

Technologie Big Data - opis przedmiotu

Informacje ogólne	
Nazwa przedmiotu	Technologie Big Data
Kod przedmiotu	11.3-WE-INFD-TechBD
Wydział	Wydział Informatyki, Elektrotechniki i Automatyki
Kierunek	Informatyka
Profil	ogólnoakademicki
Rodzaj studiów	drugiego stopnia z tyt. magistra inżyniera
Semestr rozpoczęcia	semestr zimowy 2018/2019

Informacje o przedmiocie	
Semestr	2
Liczba punktów ECTS do zdobycia	6
Typ przedmiotu	obowiązkowy
Język nauczania	polski
Sylabus opracował	dr hab. inż. Artur Gramacki, prof. UZ

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	30	2	18	1,2	Egzamin
Laboratorium	30	2	18	1,2	Zaliczenie na ocenę

Cel przedmiotu

Nauczenie studentów doboru odpowiednich technik analizy danych w zależności od skali rozpatrywanego problemu oraz rodzaju przeprowadzanej analizy.

Nauczenie studentów pracy z wykorzystaniem nowoczesnych platform do składowania i przetwarzania danych.

Zapoznanie studentów z wybranymi technikami analizowania dużych zbiorów danych, głównie tekstowych.

Wymagania wstępne

Bazy danych.

Znajomość podstaw statystyki.

Umiejętność programowania w języku Java.

Zakres tematyczny

Analiza danych nieustrukturyzowanych: text mining.

Analiza struktury sieci WWW: wyszukiwanie informacji tekstowych i wyszukiwanie w Internecie; ranking oparty o strukturę połączeń.

Analiza użytkowania sieci WWW: wstępne przetwarzanie danych; eksploracyjna analiza użytkowania sieci; grupowanie, reguły asocjacyjne i klasyfikacja.

Fenomen Big Data, jego charakterystyka oraz wpływ na istniejące rozwiązania analityczne.

Analityka biznesowa na dużą skalę; nowoczesne rozwiązania wykorzystywane do przesyłania, składowania oraz przetwarzania dużych zbiorów danych.

Architektura nowoczesnych systemów do składowania i przetwarzania Big Data na przykładzie platform Elasticsearch, Apache Hadoop, Apache Hbase.

Analityka danych tekstowych w czasie rzeczywistym z wykorzystaniem platformy ElasticSearch.

Metody kształcenia

Wykład konwencjonalny, ćwiczenia laboratoryjne, praca w zespole

Efekty uczenia się i metody weryfikacji osiągnięcia efektów uczenia się

Opis efektu	Symbole efektów	Metody weryfikacji	Forma zajęć
Student potrafi scharakteryzować cechy nowoczesnych platform Big Data	- K_W09 - K_W11	sprawdzian z progami punktowymi	Wykład
Student potrafi w praktyczny sposób wykorzystać narzędzia oferowane przez platformę Elasticsearch w celu przeprowadzenia zaawansowanej analizy danych oraz ich eksploracji w czasie rzeczywistym.	- K_W09 - K_W11	projekt	Laboratorium

Opis efektu	Symbole efektów	Metody weryfikacji	Forma zajęć
Student jest w stanie opisać tradycyjny model analizy danych wykorzystywany w data i text miningu.	• K_W09 • K_W12	• sprawdzian z progami punktowymi	• Wykład
Student potrafi w praktyczny sposób wykorzystać wybraną platformę do przeprowadzenia analizy big data.	• K_W09 • K_W11	• projekt	• Laboratorium
Student potrafi wyjaśnić pojęcie Big Data oraz przedstawić główne cechy jakie zawierają zbiory danych określane tą nazwą.	• K_W12	• sprawdzian z progami punktowymi	• Wykład

Warunki zaliczenia

Wykład - warunkiem zaliczenia jest uzyskanie pozytywnej oceny z egzaminu przeprowadzonego w formie zaproponowanej przez prowadzącego

Laboratorium - warunkiem zaliczenia jest uzyskanie ocen pozytywnych z wszystkich ćwiczeń laboratoryjnych

Składowe oceny końcowej = wykład: 50% + laboratorium: 50%

Literatura podstawowa

Larose D.T.: Metody i modele eksploracji danych, PWN, Warszawa, 2008

Markov Z., Larose D.T.: Eksploracja zasobów internetowych, PWN, Warszawa, 2009

Gormley C., Tong Z.: Elasticsearch: The Definitive Guide, 2015

George L., HBase: The Definitive Guide, O'Reilly Media, 2011

White T., Hadoop: The Definitive Guide, 3rd Edition, O'Reilly Media / Yahoo Press, 2012.

Literatura uzupełniająca

Uwagi

Zmodyfikowane przez dr hab. inż. Artur Gramacki, prof. UZ (ostatnia modyfikacja: 30-04-2018 22:27)

Wygenerowano automatycznie z systemu SylabUZ