

Big Data i analityka biznesowa - opis przedmiotu

Informacje ogólne	
Nazwa przedmiotu	Big Data i analityka biznesowa
Kod przedmiotu	11.3-WI-INFD-BDiAB
Wydział	Wydział Informatyki, Elektrotechniki i Automatyki
Kierunek	Informatyka / Przemysłowe Systemy Informatyczne
Profil	ogólnoakademicki
Rodzaj studiów	drugiego stopnia z tyt. magistra inżyniera
Semestr rozpoczęcia	semestr letni 2016/2017

Informacje o przedmiocie	
Semestr	1
Liczba punktów ECTS do zdobycia	6
Typ przedmiotu	obowiązkowy
Język nauczania	polski
Sylabus opracował	dr inż. Mariusz Jacyno

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

Cel przedmiotu

Zapoznanie studentów z pojęciem analityki biznesowej oraz jej zastosowaniem do analizy dużych zbiorów danych zawartych w mediach społecznościowych, systemach ERP oraz nowoczesnych aplikacjach e-biznesowych.

Nauczenie studentów doboru odpowiednich technik analizy danych w zależności od skali rozpatrywanego problemu oraz rodzaju przeprowadzanej analizy (w czasie rzeczywistym, w trybie batch, przetwarzanie strumieni danych).

Nauczenie studentów pracy z wykorzystaniem nowoczesnych platform analitycznych takich jak: Elastic Search, Apache Hadoop, Apache Spark, Apache Storm oraz SAS

Wymagania wstępne

Bazy danych,

Znajomość podstaw statystyki

Umiejętność programowania w języku Java

Zakres tematyczny

Definicja analityki biznesowej. Jej rola i zastosowanie we współczesnych systemach informatycznych.

Przegląd tradycyjnych pojęć i narzędzi analityki biznesowej. Jakość i czyszczenie danych.

Metody redukcji wymiaru: analiza składowych głównych i analiza czynnikowa.

Wnioskowanie statystyczne. Analiza regresji, współzależności i wariancji z punktu widzenia data mining.

Regresja logistyczna. Naiwna estymacja bayesowska i sieci bayesowskie. Analiza i prognozowanie szeregów czasowych.

Analiza danych nieustrukturyzowanych: analiza sentymentu, tworzenie ontologii, kategoryzacja treści, 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, regułyasocjacyjne i klasyfikacja.

Nowoczesne narzędzia do zaawansowanej wizualizacji oraz eksploracji danych na przykładzie platformy SAS.

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 przetwarzania Big Data na przykładzie platform Elasticsearch, Apache Hadoop, Apache Spark oraz Apache Storm.

Podstawowe techniki uczenia maszynowego wykorzystywane w analityce Big Data.

Metody kształcenia

Wykład wykład konwencjonalny, dyskusja, wykład problemowy, 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 do analizy Big Data	K_W09 K_W11	sprawdzian z progami punktowymi	Wykład
Student potrafi w praktyczny sposób wykorzystać narzędzia oferowane przez platformę SAS oraz Elastic Search w celu przeprowadzenia zaawansowanej analizy danych oraz ich eksploracji w czasie rzeczywistym.	K_W09 K_W11	projekt	Laboratorium
Student potrafi zdefiniować pojęcie analityki biznesowej oraz podać przykłady jej zastosowania.	K_W12	sprawdzian z progami punktowymi	Wykład
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 rozumie dlaczego dotychczasowe rozwiązania analityczne nie nadają się do analizy Big Data.	K_W12 K_U01	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 zna podstawowe techniki uczenia maszynowego wykorzystywane podczas analizy danych.	K_W08 K_W12 K_U01 K_U05	sprawdzian z progami punktowymi	Wykład
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

Morzy T.: Eksploracja danych. Metody i algorytmy, PWN, Warszawa, 2013

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

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

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

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

Literatura uzupełniająca

Stanton J.M.: Introduction to Data Science, E-book, 2013.

Uwagi

Zmodyfikowane przez prof. dr hab. inż. Krzysztof Patan (ostatnia modyfikacja: 22-09-2016 09:13)