

KARTA MODUŁU / KARTA PRZEDMIOTU

Kod modułu	Z-ID-509a
Nazwa modułu	Odkrywanie związków w danych wielowymiarowych
Nazwa modułu w języku angielskim	Discovering Relationships within Multivariate Data
Obowiązuje od roku akademickiego	2015/2016

A. USYTUOWANIE MODUŁU W SYSTEMIE STUDIÓW

Kierunek studiów	Inżynieria danych
Poziom kształcenia	I stopień
Profil studiów	Praktyczny
Forma i tryb prowadzenia studiów	Studia stacjonarne
Specjalność	Analityka danych i modelowanie
Jednostka prowadząca moduł	Katedra Informatyki i Matematyki Stosowanej
Koordinator modułu	Dr inż. Damian Krzesimowski
Zatwierdził	

B. OGÓLNA CHARAKTERYSTYKA PRZEDMIOTU

Przynależność do grupy/bloku przedmiotów	Specjalnościowy
Status modułu	Obowiązkowy
Język prowadzenia zajęć	Polski
Usytuowanie modułu w planie studiów - semestr	Semestr V
Usytuowanie realizacji przedmiotu w roku akademickim	Semestr zimowy
Wymagania wstępne	Algebra, Wnioskowanie statystyczne, Akwizycja danych pomiarowych
Egzamin (TAK/NIE)	NIE
Liczba punktów ECTS	3

Forma prowadzenia zajęć	wykład w	ćwiczenia ć	laboratorium l	projekt p	inne i
Liczba godzin w semestrze	20		20		

C. EFEKTY KSZTAŁCENIA I METODY SPRAWDZANIA EFEKTÓW KSZTAŁCENIA

Cel modułu	Celem modułu jest zapoznanie studentów z podstawowymi pojęciami, metodami i algorytmami odkrywania związków w różnorodnych danych. Student zostanie zapoznany z podstawowymi metodami odkrywania reguł asocjacyjnych, wzorców sekwencji i klastrowania przy użyciu profesjonalnych narzędzi programistycznych.
-------------------	--

Symbol efektu	Efekty kształcenia	Forma prowadzenia zajęć	Odniesienie do efektów kierunkowych	Odniesienie do efektów obszarowych
W_01	Student ma wiedzę z zakresu probablistyki i statystyki.	w, l	K_W02	T1P_W01 X1P_W02 X1P_W04
W_02	Zna podstawowe techniki pozyskiwania, gromadzenia, weryfikacji i przetwarzania danych wielowymiarowych.	w, l	K_W05	T1P_W02 T1P_W03 T1P_W06 inzP_W02 inzP_W03
W_03	Ma podstawową wiedzę z zakresu odkrywania wzorców sekwencji, odkrywania reguł asocjacyjnych i ich oceny.	w, l	K_W13	T1P_W03 T1P_W06 X1P_W02 X1P_W04
W_04	Ma podstawową wiedzę z zakresu grupowania i podziału zbioru danych oraz miar podobieństwa i odległości.	w, l	K_W02 K_W13	T1P_W01 T1P_W03 T1P_W06 X1P_W02 X1P_W04
U_01	Student potrafi przygotować dane do analizy.	l	K_U04 K_U06	T1P_U09 T1P_U16 X1P_U03
U_02	Umie zastosować system SAS® do odkrywania reguł asocjacyjnych, odkrywania wzorców sekwencji, grupowania i podziału.	l	K_U06 K_U16	T1P_U05 T1P_U08 T1P_U09 T1P_U16
U_03	Potrafi zinterpretować wyniki analiz i ocenić ich jakość.	l	K_U04	T1P_U08 T1P_U09 X1P_U02 X1P_U03
K_01	Rozumie potrzebę i zna możliwości ciągłego doskonalenia się.	l	K_K01	T1P_K01 X1P_K01 X1P_K05 inzP_K01 inzP_K02
K_02	Potrafi komunikować się w zespole również w zakresie wykraczającym poza zagadnienia techniczne.	l	K_K05	T1P_K03 X1P_K02 inzP_K01

Treści kształcenia

1. Treści kształcenia w zakresie wykładu

Nr wykładu	Treści kształcenia	Odniesienie do efektów kształcenia dla modułu
1	Źródła danych. Pojęcie odkrywania wiedzy z danych. Wybrane narzędzia programistyczne do przetwarzania danych wielowymiarowych: Ms Excel, MatLAB, SAS® i ich wykorzystanie w podstawowej analizie statystycznej danych wielowymiarowych.	W_01 W_02
2	Przygotowanie struktury danych do analiz asocjacyjnych. Podstawy analizy koszykowej. Klasyfikacja reguł asocjacyjnych. Wybrane miary oceny reguł asocjacyjnych i ich interpretacja.	W_01 W_02 W_03
3-4	Odkrywanie reguł asocjacyjnych z wykorzystaniem profesjonalnego oprogramowania (SAS). Ocena jakości reguł asocjacyjnych. Ogólny algorytm odkrywania reguł asocjacyjnych – budowa programu do analizy koszykowej w środowisku Ms Excel.	W_02 W_03
5-6	Przygotowanie struktury danych do analizy sekwencji, rola zmiennej czasowej. Klasyfikacja problemów odkrywania wzorców sekwencji; uogólnione wzorce sekwencji. Odkrywanie wzorców sekwencji z wykorzystaniem profesjonalnego oprogramowania (SAS). Ocena jakości sekwencji.	W_02 W_03
7	Pojęcie metryki. Miary podobieństw i odległości dla atrybutów binarnych, nominalnych, porządkowych i ilościowych. Obliczanie miar podobieństw i odległości z wykorzystaniem środowiska Ms Excel, MatLAB i SAS.	W_01 W_04
8-9	Pojęcie klastra; omówienie podstawowych typów grupowania i podziału. Hierarchiczna i niehierarchiczna analiza skupień. Zagadnienie normalizacji zmiennych. Przykładowy algorytm aglomeracyjny – metoda najbliższego sąsiada (SAS).	W_02 W_04
10	Problem podziału zbioru i grupowania pojęciowego. Ocena jakości grupowania. Przykładowe metody podziału: k-średnich, k-medoid, SSN Kohonena.	W_02 W_04

2. Treści kształcenia w zakresie ćwiczeń

3. Treści kształcenia w zakresie zadań laboratoryjnych

Nr zajęć lab.	Treści kształcenia	Odniesienie do efektów kształcenia dla modułu
1	Przegląd materiałów dotyczących odkrywania związków w danych wielowymiarowych dostępnych w domenie publicznej. Przegląd wybranych serwisów i repozytoriów danych. Zapoznanie z przykładowymi danymi pod kątem analiz asocjacyjnych i analiz skupień.	U_01 K_01
2	Pozyskanie i przygotowanie danych do analiz. Wykonanie analityki prostej dla danych – eksploatacja wybranych programów komputerowych. <i>Opracowanie raportu z eksploracji danych wybranych do analiz.</i>	W_01 U_01 U_03 K_01 K_02
3	Przekształcenie struktury danych dla potrzeb analiz asocjacyjnych – dobór i wykorzystanie właściwego oprogramowanie do wykonania zadania. Przeprowadzenie analizy koszykowej na przygotowanych danych z wykorzystaniem profesjonalnego oprogramowania (SAS).	W_03 U_02

4	Analiza i interpretacja wyników analizy koszykowej. Ocena jakości reguł asocjacyjnych, kryteria wyboru reguł najlepszych. <i>Opracowanie raportu z wykonania prac z ćwiczeń nr 3 i 4, sformułowanie wniosków i zaleceń.</i>	U_02 U_03 K_02
5-6	Przekształcenie struktury danych dla potrzeb analiz sekwencyjnych – dobór i wykorzystanie właściwego oprogramowania do wykonania zadania. Przeprowadzenie analizy sekwencyjnej na przygotowanych danych z wykorzystaniem profesjonalnego oprogramowania (SAS). <i>Opracowanie raportu z wykonania prac z ćwiczeń nr 5 i 6, sformułowanie wniosków i zaleceń.</i>	W_03 U_02 K_02
7	Obliczanie wartości miar odległości danych ilościowych z wykorzystaniem różnych metryk. Uwzględnianie aspektu wielowymiarowości danych w hierarchicznej analizie skupień. Normalizacja zmiennych.	W_04 U_02 U_03
8	Klastrowanie ilościowych danych wielowymiarowych z wykorzystaniem oprogramowania SAS – algorytm najbliższego sąsiada. Kryteria przerwania pracy algorytmu aglomeracyjnego. <i>Opracowanie raportu z wykonania prac z ćwiczeń nr 7 i 8, sformułowanie wniosków i zaleceń.</i>	W_04 U_02 U_03 K_02
9-10	Obliczanie wartości miar podobieństwa danych jakościowych. Podział zbioru danych ilościowych i jakościowych z wykorzystaniem oprogramowania SAS – algorytm k-średnich, SSN Kohonena. Ocena jakości podziału zbioru. <i>Opracowanie raportu z wykonania prac z ćwiczeń nr 9 i 10, sformułowanie wniosków i zaleceń.</i>	W_04 U_02 U_03 K_02

4. Charakterystyka zadań projektowych

5. Charakterystyka zadań w ramach innych typów zajęć dydaktycznych

Metody sprawdzania efektów kształcenia

Symbol efektu	Metody sprawdzania efektów kształcenia (sposób sprawdzenia, w tym dla umiejętności – odwołanie do konkretnych zadań projektowych, laboratoryjnych, itp.)
W_01	Aktywność na ćwiczeniach. Poprawne wykonanie raportów.
W_02	Aktywność na ćwiczeniach. Poprawne wykonanie raportów.
W_03	Aktywność na ćwiczeniach. Poprawne wykonanie raportów.
W_04	Aktywność na ćwiczeniach. Poprawne wykonanie raportów.
U_01	Poprawne wykonanie wszystkich ćwiczeń praktycznych na ćwiczeniach laboratoryjnych. Poprawne wykonanie raportów.
U_02	Poprawne wykonanie wszystkich ćwiczeń praktycznych na ćwiczeniach laboratoryjnych. Poprawne wykonanie raportów.
U_03	Poprawne wykonanie wszystkich ćwiczeń praktycznych na ćwiczeniach laboratoryjnych. Poprawne wykonanie raportów.
K_01	Aktywność na ćwiczeniach. Wnioski i zalecenia zamieszczone w raportach.
K_02	Aktywność na ćwiczeniach.

D. NAKŁAD PRACY STUDENTA

Bilans punktów ECTS			
Lp.	Rodzaj aktywności	Obciążenie studenta	Jednostka
1.	Udział w wykładach	20	h
2.	Udział w ćwiczeniach		
3.	Udział w laboratoriach	20	h
4.	Udział w zajęciach projektowych		
5.	Udział w konsultacjach (2-3 razy w semestrze)	6	h
6.	Konsultacje projektowe		
7.	Udział w egzaminie		
8.			
9.	Liczba godzin realizowanych przy bezpośrednim udziale nauczyciela akademickiego	46	h
10.	Liczba punktów ECTS, którą student uzyskuje na zajęciach wymagających bezpośredniego udziału nauczyciela akademickiego <i>(1 punkt ECTS=27 godzin obciążenia studenta)</i>	1,7	ECTS
11.	Samodzielne studiowanie tematyki wykładów	7	h
12.	Samodzielne przygotowanie się do ćwiczeń		
13.	Samodzielne przygotowanie się do kolokwium		
14.	Samodzielne przygotowanie się do laboratoriów	18	h
15.	Wykonanie sprawozdań	10	h
16.	Przygotowanie do kolokwium końcowego z laboratorium		
17.	Wykonanie projektu lub dokumentacji		
18.	Przygotowanie do egzaminu		
19.			
20.	Liczba godzin samodzielnej pracy studenta	35	h
21.	Liczba punktów ECTS, którą student uzyskuje w ramach samodzielnej pracy <i>(1 punkt ECTS=27 godzin obciążenia studenta)</i>	1,3	ECTS
22.	Sumaryczne obciążenie pracą studenta	81	h
23.	Punkty ECTS za moduł <i>1 punkt ECTS=27 godzin obciążenia studenta</i>	3	ECTS
24.	Nakład pracy związany z zajęciami o charakterze praktycznym <i>Suma godzin związanych z zajęciami praktycznymi</i>	54	h
25.	Liczba punktów ECTS, którą student uzyskuje w ramach zajęć o charakterze praktycznym <i>1 punkt ECTS=27 godzin obciążenia studenta</i>	2	ECTS

E. LITERATURA

Wykaz literatury	<ol style="list-style-type: none">1. Andrienko N., Andrienko G., <i>Exploratory Analysis of Spatial and Temporal Data. A systematic Approach</i>, Springer, 2005.2. Ballingsley P., <i>Prawdopodobieństwo i miara</i>, PWN, 2009.3. Hoaglin D. C., Mosteller F., Tukey J. W., <i>Exploring Data Tables, Trends and Shapes</i>, 1985.4. Martinez W. L., Martunez A. R., Solka J., <i>Exploratory Data Analysis with MATLAB</i>, Chapman & Hall/CRC, 2010.5. Elliot J., Marsh C., <i>Exploring Data: An introduction to Data Analysis for Social Scientists</i>, 2009.6. Hand D., Mannila H., Smyth P., <i>Eksploracja danych</i>, WNT, Warszawa 2005.7. Krzyśko M., Wołyński W., Górecki T., Skorzybut M., <i>Systemy uczące się. Rozpoznawanie wzorców, analiza skupień i redukcja wymiarowości</i>, WNT, Warszawa 2008.8. Larose D. T., <i>Odkrywanie wiedzy z danych</i>, PWN, 2006.9. Larose D. T., <i>Metody i modele eksploracji danych</i>, PWN, 2012.10. Luszczewicz A., <i>Statystyka ogólna</i>, PWE, 1987.11. Morzy T., <i>Eksploracja danych</i>, PWN, 2013.12. Osowski S., <i>Metody i narzędzia eksploracji danych</i>, Wyd. Helion, Gliwice 2013.
Witryna WWW modułu/przedmiotu	