**Nazwa przedmiotu:**

Optimisation in Data Analysis

**Koordynator przedmiotu:**

Prof. dr hab. Radosław Pytlak

**Status przedmiotu:**

Obowiązkowy

**Poziom kształcenia:**

Studia II stopnia

**Program:**

Data Science

**Grupa przedmiotów:**

Współny

**Kod przedmiotu:**

.

**Semestr nominalny:**

2 / rok ak. 2022/2023

**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) konsultacje – 5 h
d) obecność na egzaminie – 3 h
2. praca własna studenta – 85 h; w tym
a) zapoznanie się z literaturą – 15 h
b) przygotowanie do zajęć projektowych – 50 h
c) przygotowanie do egzaminu – 20 h
Razem 168 h, co odpowiada 6 pkt. ECTS

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

1. obecność na wykładach – 30 h
2. obecność na zajęciach projektowych – 45 h
3. konsultacje – 5 h
4. obecność na egzaminie – 3 h
Razem 83 h, co odpowiada 3 pkt. ECTS

**Język prowadzenia zajęć:**

angielski

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

1. obecność na zajęciach projektowych – 45 h
2. przygotowanie do zajęć projektowych – 50 h
Razem 95 h, co odpowiada 4 pkt. ECTS

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

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

**Wymagania wstępne:**

none

**Limit liczby studentów:**

Project classes – the number of students in a group matches the limits defined by the Warsaw University of Technology

**Cel przedmiotu:**

Introduction to convex functions, convex programming and duality theory. Introduction to computational methods for unconstrained and constrained optimization problems. Introduction to computational methods for optimization problems with simple constraints. Introduction to computational methods for mixed integer optimization problems. Presentation of computational methods for solving linear and nonlinear least squares problems. Presentation of optimization methods for solving least squares problems in data analysis: incremental gradient methods, Kalman filter, stochastic gradient methods. Introduction to SVM models used in binary classification. Introduction to computational methods for building SVM models: cutting plane methods, subgradient methods, optimization methods for problems with simple constraints, ‘coordinate descent’ methods. Discussion on LASSO models and computational methods for their building: proximal methods, mixed integer optimization methods. Introduction to parallel and distributed computational methods for solving optimization problems applied in data analysis. Student’s projects will concentrate on: building regression models based on LASSO approach; building SVM models for binary data classification; solving optimization problems with the help of computational environments implemented in R and Python.

**Treści kształcenia:**

Lecture:
1. Introduction to numerical methods for unconstrained optimization.
2. Gaussa-Newton and Levenberg-Marquardt methods for nonlinear least squares problems.
3. Nonlinear least squares methods applied to neural networks training.
4. Incremental gradient methods applied to nonlinear least squares problems.
5. Kalman filter and its application to least squares problems.
6. Optimality conditions in constrained optimization.
7. Convex programming and duality.
8. Interior point methods in cone programming.
9. Introduction to mixed integer optimization methods.
10. SVM approach to binary classification, loss function in primal and dual spaces.
11. Optimization methods for building SVM models.
12. Optimization methods for building linear regression models based on LASSO approach.
13. Subgradient proximal methods in data analysis.
14. Optimization methods with stochastic gradients.
15. Parallel and distributed optimization methods in data analysis.
Project classes:
Student carries out the project associated with the use of some optimization methods for building regression models based on the LASSO approach, or SVM classification models in primal, or dual spaces. The project includes the effectiveness analysis of applied optimization methods.

**Metody oceny:**

The final grade is based on two partial grades: 50% allocated to the project’s assessment; 50% allocated to oral exam.

**Egzamin:**

tak

**Literatura:**

1. S. Boyd, L. Vandenberghe, Convex Optimization, Cambridge University Press, 2004
2. Y. Nesterov, Lectures on convex optimization, Springer-Verlag, 2018.
3. D..P. Bertsekas, Nonlinear Programming, Athena Scientific, 1999..
4. J. Nocedal, S.J. Wright, Numerical Optimization, Springer, 2006.

**Witryna www przedmiotu:**

.

**Uwagi:**

.

## Charakterystyki przedmiotowe

### Profil ogólnoakademicki - wiedza

**Charakterystyka W01:**

He has knowledge about optimization methods applied in data analysis.

Weryfikacja:

Oral exam

**Powiązane charakterystyki kierunkowe:** DS2\_W01, DS2\_W02, DS2\_W03, DS2\_W10

**Powiązane charakterystyki obszarowe:**

**Charakterystyka W02:**

He has knowledge about numerical methods for optimization problems applied in data analysis.

Weryfikacja:

Oral exam

**Powiązane charakterystyki kierunkowe:** DS2\_W03, DS2\_W10, DS2\_W02

**Powiązane charakterystyki obszarowe:**

### Profil ogólnoakademicki - umiejętności

**Charakterystyka U01:**

He can solve optimization problem occurring in data analysis with the help of the appropriate numerical software.

Weryfikacja:

Assessment of the project’s report

**Powiązane charakterystyki kierunkowe:** DS2\_U02, DS2\_U03, DS2\_U04, DS2\_U11

**Powiązane charakterystyki obszarowe:**

**Charakterystyka U02:**

He can formulate and solve optimization problem occurring in data analysis with the help of optimization modelling language.

Weryfikacja:

Assessment of the project’s report

**Powiązane charakterystyki kierunkowe:** DS2\_U02, DS2\_U03, DS2\_U11

**Powiązane charakterystyki obszarowe:**

**Charakterystyka U03:**

He can implement optimization methods for solving problems occurring in data analysis.

Weryfikacja:

Assessment of the project’s report

**Powiązane charakterystyki kierunkowe:** DS2\_U01

**Powiązane charakterystyki obszarowe:**

**Charakterystyka U04:**

He can evaluate effectiveness analysis of applied optimization methods in data analysis.

Weryfikacja:

Assessment of the project’s report

**Powiązane charakterystyki kierunkowe:** DS2\_U12, DS2\_U13

**Powiązane charakterystyki obszarowe:**

### Profil ogólnoakademicki - kompetencje społeczne

**Charakterystyka K01:**

He is aware of the developments in the field of data analysis and associated with that the necessity of continuing education.

Weryfikacja:

Oral exam, assessment of the project’s report

**Powiązane charakterystyki kierunkowe:** DS2\_K01

**Powiązane charakterystyki obszarowe:**