**Nazwa przedmiotu:**

Eksploracja i przetwarzanie dużych zbiorów danych

**Koordynator przedmiotu:**

Anna Fiedukowicz, Robert Olszewski

**Status przedmiotu:**

Obowiązkowy

**Poziom kształcenia:**

Studia II stopnia

**Program:**

Geodesy and Cartography

**Grupa przedmiotów:**

Obowiązkowe

**Kod przedmiotu:**

1060-GK000-MSA-3000

**Semestr nominalny:**

3 / rok ak. 2020/2021

**Liczba punktów ECTS:**

3

**Liczba godzin pracy studenta związanych z osiągnięciem efektów uczenia się:**

90 godzin: w tym: wykład 15 h, ćwiczenia 15 h, studiowanie literatury: 20 godz., konsultacje 5h, samodzielne wykonanie projektów 35 h

**Liczba punktów ECTS na zajęciach wymagających bezpośredniego udziału nauczycieli akademickich:**

1 pkt. ECTS (35 godzin, w tym: wykład, ćwiczenia, konsultacje)

**Język prowadzenia zajęć:**

angielski

**Liczba punktów ECTS, którą student uzyskuje w ramach zajęć o charakterze praktycznym:**

2 pkt. ECTS w tym: ćwiczenia 15 h, studiowanie literatury: 20 godz., konsultacje 5h, samodzielne wykonanie projektów 35 h

**Formy zajęć i ich wymiar w semestrze:**

|  |  |
| --- | --- |
| Wykład: | 15h |
| Ćwiczenia: | 0h |
| Laboratorium: | 0h |
| Projekt: | 15h |
| Lekcje komputerowe: | 0h |

**Wymagania wstępne:**

Podstawowa wiedza i umiejętności w zakresie statystyki i technologii GIS

**Limit liczby studentów:**

100

**Cel przedmiotu:**

Celem kursu jest rozwinięcie umiejętności w zakresie data mining i ich zastosowań do analiz danych przestrzennych, jak również wizualizacji i interpretacji wyników Spatial Data Mining.
Kurs pozwala na zapoznanie się z wiedzą dotyczącą Big Data oraz baz danych typu NoSQL.

**Treści kształcenia:**

Statistics vs. Data Mining. Big data and spatial big data. NoSQL databases. Basic Data Mining Methods: Decision Trees (Regression and Classification Trees), Random Forest, Boosted Trees, Artificial Neural Networks, Multivariate Adaptive Regression Splines (MARS), Support Vector Machines (SVM), Association Rules, Naive Bayesian Classifier. Spatial Statistics and Geostatistics. Application of Data Mining Methods on Spatial Data. Cartographical Visualization of SDM results and their interpretation

**Metody oceny:**

Studenci są oceniani na podstawie projektów i zaliczenia wykładu (próg zaliczenia wynosi 50%). Jeżeli student jest nieobecny na zajęciach (maks. 20% zajęć), jest zobowiązany do uzupełnienia wiedzy.

**Egzamin:**

nie

**Literatura:**

Gorunescu, F. (2011). Data Mining: Concepts, models and techniques (Vol. 12). Springer Science &amp; Business Media.
Berman, J. J. (2013). Principles of big data: preparing, sharing, and analyzing complex information. Newnes.
Leung, Y. (2010). Knowledge discovery in spatial data (pp. 223-276). Berlin, Germany:: Springer.
Li, D., Wang, S., &amp; Li, D. (2015). Spatial data mining. Springer Berlin Heidelberg.

**Witryna www przedmiotu:**

brak

**Uwagi:**

brak

## Efekty przedmiotowe

### Profil ogólnoakademicki - wiedza

**Efekt :**

ma pogłębioną wiedzę na temat technologii Big Data i Spatial Big Data

Weryfikacja:

**Powiązane efekty kierunkowe:** K\_W14

**Powiązane efekty obszarowe:** T2A\_W01, T2A\_W02, T2A\_W05, T2A\_W07

### Profil ogólnoakademicki - umiejętności

**Efekt :**

potrafi zrealizować analizę Spatial Data Mining, zwizualizować jej wyniki na mapie i zinterpretować uzyskane wyniki

Weryfikacja:

**Powiązane efekty kierunkowe:** K\_U18

**Powiązane efekty obszarowe:** T2A\_U07, T2A\_U08, T2A\_U09, T2A\_U11

### Profil ogólnoakademicki - kompetencje społeczne

**Efekt :**

Potrafi działać kreatywnie i nieszablonowo podczas rozwiązywania problemów badawczych

Weryfikacja:

**Powiązane efekty kierunkowe:** K\_K01

**Powiązane efekty obszarowe:** T2A\_K06