

Nazwa przedmiotu **Techniki eksploracji danych na platformie Java EE**

Nazwa w języku angielskim **Data processing and analysis on Java EE platform**

Język prowadzenia zajęć polski

Kierunek studiów Informatyka

Poziom studiów studia II stopnia magisterskie

Jednostka prowadząca Katedra Mikroelektroniki i Technik Informatycznych, K-25

<i>Kierownik i realizatorzy</i>	dr inż. Marek Kamiński	kaminski@dmcs.pl
	mgr inż. Robert Ritter	ritter@dmcs.pl

<i>Formy zajęć i liczba godzin w semestrze</i>	Wyk.	Ćw.	Lab.	Proj.	Sem.	Inne	Suma godzin w semestrze
	15	0	15	0	0	0	30

Efekty kształcenia

Po zakończeniu zajęć student będzie:

- Potrafił ocenić zasadność wykorzystania danej techniki przetwarzania oraz analizy danych dla aplikacji opartej na platformie Java EE
- Potrafił zaimplementować proste algorytmy drążenia danych w języku Java bez konieczności stosowania zewnętrznych bibliotek
- Potrafił przetwarzać i duże zbiory danych stosując narzędzia dedykowane dla platformy Java EE

Wymagania wstępne

Umiejętność programowania obiektowego w języku Java, podstawowa znajomość baz danych, znajomość bazy danych Apache Hadoop oraz SOLR

Organizacja przedmiotu i treści kształcenia

Wykład:

Podstawowe techniki analizy dużych zbiorów danych, klasyfikacja, grupowanie oraz asocjacja: sieci Bayesa, SVM, A-Priori, FP-Growth, kNN, SGD, Random Forests

Implementacja klasyfikacji w języku Java na przykładzie biblioteki Classifier4j

Przetwarzanie dużych zbiorów danych oraz analiza tekstu na platformie Java EE z zastosowaniem bazy danych Apache SOLR

Przetwarzanie oraz analiza dużych zbiorów danych z wykorzystaniem bazy danych Apache Hadoop, możliwości zrównoleglania działania algorytmów

Biblioteka Apache Mahout jako przykład rozwiązania do analizy zbiorów danych w środowisku rozproszonym: klasyfikacja, grupowanie, asocjacja

Ćwiczenia laboratoryjne:

- Implementacja prostych technik analizy zbiorów danych w języku Java
- Zastosowanie biblioteki Classifier4j w celu klasyfikacji danych.
- Analiza tekstu w aplikacji Java EE z wykorzystaniem Apache SOLR
- Analiza tekstu oraz danych w aplikacji Java EE w środowisku rozproszonym na przykładzie bazy danych Apache Hadoop
- Zastosowanie biblioteki Apache Mahout w celu przetwarzania dużych zbiorów danych: klasyfikacja, grupowanie oraz asocjacja

Forma zaliczenia - sprawdzenia osiągnięcia efektów kształcenia

ocena z kolokwium wykładowego oraz ocena postępów pracy w laboratorium

*Literatura
podstawowa*

Materiały wykładowe

*Literatura
uzupełniająca*

*Przeciętne
obciążenie studenta
pracą własną*

010

*Całkowite obciążenie
studenta pracą*

55

Uwagi

-

Aktualizacja

2011-10-17