analizy i eksploracji danych. - [K_W08] | | |
| Skills: 1. Potrafi pozyskiwać informacje z literatury, baz danych i innych źródeł; potrafi integrować uzyskane informacje, dokonywać ich interpretacji, a także wyciągać wnioski oraz formułować i uzasadniać opinie. - [K_U01] 2. Potrafi zaprojektować oraz zrealizować prosty system integracji danych; potrafi analizować i eksplorować dane - [K_U12] 3. Potrafi - przy formułowaniu i rozwiązywaniu informatycznych zadań inżynierskich - dostrzegać ich aspekty pozatechniczne, ekonomiczne i prawne. - [K_U21] | | |
| Social competencies: 1. Ma świadomość odpowiedzialności za pracę własną oraz gotowość podporządkowania się zasadom pracy w zespole i ponoszenia odpowiedzialności za wspólnie realizowane zadania. - [K_K04] | | |
| Assessment methods of study outcomes | | |
| Egzamin pisemny, zaliczenie zajęć laboratoryjnych i projektowych na podstawie zrealizowanych zadań i projektów. | | |
| Course description | | |

| | | |
|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|---------------------|------------------------------------|
| <p>Wykłady</p> <p>Znaczenie eksploracji danych (data mining). Eksploracja danych na tle rozwoju nauki i technologii przetwarzania danych. Proces odkrywania wiedzy z danych. Metody integracji danych, modelowania danych wielowymiarowych, budowy hurtowni danych. Język MDX i jego znaczenie w przetwarzaniu danych wielowymiarowych. Drzewa decyzyjne jako metoda klasyfikacji obiektów. Wybór zbioru treningowego, problem analizy i predykcji za pomocą drzew decyzyjnych, dziedziny zastosowań. Budowa drzewa decyzyjnego: podstawy teorii informacji (entropia układu); analiza informacyjna ? miara ilości informacji. Klasyfikacja metodą Bayesa. Odkrywanie reguł asocjacyjnych. Algorytm a priori ? odkrywanie zbiorów częstych. Grupowanie obiektów, analiza skupień (klastrowanie obiektów). Klasyfikacja metod grupowania. Grupowanie hierarchiczne. Miary odległości między klastrami (grupami). Algorytm k-średnich. Algorytm k-medoidów. Norma kosinusowa i jej znaczenie w systemach wyszukiwania informacji. Boolowskie systemy wyszukiwania informacji. Eksploracja danych tekstowych. Wykorzystanie metod statystycznych do opisu i analizy danych.</p> | | |
| <p>Zajęcia laboratoryjne</p> <p>Poznanwanie narzędzi integracji i eksploracji danych i zaznajamianie się z ich praktycznym wykorzystaniem. SSIS (SQL Server Integration Services - z tworzeniem pakietów przepływu włączanie); SSAS (SQL Server Analysis Services - język MDX, kostki, wyodrębnianie miar, wymiarów, agregacje); język DMX; SSRS (SQL Server Reporting Services - integracją z Excelem 2007. Wykorzystanie wspomnianych narzędzi do eksploracji danych: budowa drzew decyzyjnych, klasyfikacja obiektów metodą Baysa, odkrywania reguł asocjacyjnych, eksploracja tekstów (text mining).</p> | | |
| <p>Zajęcia projektowe</p> <p>Celem zajęć projektowych jest pogłębienie wiedzy w zakresie stosowania metod statystycznych do opisu i analizy danych. Wykonywane są projekty w zakresie: statystyki opisowej, badania zgodności rozkładów, estymacji parametrów, wnioskowania statystycznego, analizy wariancji, analizy korelacji i regresji. Wspomniane zadania i projekty realizowane są z wykorzystaniem systemu Statistica i Excel 2010.</p> | | |
| <p>Basic bibliography:</p> <ol style="list-style-type: none"> 1. Sturm J., Hurtownie danych. Microsoft SQL Server. Przewodnik Techniczny, Microsoft Press/APN PROMISE, Warszawa. 2. Han J., Kamber M., Data Mining. Concepts and Techniques, Morgan Kaufmann, 2001. 3. C. Seidman, Zgłębianie i analiza danych w Microsoft SQL Server 2000. Przewodnik Techniczny, Microsoft Press/APN PROMISE, Warszawa, 2002. 4. M. Rabiej, Statystyka z programem Statistica, Helion, Gliwice, 2012. | | |
| <p>Additional bibliography:</p> <ol style="list-style-type: none"> 1. D. Mendrala, M. Szeliga, Server SQL 2008. Usługi biznesowe, Helion, Gliwice, 2009. 2. R. Baeza-Yates, B. Ribeiro-Neto, Modern Information Retrieval, ACM Press, 1999. | | |
| <p>Result of average student's workload</p> | | |
| <p>Activity</p> | | <p>Time (working hours)</p> |
| 1. Uczestnictwo w wykładach | | 30 |
| 2. Udział w zajęciach laboratoryjnych | | 15 |
| 3. Udział w zajęciach projektowych | | 15 |
| 4. Przygotowanie projektów i ich realizacja | | 30 |
| 5. Studiowanie notatek z zajęć oraz zalecanej literatury przedmiotu | | 35 |
| <p>Student's workload</p> | | |
| <p>Source of workload</p> | <p>hours</p> | <p>ECTS</p> |
| Total workload | 125 | 5 |
| Contact hours | 30 | 2 |
| Practical activities | 30 | 3 |