

KARTA OPISU MODUŁU KSZTAŁCENIA		
Nazwa modułu/przedmiotu Analiza danych i technologie semantyczne		Kod 1010515331010510112
Kierunek studiów Informatyka	Profil kształcenia (ogólnoakademicki, praktyczny) ogólnoakademicki	Rok / Semestr 2 / 3
Ścieżka obieralności/specjalność Aplikacje mobilne i wbudowane dla	Przedmiot oferowany w języku: polski	Kurs (obligatoryjny/obieralny) obieralny
Stopień studiów: II stopień	Forma studiów (stacjonarna/niestacjonarna) niestacjonarna	
Godziny Wykłady: 16 Ćwiczenia: - Laboratoria: 16 Projekty/seminaria: -		Liczba punktów 3
Status przedmiotu w programie studiów (podstawowy, kierunkowy, inny) (ogólnouczelniany, z innego kierunku) kierunkowy z danego kierunku		
Obszar(y) kształcenia i dziedzina(y) nauki i sztuki nauki techniczne		Podział ECTS (liczba i %) 3 100%
Odpowiedzialny za przedmiot / wykładowca: dr inż. Tomasz Łukaszewski email: Tomasz.Lukaszewski@cs.put.poznan.pl tel. 61 6652920 Instytut Informatyki ul. Piotrowo 2, 60-965 Poznań		
Wymagania wstępne w zakresie wiedzy, umiejętności, kompetencji społecznych:		
1	Wiedza:	Student rozpoczynający ten przedmiot powinien posiadać podstawową wiedzę w zakresie technologii internetowych (w tym języka XML), podstaw logiki i baz danych.
2	Umiejętności:	Powinien posiadać umiejętność rozwiązywania podstawowych problemów w zakresie projektowania systemów informatycznych i ich realizacji oraz umiejętność pozyskiwania informacji ze wskazanych źródeł.
3	Kompetencje społeczne	Powinien również rozumieć konieczność poszerzania swoich kompetencji. Ponadto w zakresie kompetencji społecznych student musi prezentować takie postawy jak uczciwość, odpowiedzialność, wytrwałość, ciekawość poznawcza, kreatywność, kultura osobista, szacunek dla innych ludzi.
Cel przedmiotu: 1. Wprowadzenie do analizy danych w zakresie metod statystycznych oraz uczenia nadzorowanego z przykładów 2. Przekazanie studentom podstawowej wiedzy w zakresie technologii semantycznych, w tym koncepcji Web 3.0 (semantyczna sieć WWW). 3. Rozwijanie u studentów umiejętności rozwiązywania problemów w zakresie sposobów użytkowania i projektowania systemów wykorzystujących technologie semantyczne.		
Efekty kształcenia i odniesienie do kierunkowych efektów kształcenia		
Wiedza: 1. ma uporządkowaną, podbudowaną teoretycznie wiedzę ogólną w zakresie języków i paradygmatów programowania, sztucznej inteligencji, baz danych - [K_W4] 2. ma podbudowaną teoretycznie szczegółową wiedzę związaną z wybranymi zagadnieniami z zakresu informatyki, takimi jak: uczenie nadzorowane, technologie semantyczne, Web 3.0 - [K_W5] 3. zna podstawowe metody, techniki i narzędzia stosowane przy rozwiązywaniu złożonych zadań inżynierskich z obszaru informatyki dotyczącego analizy danych i technologii semantycznych - [K_W8]		
Umiejętności: 1. potrafi pozyskiwać informacje z literatury, baz danych oraz innych źródeł (w języku ojczystym i angielskim), integrować je, dokonywać ich interpretacji i krytycznej oceny, wyciągać wnioski oraz formułować i wyczerpująco uzasadniać opinie, - [K_U1] 2. potrafi określić kierunki dalszego uczenia się i zrealizować proces samokształcenia - [K_U5] 3. potrafi wykorzystać do formułowania i rozwiązywania zadań inżynierskich i prostych problemów badawczych metody analityczne - [K_U9] 4. potrafi formułować i testować hipotezy związane z problemami inżynierskimi i prostymi problemami badawczymi - [K_U12]		
Kompetencje społeczne:		

1. rozumie, że w informatyce wiedza i umiejętności bardzo szybko stają się przestarzałe - [K_K1]
2. zna przykłady i rozumie przyczyny wadliwie działających systemów informatycznych, które doprowadziły do poważnych strat finansowych, społecznych lub też do poważnej utraty zdrowia, a nawet życia - [K_K4]
3. potrafi odpowiednio określić priorytety służące realizacji określonego przez siebie lub innych zadania - [K_K6]
4. ma świadomość roli społecznej absolwenta uczelni technicznej, a zwłaszcza rozumie potrzebę formułowania i przekazywania społeczeństwu, w szczególności poprzez środki masowego przekazu, informacji i opinii dotyczących osiągnięć techniki i innych aspektów działalności inżynierskiej; - [K_K9]
5. podejmuje starania, aby przekazać takie informacje i opinie w sposób powszechnie zrozumiały, z uzasadnieniem różnych punktów widzenia - [K_K9]

Sposoby sprawdzenia efektów kształcenia

Efekty kształcenia przedstawione wyżej weryfikowane są w następujący sposób:

Ocena formująca:

- a) w zakresie wykładów:
 - na podstawie odpowiedzi na pytania dotyczące materiału omówionego na poprzednich wykładach.
- b) w zakresie laboratoriów:
 - na podstawie oceny bieżącego postępu realizacji zadań cząstkowych.

Ocena podsumowująca:

- a) w zakresie wykładów weryfikowanie założonych efektów kształcenia realizowane jest przez:
 - ocenę wiedzy z zakresu analizy danych na podstawie kolokwium zaliczeniowego
 - ocenę wiedzy i umiejętności podczas realizacji w grupie projektu informatycznego zawierającego elementy technologii semantycznych.
- b) w zakresie laboratoriów weryfikowanie założonych efektów kształcenia realizowane jest przez:
 - ocenę prezentowanego przez studenta sprawozdania z realizacji projektu w ramach publicznej prezentacji projektu.

Uzyskiwanie punktów dodatkowych za aktywność podczas zajęć, a szczególnie za:

- omówienia dodatkowych aspektów zagadnienia,
- efektywność zastosowania zdobytej wiedzy podczas rozwiązywania zadanych problemów,
- uwagi prowadzące do udoskonalenia materiałów dydaktycznych lub procesu dydaktycznego.

Treści programowe

Program wykładu obejmuje następujące zagadnienia:

Analiza danych w zakresie statystyki opisowej (miary opisu, rozproszenia) oraz badanie zależności między zmiennymi liczbowymi. Analiza danych w zakresie uczenia nadzorowanego z przykładów (cel: zrozumienie danych i generalizacji danych na nowe przypadki). Podejścia: drzewa decyzyjne, podejścia regułowe, naiwny klasyfikator Bayesowski.

Pojęcie Internetu Semantycznego. Warstwowa architektura języków Internetu Semantycznego. Klasyczne sieci semantyczne i trójkowy model danych. Język opisu zasobów RDF Reprezentacja wiedzy za pomocą ontologii. Język zapytań SPARQL. Pojęcie końcówki SPARQL. Silniki zapytań SPARQL. Semantyczne systemy Wiki. Przykłady aktualnych inicjatyw związanych z semantyczną integracją danych i wiedzy (<http://schema.org>, Graf Wiedzy Google) oraz wykorzystanie technologii semantycznych w Internecie Przedmiotów.

Program laboratorium obejmuje następujące zagadnienia:

Analiza danych (statystyka opisowa i analiza zależności między zmiennymi) z wykorzystaniem arkusza Excel. Analiza danych (uczenie nadzorowane) z wykorzystaniem narzędzia WEKA (drzewa decyzyjne, systemy regułowe, naiwny klasyfikator Bayesowski). Modelowanie ontologii z wykorzystaniem edytora ontologii Protégé. Odpytywanie heterogenicznych źródeł wiedzy za pomocą języka SPARQL. Tworzenie i zarządzanie treścią oraz modelowanie i wyszukiwanie wiedzy w środowisku Semantic Media Wiki.

Metody dydaktyczne:

1. wykład: prezentacja multimedialna, prezentacja ilustrowana przykładami, demonstracja, dyskusja.
2. ćwiczenia laboratoryjne: ćwiczenia praktyczne, praca w zespole, pokaz multimedialny, gry integracyjne, demonstracja, dyskusja.

Literatura podstawowa:

1. Analiza statystyczna. Microsoft Excel 2010 PL, Carlberg C., Helion, 2012.
2. Eksploracja danych. Metody i algorytmy. Morzy T., PWN, 2013.
3. Ontologie w systemach informatycznych, Goczyła K., EXIT, 2011

Literatura uzupełniająca:		
1. Semantic Web Programming, Hebel J., Fisher M., Blace R., Perez-Lopez A., Dean M., Wiley, 2009		
2. Semantic Web for the Working Ontologist, Allemang D. and Hendler J., Morgan Kaufmann , 2008		
3. Handbook on Ontologies, Staab S., Studer R., Springer, 2009		
4. Working with MediaWiki, Koren Y., WikiWorks Press, 2012 (aktualizacja 2014)		
Bilans nakładu pracy przeciętnego studenta		
Czynność	Czas (godz.)	
1. udział w wykładach	16	
2. udział w zajęciach laboratoryjnych	16	
3. przygotowanie do zajęć laboratoryjnych	8	
4. dokończenie (w ramach pracy własnej) ćwiczeń laboratoryjnych	8	
5. udział w konsultacjach związanych z realizacją ćwiczeń laboratoryjnych i projektu (również drogą elektroniczną)	4	
6. napisanie projektu zaliczeniowego, przygotowanie prezentacji multimedialnej projektu	20	
7. zapoznanie się ze wskazaną literaturą i materiałami dydaktycznymi	16	
8. przygotowanie do kolokwium zaliczeniowego	2	
Obciążenie pracą studenta		
forma aktywności	godzin	ECTS
Łączny nakład pracy	90	3
Zajęcia wymagające bezpośredniego kontaktu z nauczycielem	40	2
Zajęcia o charakterze praktycznym	40	2