**Nazwa przedmiotu:**

Metody optymalizacji w analizie danych

**Koordynator przedmiotu:**

prof. dr hab. inż. Radosław Pytlak

**Status przedmiotu:**

Obowiązkowy

**Poziom kształcenia:**

Studia II stopnia

**Program:**

Inżynieria i Analiza Danych

**Grupa przedmiotów:**

Wspólne

**Kod przedmiotu:**

.

**Semestr nominalny:**

3 / rok ak. 2019/2020

**Liczba punktów ECTS:**

6

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

1. godziny kontaktowe – 83 h; w tym
a) obecność na wykładach – 30 h
b) obecność na zajęciach projektowych – 45 h
c) obecność na egzaminie – 3 h
d) konsultacje – 5 h
2. praca własna studenta – 75 h; w tym
a) przygotowanie projektu – 40 h
b) przygotowanie do egzaminu – 25 h
Razem 158 h, co odpowiada 6 pkt. ECTS

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

a) obecność na wykładach – 30 h
b) obecność na zajęciach projektowych – 45 h
c) obecność na egzaminie – 3 h
d) konsultacje – 5 h
Razem 80 h, co odpowiada 3 pkt. ECTS

**Język prowadzenia zajęć:**

polski

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

a) przygotowanie projektu – 40 h
b) obecność na zajęciach projektowych – 45 h
Razem 85 h, co odpowiada 3 pkt. ECTS

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

|  |  |
| --- | --- |
| Wykład:  | 30h |
| Ćwiczenia:  | 0h |
| Laboratorium:  | 0h |
| Projekt:  | 45h |
| Lekcje komputerowe:  | 0h |

**Wymagania wstępne:**

Analiza matematyczna (rachunek różniczkowy)
Algebra liniowa (rachunek macierzowy)
Metody numeryczne
Ekonometria

**Limit liczby studentów:**

.

**Cel przedmiotu:**

Wprowadzenie do podstawowych metod obliczeniowych optymalizacji dla zadań z ograniczeniami i bez ograniczeń. Wprowadzenie do podstawowych metod optymalizacji całkowitoliczbowej. Przedstawienie metod obliczeniowych optymalizacji stosowanych do rozwiązywania zadań najmniejszych kwadratów: liniowych i nieliniowych. Omówienie metod obliczeniowych stosowanych do rozwiązywania zadań najmniejszych występujących w analizie danych. Omówienie zadania LASSO oraz metod wykorzystywanych do jego rozwiązania: metod proksymalnych opartych na technikach optymalizacji nieróżniczkowalnej; metod optymalizacji całkowitoliczbowej. Omówienie zastosowania obliczeń równoległych i rozproszonych do rozwiązywania zadań optymalizacji występujących w analizie danych. W ramach projektu zrealizowane zostaną: rozwiązywanie zadań estymacji parametrów modeli regresji poprzez rozwiązania zadania LASSO; wykorzystanie serwera NEOS, środowiska obliczeniowego IBM ILOG.

**Treści kształcenia:**

1. Wprowadzenie do podstawowych metod optymalizacji dla zadań bez ograniczeń.
2. Metoda Gaussa-Newtona i Levenberga-Marquardta rozwiązywania nieliniowego zadania najmniejszych kwadratów.
3. Zastosowanie metod dla zadania najmniejszych kwadratów do klasyfikacji danych oraz uczenia sieci neuronowych.
4. Zastosowanie przyrostowej metody gradientowej do rozwiązywania zadania najmniejszych kwadratów.
5. Filtr Kalmana w zastosowaniu do rozwiązywania zadania najmniejszych kwadratów.
6. Warunki optymalności dla zadań z ograniczeniami.
7. Programowanie wypukłe i dualność.
8. Metoda sympleks.
9. Metody punktu wewnętrznego.
10. Warunki konieczne optymalności dla zadań optymalizacji całkowitoliczbowej. Metoda podziału i ograniczeń.
11. Zadania optymalizacji nieróżniczkowalnej. Metody subgradientowe oraz metody proksymalne do rozwiązywania zadania estymacji parametrów zgodnie z podejściem LASSO.
12. Zastosowanie obliczeń równoległych i rozproszonych w rozwiązywaniu zadań optymalizacji metodami proksymalnymi.
13. Zastosowanie metod optymalizacji całkowitoliczbowej do realizacji podejścia LASSO..
14. Asynchroniczne metody pierwszego rzędu z decentralizowaną komunikacją.
15. Wykorzystanie środowisk obliczeniowych optymalizacji w analizie danych.

**Metody oceny:**

Ocena przedmiotu składa się z ocen cząstkowych: sprawozdanie z projektu – 50%; egzamin – 50%.

**Egzamin:**

tak

**Literatura:**

1. S. Boyd, L. Vandenberghe, Convex Optimization, Cambridge University Press, 2004
2. L.A. Wolsey. Integer Programming, J. Wiley & Sons, 1998.
3. D..P. Bertsekas, Nonlinear Programming, Athena Scientific, 1999.
4. J. Nocedal, S.J. Wright, Numerical Optimization, Springer, 2006.

**Witryna www przedmiotu:**

.

**Uwagi:**

program 4 semestralny - 3 semestr
program 3 semestralny - 2 semestr

## Charakterystyki przedmiotowe

### Profil ogólnoakademicki - wiedza

**Charakterystyka W01:**

Ma podstawową wiedzę z zakresu metod optymalizacji wykorzystywanych w analizie danych.

Weryfikacja:

Kolokwium

**Powiązane charakterystyki kierunkowe:** DS2\_W01, DS2\_W03, DS2\_W14

**Powiązane charakterystyki obszarowe:** I.P7S\_WG, I.P7S\_WK

**Charakterystyka W02:**

Ma wiedzę z zakresu metod numerycznych dla zadań optymalizacji występujących w analizie danych.

Weryfikacja:

Kolokwium

**Powiązane charakterystyki kierunkowe:** DS2\_W01, DS2\_W03, DS2\_W14

**Powiązane charakterystyki obszarowe:** I.P7S\_WG, I.P7S\_WK

### Profil ogólnoakademicki - umiejętności

**Charakterystyka U01:**

Potrafi rozwiązywać zadania optymalizacji w analizie danych z wykorzystaniem właściwego pakietu numerycznego.

Weryfikacja:

Sprawozdanie z projektu

**Powiązane charakterystyki kierunkowe:** DS2\_U03, DS2\_U18

**Powiązane charakterystyki obszarowe:** I.P7S\_UW, I.P7S\_UU

**Charakterystyka U02:**

Potrafi sformułować i rozwiązać zadanie optymalizacji analizy danych z wykorzystaniem języka modelowania optymalizacji.

Weryfikacja:

Sprawozdanie z projektu

**Powiązane charakterystyki kierunkowe:** DS2\_U17, DS2\_U02

**Powiązane charakterystyki obszarowe:** I.P7S\_UU, I.P7S\_UW

### Profil ogólnoakademicki - kompetencje społeczne

**Charakterystyka K01:**

Rozumie proces postępu w dziedzinie analizy danych i konieczność ciągłego samokształcenia.

Weryfikacja:

.

**Powiązane charakterystyki kierunkowe:** DS2\_K01

**Powiązane charakterystyki obszarowe:** I.P7S\_KK