Big data and business intelligence - opis przedmiotu

Informacje ogólne

Nazwa przedmiotu	Big data and business intelligence
Kod przedmiotu	11.3-WE-INFD-BDial-Er
Wydział	Wydział Informatyki, Elektrotechniki i Automatyki
Kierunek	WIEiA - oferta ERASMUS / Informatyka
Profil	-
Rodzaj studiów	Program Erasmus drugiego stopnia
Semestr rozpoczęcia	semestr zimowy 2018/2019

Informacje o przedmiocie

Semestr	1
Liczba punktów ECTS do zdobycia	6
Typ przedmiotu	obowiązkowy
Język nauczania	angielski
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	30	2	-	-	Zaliczenie na ocenę
Laboratorium	30	2	-	-	Zaliczenie na ocenę

Cel przedmiotu

- Familiarize students with the concept of business analytics and its applications to the analysis of large data sets included in social media, ERP systems and modern e-business applications.
- Teach students how to select the appropriate data analysis techniques depending on the scale of the problem under consideration and the type of analysis (real-time, batch mode, data stream processing).
- Teach students to work using modern analytical platforms such as Elastic Search, Apache Hadoop, Apache Spark, Apache Storm and SAS.

Wymagania wstępne

Databases, statistics essentials, programming in Java/Python

Zakres tematyczny

Definition of business analytics. Its role and application in modern IT systems.

Review of traditional business analytics concepts and tools. Data quality and cleaning.

Dimension reduction techniques: principal component analysis and factor analysis.

Statistical inference. Analysis of regression, interdependence, and dispersion from a data mining perspective. Logistic regression. Naive Bayesian estimation and Bayesian networks. Analysis and forecasting of time series.

Analysis of unstructured data: sentiment analysis, ontology creation, content categorization, text mining.

Web structure analysis: text search; ranking based on the connection structure.

Web usage analysis: preparatory data processing; exploratory network usage analysis; grouping, associative rules and classification.

Modern tools for advanced visualization and data mining on the SAS platform.

The Big Data phenomenon, its characteristics and impact on existing analytical solutions.

Large-scale business analytics; modern solutions used to transfer, store and process large data sets.

The architecture of modern Big Data processing systems based on Elasticsearch, Apache Hadoop, Apache Spark and Apache Storm.

Basic machine learning techniques used in Big Data analytics.

Metody kształcenia

Lecture, teaching laboratory classes

Efekty uczenia się i metody weryfikacji osiągania efektów uczenia się

Opis efektu	Symbole efektów	Metody weryfikacji	Forma zajęć
Can characterize the features of modern platforms for Big Data analytics		• test końcowy • final test	• Wykład
Is able to use the tools offered by SAS and Elastic Search in a practical way to make advanced data analysis and proceed with real-time exploration.		• bieżąca kontrola na zajęciach • wykonanie sprawozdań laboratoryjnych • on-going assessment of teaching laboratory activities, writing laboratory reports	• Laboratorium
Is able to use a selected platfom in a practical way to make a big data analysis.		• bieżąca kontrola na zajęciach • wykonanie sprawozdań laboratoryjnych • on-going assessment of teaching laboratory activities, writing laboratory reports	• Laboratorium
Can define the concept of business analytics and give examples of its usage.		• test końcowy • final test	• Wykład
Understands why traditional analytical solutions are not suitable for Big Data analytics.		• test końcowy • final test	• Wykład
Is able to describe the traditional data analysis model used in data and text mining.		• test końcowy • final test	• Wykład
Student knows basic machine learning techniques used in data analysis.		• test końcowy • final test	• Wykład
Can explain the concept of Big Data and present the main features that the relevant datasets contain.		• test końcowy • final test	• Wykład

Warunki zaliczenia

Lecture - the passing criterion is a sufficient mark from the final test.
Laboratory - the passing criterion are positive marks for all laboratory exercises.
Final mark components = lecture: 50% + teaching laboratory: 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

Literatura uzupełniająca

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

Uwagi

Zmodyfikowane przez prof. dr hab. inż. Dariusz Uciński (ostatnia modyfikacja: 27-03-2018 17:00)

Wygenerowano automatycznie z systemu SylabUZ