

Eksploracja zasobów internetowych - opis przedmiotu

Informacje ogólne	
Nazwa przedmiotu	Eksploracja zasobów internetowych
Kod przedmiotu	11.3-WE-BEP-EZI
Wydział	Wydział Informatyki, Elektrotechniki i Automatyki
Kierunek	Biznes elektroniczny
Profil	praktyczny
Rodzaj studiów	pierwszego stopnia z tyt. inżyniera
Semestr rozpoczęcia	semestr zimowy 2018/2019

Informacje o przedmiocie	
Semestr	5
Liczba punktów ECTS do zdobycia	4
Typ przedmiotu	obowiązkowy
Język nauczania	polski
Sylabus opracował	prof. dr hab. inż. Dariusz Uciński

Formy zajęć					
Forma zajęć	Liczba godzin w semestrze (stacjonarne)	Liczba godzin w tygodniu (stacjonarne)	Liczba godzin w semestrze (niestacjonarne)	Liczba godzin w tygodniu (niestacjonarne)	Forma zaliczenia
Wykład	15	1	9	0,6	Zaliczenie na ocenę
Projekt	30	2	18	1,2	Zaliczenie na ocenę

Cel przedmiotu

- zapoznanie studentów z podstawowymi modelami i technikami odkrywania informacji znajdujących się w sieci Internet
- zapoznanie ze sposobami działania algorytmów text mining
- ukształtowanie umiejętności eksploracji zasobów internetowych w oparciu o oprogramowanie statystyczne.

Wymagania wstępne

podstawy analityki biznesowej

Zakres tematyczny

Analiza struktury sieci WWW. Wyszukiwanie informacji tekstowych i wyszukiwanie w Internecie: wyszukiwarki internetowe; roboty internetowe; indeksowanie i wyszukiwanie według słów kluczowych; ocena jakości wyszukiwania; wyszukiwanie miar podobieństwa. Ranking oparty o strukturę połączeń: analiza sieci społecznych; algorytm PageRank; autorytety i koncentratory; wyszukiwanie oparte na podobieństwie strukturalnym.

Analiza zawartości sieci WWW. Grupowanie: aglomeracyjne grupowanie hierarchiczne; algorytm k-średnich; grupowanie oparte na prawdopodobieństwie; techniki wspólnego filtrowania. Ocena grupowania: probabilistyczne funkcje kryterialne; minimalna długość opisu i ocena cech; ocena za pomocą odwzorowania klas do grup; entropia. Klasyfikacja: algorytm najbliższego sąsiada; wybór cech; naiwny algorytm Bayesa; relacyjne uczenie się.

Analiza użytkowania sieci WWW. Wstępne przetwarzanie danych: czyszczenie i filtrowanie; analiza kliknięć; identyfikacja użytkownika i sesji; katalogi i przypisanie kategorii. Eksploracyjna analiza użytkowania: liczba żądań w sesji; zależność między długością sesji a liczbą żądań użytkownika; średni czas na stronę; czas dla pojedynczych stron. Modelowanie użytkowania sieci WWW: grupowanie; reguły asocjacyjne; klasyfikacja.

Analityka Text Mining. Klasyfikacja dokumentów. Ontologie. Ekstrakcja informacji. Klasteryzacja. Analiza trendu. Eksploracyjny text mining. Analiza sentymentu.

Metody kształcenia

Wykład - wykład konwencjonalny.

Projekt - zajęcia praktyczne w laboratorium komputerowym.

Efekty uczenia się i metody weryfikacji osiągnięcia efektów uczenia się

Opis efektu	Symbole efektów	Metody weryfikacji	Forma zajęć
Zna metody uczenia maszynowego i eksploracji danych porządkujących sieć według zawartości	• K_W02	• sprawdzian	• Wykład
Potrafi stosować metody eksploracji danych do badań użytkowników sieci w celach marketingowych i handlowych	• K_U06 • K_U16	• przygotowanie projektu	• Projekt
Potrafi zastosować oprogramowanie statystyczne do odkrywania wzorców w strukturze, zawartości i użytkowaniu sieci WWW	• K_U06 • K_U16	• przygotowanie projektu	• Projekt

Opis efektu	Symbole efektów	Metody weryfikacji	Forma zajęć
Potrafi stosować oprogramowanie eksploracji danych tekstowych	• K_U06 • K_U16	• przygotowanie projektu	• Projekt
Zna podstawowe idee i metody wyszukiwania informacji tekstowej	• K_W02	• sprawdzian	• Wykład

Warunki zaliczenia

Wykład – uzyskanie oceny pozytywnej ze sprawdzianu w formie pisemnej i/lub ustnej, przeprowadzonego na koniec semestru.

Projekt – ocena końcowa stanowi sumę ważoną ocen uzyskanych za realizację poszczególnych elementów projektu oraz formy jego prezentacji. Wkład poszczególnych elementów oceny: ocena projektu - 75%, wizualna forma prezentacji projektu 25%.

Ocena końcowa = 50 % oceny zaliczenia z formy zajęć wykład + 50 % oceny zaliczenia z formy zajęć projekt.

Literatura podstawowa

1. Markov, Z., Larose, D.T., Eksploracja zasobów internetowych: Analiza struktury, zawartości i użytkowania sieci WWW, Warszawa, PWN, 2009.
2. Chakraborty, G., Pagolu, M., Garla, S., Text Mining and Analysis: Practical Methods, Examples and Case Studies Using SAS, Cary, SAS Press, 2013.
3. Segaran, T., Programming Collective Intelligence, Cambridge, O'Reilly, 2007.

Literatura uzupełniająca

Uwagi

Zmodyfikowane przez prof. dr hab. inż. Dariusz Uciński (ostatnia modyfikacja: 28-04-2018 10:53)

Wygenerowano automatycznie z systemu SyllabUZ