COURSE OFFERED IN THE DOCTORAL SCHOOL

Code of the course	4606-ES-000BCHJ-0371	Name of the course	Polish	**Metody optymalizacji w analizie danych**
			English	**Optimization Methods in Data Science**
Type of the course	specialized			
Course coordinator	prof. dr hab. inż. Radosław Pytlak	Course teacher	prof dr hab. inż. R. Pytlak	
Implementing unit	Faculty of Mathematics and Information Science	Scientific discipline / disciplines*	information and communication technology; automation, electronics, electrical engineering and space technologies; mechanical engineering; mathematics	
Level of education	doctoral studies	Semester	winter	
Language of the course	English			

Type of assessment	grade	Number of hours in a semester	52	ECTS credits	4
Minimum number of participants	10	Maximum number of participants	45	Available for students (BSc, MSc)	Yes

Type of classes		Lecture	Auditory classes	Project classes	Laboratory	Seminar
Number of hours	in a week	2		2		
	in a semester	26		26		

* does not apply to the Researcher's Workshop

1. Prerequisites

basic knowledge of mathematical analysis and linear algebra

2. Course objectives

The main goal of the course is to present optimization methods used in machine learning. The course will review computational methods which are suitable for solving optimization problems which are related to building models based on data. Both classical methods such as steepest descent, conjugate gradient, Newton and Gauss-Newton methods will be discussed with respect to their application to optimization problems which are large in their dimension and the complexity of their objective functions requires using stochastic approximations to their gradients, and specialized optimization methods developed for solving problems in machine learning (such as Kalman filter, coordinate, or mirror descent methods).

Necessary and sufficient conditions for unconstrained optimization problems. Convex functions, necessary and sufficient conditions for the convexity of functions.

Introduction to basic computational optimization methods for unconstrained problems: steepest descent, conjugate gradient method, limited-memory variable metric method, and globally convergent Newton's method using conjugate gradient to determine Newton's directions. Optimization methods for problems with simple constraints.

Presentation of computational optimization methods used to solve least squares problems: linear and nonlinear. Discussion of computational methods used to solve least squares problems occurring in data analysis: 'incremental' gradient methods, computational methods based on the Kalman filter.

Global convergence of the steepest descent method with a constant step. Overview of first-order accelerated methods, the Polyak–Heavy Ball method and its rate of convergence. Nesterov's first-order accelerated method with gradient prediction. Nesterov's accelerated method based on sequences of estimating functions. The impact of inaccurate gradient evaluation on the convergence of gradient methods. Application of the Lyapunov method to analyze the convergence of optimization algorithms. The impact of inaccurate gradient determination on the convergence rate of gradient methods. The 'coordinate descent' method, global convergence conditions and its convergence rate. Optimization methods with stochastic gradient approximation. Stochastic optimization problem, global convergence conditions of the subgradient method for a stochastic optimization problem. Constrained optimization problem, necessary and sufficient conditions for optimality. Dual problem. Subgradient method for solving dual problem. Methods for building classification models using the SVM approach. The primal and dual problems used to build the SVM model. Building a linear regression model using l_0 regularization. First-order optimization method and mixed integer optimization method for building a linear regression model based on l_0 regularization.

3. Course content (separate for each type of classes)
Lecture
1. Necessary and sufficient conditions for unconstrained optimization problems. Convex functions, necessary conditions for the convexity of functions. 2. Basic optimization methods for unconstrained problems, methods with line search, trust region methods. Conjugate gradient methods, variable metric methods. Rates of convergence of basic optimization methods. 3. Globally convergent Newton method using the conjugate gradient method to determine the Newton direction. Rates of convergence of the Newton method. Optimization methods for problems with simple constraints. 4. Gauss–Newton method for solving nonlinear least squares problems. Incremental gradient method for solving nonlinear least squares problems. Convergence conditions for the incremental gradient method. 5. Gauss–Newton version of the incremental gradient method. Kalman filter for linear least squares problems. Extended Kalman filter for nonlinear least squares problems. 6. Rate of convergence of the steepest descent method with a constant step. Accelerated first-order Polyak method. Rate of convergence of the accelerated first-order method. 7. Analysis of the rate of convergence of the steepest descent method in the space of objective function values. Accelerated first-order Nesterov method and its rate of convergence for objective function values. Accelerated first-order Nesterov method based on sequences of estimating functions. 8. The effect of imprecise gradients on the convergence and rates of convergence of first-order optimization methods. Ap-plication of the Lyapunov function to analyze the convergence of optimization methods. 9. The 'coordinate descent' method: conditions for the global convergence of the method and its rate of convergence. First-order optimization methods with stochastic gradient approximation. Globally convergent stochastic subgradient methods for stochastic optimization problems. 10. Necessary and sufficient conditions for optimality for constrained problems. Dual problem, subgradient method for solving the dual problem. 11. Primal and dual problems for building a classification SVM model. Optimization methods for solving the primal and dual problems. 12. First-order method for building a linear regression model with l_0 regularization. A method based on solving a mixed integer optimization problem for building a linear regression model with l_0 regularization. 13. Mirror methods for solving stochastic optimization problems with simple constraints.

.
Project
The student completes a project involving the application of selected optimization methods to solve a selected data analysis problem. The student implements optimization methods described in a scientific publication in the field of optimization or machine learning. The project includes an analysis of the effectiveness of the optimization methods used, in particular their comparison to other optimization methods used to solve the selected problem.

4. Learning outcomes

Type of learning outcomes	Learning outcomes description	Reference to the learning outcomes of the WUT DS	Learning outcomes verification methods*
Knowledge			
W01	He has knowledge about optimization methods applied in data analysis.	SD_W2	report evaluation
W02	He has knowledge about numerical methods for optimization problems applied in data analysis.	SD_W3	report evaluation
W03			
Skills			
U01	He can solve optimization problem occurring in data analysis with the help of the appropriate numerical software.	SD_U1	report evaluation
U02	He can formulate and solve optimization problem occurring in data analysis with the help of optimization modelling language.	SD_U2, SD_U4	report evaluation
U03	He can implement optimization methods for solving problems occurring in data analysis.	SD_U2, SD_U4	report evaluation
U04	He can evaluate effectiveness analysis of applied optimization methods in data analysis.	SD_U2, SD_U4	report evaluation
Social competences			
K01	He is aware of the developments in the field of data analysis and associated with that the necessity of continuing education.	SD_K2	
K02			
K03			

*Allowed learning outcomes verification methods: exam; oral exam; written test; oral test; project evaluation; report evaluation; presentation evaluation; active participation during classes; homework; tests

5. Assessment criteria

The final grade is entirely based on the project's assessment.

6. Literature

Primary references:

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.

Secondary references:

1.L. A. Wolsey, Integer Programming (Second Edition), John Wiley & Sons, 2021.

7. PhD student's workload necessary to achieve the learning outcomes**		
No.	Description	Number of hours
1	Hours of scheduled instruction given by the academic teacher in the classroom	52
2	Hours of consultations with the academic teacher, exams, tests, etc.	5
3	Amount of time devoted to the preparation for classes, preparation of presentations, reports, projects, homework	37
4	Amount of time devoted to the preparation for exams, test, assessments	10
	Total number of hours	**104**
	ECTS credits	**4**

** 1 ECTS = 25-30 hours of the PhD students work (2 ECTS = 60 hours; 4 ECTS = 110 hours, etc.)

8. Additional information	
Number of ECTS credits for classes requiring direct participation of academic teachers	2
Number of ECTS credits earned by a student in a practical course	2