



KAPITAŁ LUDZKI
NARODOWA STRATEGIA SPÓJNOŚCI

Projekt współfinansowany przez
Unię Europejską w ramach
Europejskiego Funduszu
Społecznego

UNIA EUROPEJSKA
EUROPEJSKI
FUNDUSZ SPOŁECZNY



Nazwa przedmiotu		Kod ECTS	
Komputerowe przetwarzanie tekstów		11.3.1392	
Nazwa jednostki prowadzącej przedmiot			
Instytut Informatyki			
Studia			
wydział	kierunek	poziom	pierwszego stopnia
Wydział Matematyki, Fizyki i Informatyki	Informatyka	forma	stacjonarne
		moduł	wszystkie
		specjalnościowy	wszystkie
		specjalizacja	wszystkie
Nazwisko osoby prowadzącej (osób prowadzących)			
prof. UG, dr hab. Joanna Jędrzejowicz			
Formy zajęć, sposób ich realizacji i przypisana im liczba godzin		Liczba punktów ECTS	
Formy zajęć		3 Przedmiot fakultatywny w wymiarze 15h wykładu i 15h laboratorium + praca własna studenta.	
Wykład, Ćw. laboratoryjne			
Sposób realizacji zajęć			
zajęcia w sali dydaktycznej			
Liczba godzin			
Wykład: 15 godz., Ćw. laboratoryjne: 15 godz.			
Termin realizacji przedmiotu			
2020/2021 zimowy			
Status przedmiotu		Język wykładowy	
fakultatywny (do wyboru)		polski	
Metody dydaktyczne		Forma i sposób zaliczenia oraz podstawowe kryteria oceny lub wymagania egzaminacyjne	
<ul style="list-style-type: none"> - Wykład z prezentacją multimedialną - używanie wyspecjalizowanych programów 		Sposób zaliczenia	
		<ul style="list-style-type: none"> - Zaliczenie na ocenę - Zaliczenie (zał) 	
		Formy zaliczenia	
		<ul style="list-style-type: none"> - wykonanie pracy zaliczeniowej - projekt lub prezentacja - ustalenie oceny zaliczeniowej na podstawie ocen cząstkowych otrzymywanych w trakcie trwania semestru - kolokwium 	
		Podstawowe kryteria oceny	
		Potwierdzone umiejętności oraz posiadanie wiedzy na temat objęty przedmiotem. Wykonanie projektu.	
Sposób weryfikacji założonych efektów kształcenia			

zakładany efekt kształcenia	egzamin	kolokwium	program	zadania na laboratorium	raport	aktywność w dyskusji	obserwacja postawy
Wiedza							
W01		X		X			
W02		X		X			
W03		X		X			
Umiejętności							
U01			X	X			X
U02			X	X			
Kompetencje							
K01			X	X			X

Określenie przedmiotów wprowadzających wraz z wymogami wstępnymi**A. Wymagania formalne**

brak

B. Wymagania wstępne**Cele kształcenia**

Celem przedmiotu jest przekazanie studentom podstawowej wiedzy w zakresie zastosowania metod informatycznych do przetwarzania dokumentów: reprezentacja dokumentu, metody porównywania dokumentów, wyszukiwania dokumentów pasujących do zapytania, klasyfikacja dokumentów, korzystanie z cyfrowych zasobów językowych, analiza opinii.

Treści programowe

Metody reprezentacji dokumentów i struktury danych.
Wyszukiwanie informacji w tekście, miary odległości. Statystyczne prawa językowe.
Etapy przetwarzania tekstów (usuwanie stop-words, stemming, podział na części mowy).
Klasyfikacja dokumentów.
Cyfrowe zasoby językowe: WordNet, SentiWordNet.
Analiza opinii.
Ukryta analiza semantyczna.

Wykaz literatury

S. Ghosh, D. Gunning - Natural Language Processing Fundamentals, Packt Publishing 2019
B. Liu - Web data mining, Springer 2007
S. Bird, E. Klein, E. Loper - Natural language processing with Python, O'Reilly 2009
A. Mykowiecka - Inżynieria lingwistyczna. Komputerowe przetwarzanie tekstów w języku naturalnym, Wyd. PJWSTK 2007

Kierunkowe efekty kształcenia

K_W02: posiada wiedzę w zakresie matematyki dyskretnej oraz metod probabilistycznych i statystyki
K_W03: ma uporządkowaną, podbudowaną teoretycznie wiedzę ogólną w zakresie algorytmów i struktur danych, języków formalnych, teorii automatów i złożoności obliczeniowej
K_U01: potrafi zastosować wiedzę matematyczną do formułowania, analizowania i rozwiązywania problemów związanych z informatyką
K_U02: potrafi projektować i analizować algorytmy pod kątem ich poprawności i złożoności obliczeniowej wykorzystując odpowiednie techniki algorytmiczne i struktury danych
K_U04: potrafi precyzyjnie formułować pytania, służące pogłębieniu własnego zrozumienia danego tematu lub odnalezieniu brakujących elementów rozumowania
K_U06: potrafi pozyskiwać informacje z literatury, Internetu oraz innych źródeł, integrować je, oceniać ich wiarygodność, dokonywać interpretacji oraz wyciągać

Wiedza

W01: student zna struktury danych i nazywa metody reprezentowania dokumentu,
W02: student wymienia różne metody określania wag słów oraz liczenia odległości pomiędzy dokumentami
W03: student potrafi omówić metody klasyfikacji dokumentów

Umiejętności

U01: student modeluje dokumenty przy pomocy wybranych słów kluczowych
U02: student pracuje w zespole nad zadaniami przygotowanymi przez prowadzącego laboratorium

Kompetencje społeczne (postawy)

K01: student pracuje w zespole nad zadaniem przez prowadzącego projektem, dyskutuje o możliwych rozwiązaniach i poszukuje rozwiązań alternatywnych

<p>wnioski i formułować opinie K_K01: zna ograniczenia własnej wiedzy i rozumie potrzebę dalszego uczenia się K_K03: potrafi i jest gotów formułować opinie na temat podstawowych zagadnień informatycznych</p>	
Kontakt jj@inf.ug.edu.pl	