POZNAN UNIVERSITY OF TECHNOLOGY

EUROPEAN CREDIT TRANSFER AND ACCUMULATION SYSTEM (ECTS)

COURSE DESCRIPTION CARD - SYLLABUS

Course name
Machine Learning Theory [S2SI1E>TUM]

Course

Field of study
Artificial Intelligence

Year/Semester
1/1

Area of study (specialization)
–

Profile of study
general academic

Level of study
second-cycle

Course offered in
English

Form of study
full-time

Requirements
compulsory

Number of hours

Lecture
15

Laboratory classes
0

Other
0

Tutorials
15

Projects/seminars
0

Number of credit points

3,00

Coordinators

dr hab. inż. Wojciech Kotłowski prof. PP
wojciech.kotlowski@put.poznan.pl

Lecturers

dr hab. inż. Wojciech Kotłowski prof. PP
wojciech.kotlowski@put.poznan.pl

Prerequisites

The student at the beginning of the course should have a basic knowledge of the probability calculus (axioms and properties of the probability measure, discrete and continuous random variables, moments of random variables, multidimensional random variables, probabilistic inequalities: Markov and Chebyshev), mathematical statistics (problems of parameter estimation, regression), and machine learning (overfitting, validation of learning systems, linear models, boosting, neural networks) and the ability to solve basic problems in these areas. In terms of social competences, the student must understand the importance of using the latest knowledge in the field of computer science in solving research problems, as well as present attitudes such as honesty, responsibility, perseverance, cognitive curiosity, creativity, personal culture, respect for other people.

Course objective

The aim of the course is to familiarize the students with the most important results in the field of machine learning theory. Lectures focus on discussing the basics of statistical learning theory (formulation of the learning problem, elements of statistical decision theory, minimizing the empirical risk, generalization theory, bias /variance decomposition) and the online theory framework (predictions with expert advice, online convex optimization).

Course-related learning outcomes

Knowledge 1. Has a structured and theoretically founded general knowledge related to key issues in the field of machine learning [K2st_W2]
2. Has advanced detailed knowledge regarding the fundamentals of machine learning [K2st_W3]
3. Has knowledge about development trends and the most important cutting edge achievements in computer science and statistics [K2st_W4]
4. Knows advanced methods, techniques and tools used to solve complex engineering tasks and conduct research in the theory of machine learning [K2st_W6]
Skills 1. Is able to obtain information from literature, databases and other sources (both in Polish and English), integrate them, interpret and critically evaluate them, draw conclusions and formulate and fully justify opinions [K2st_U1]
2. Can use analytical, simulation and experimental methods to formulate and solve machine learning theory problems [K2st_U4]
3. Can - when formulating and solving machine learning tasks - integrate knowledge from different areas of computer science (and if necessary also knowledge from other scientific disciplines) and apply a systemic approach, also taking into account non-technical aspects [K2st_U5]
4. Is able to assess the suitability and the possibility of using new achievements (methods and tools) in machine learning theory [K2st_U6]
5. Is able - using among others conceptually new methods - to solve complex tasks in machine learning theory [K2st_U10]
6. Can determine the directions of further learning and implement the process of self-education, including other people [K2st_U16]
Social competences 1. Understands that in the field of IT the knowledge and skills quickly become obsolete [K2st_K1]
2. Understands the importance of using the latest knowledge in the field of computer science in solving research and practical problems [K2st_K2]

Methods for verifying learning outcomes and assessment criteria

Learning outcomes presented above are verified as follows:

Learning outcomes presented above are verified as follows:
1. Lectures:
- assessment of knowledge and skills on a final written test containing open exercises and / or multiple-choice questions;
- discussion of the test results.
2. Tutorial sessions:
- continuous assessment, at each class in the form of short test or open questions,
- obtaining additional points for activity during exercises,
- obtaining additional points by discussing and presenting scientific articles;
For both lectures and tutorials, the following grading scale is used: over 50% - satisfactory, 60% - sufficient plus, 70% - good, 80% - good plus, 90% - very good.

Programme content

1. Formal presentation of the learning problem: statistical model of the learning problem, loss function, risk, Bayesian classifier, elements of statistical decision theory, basic problem of learning from data.
2. Empirical risk, empirical risk minimization, generalization error, estimation and approximation error, No-Free-Lunch theorem, basic probabilistic inequalities (Markov inequality, Chebyshev inequality, union bound, Hoeffding inequality), derivation of the bound on the generalization error of the finite function classes, PAC model.
3. Uniform convergence within the class of predictive functions, Rademacher complexity, growth function.
4. Vapnik-Chervonenkis (VC) dimension and fundamental learning theorem, VC dimension for popular function classes.
5. Linear classification, SVM methods, boosting methods, surrogate convex loss functions.
6. Online learning, regret minimization, the problem of prediction with expert advice, Follow-the-Leader and Hedge algorithms, regret bounds,, optimal algorithms.
7. Online convex optimization, Stochastic Gradient Descent (SGD) algorithm, regret analysis for SGD algorithm.

Course topics

1. Formal presentation of the learning problem: statistical model of the learning problem, loss function, risk, Bayesian classifier, elements of statistical decision theory, basic problem of learning from data.
2. Empirical risk, empirical risk minimization, generalization error, estimation and approximation error, No-Free-Lunch theorem, basic probabilistic inequalities (Markov inequality, Chebyshev inequality, union bound, Hoeffding inequality), derivation of the bound on the generalization error of the finite function classes, PAC model.
3. Uniform convergence within the class of predictive functions, Rademacher complexity, growth function.
4. Vapnik-Chervonenkis (VC) dimension and fundamental learning theorem, VC dimension for popular function classes.
5. Linear classification, SVM methods, boosting methods, surrogate convex loss functions.
6. Online learning, regret minimization, the problem of prediction with expert advice, Follow-the-Leader and Hedge algorithms, regret bounds,, optimal algorithms.
7. Online convex optimization, Stochastic Gradient Descent (SGD) algorithm, regret analysis for SGD algorithm.

Teaching methods

1. Lecture: multimedia presentation, illustrated with examples given on the blackboard, practical exercises (including calculation on the blackboard).
2. Tutorials: solving tasks and problems related to the content discussed in the lecture.

Bibliography

Basic
1. S. Shalev-Shwartz and S. Ben-David. Understanding Machine Learning: From Theory to Algorithms. Cambridge University Press, 2014
2. Yaser S. Abu-Mostafa, Malik Magdon-Ismail, Hsuan-Tien Lin: Learning From Data. AMLBook, 2012.
3. Mehryar Mohri, Afshin Rostamizadeh, Ameet Talwalkar: Foundations of Machine Learning, MIT Press, 2012.
4. Elad Hazan: Introduction to Online Convex Optimization. Foundations and Trends® in Optimization, Vol. 2, no. 3-4, pp 157-325.

Additional
1. O. Bousquet, S. Boucheron, and G. Lugosi: Introduction to statistical learning theory. Advanced Lectures on Machine Learning, pp. 169-207. Springer Berlin Heidelberg, 2004.
2. L. Devroye, L. Gyorfi, and G. Lugosi: A Probabilistic Theory of Pattern Recognition. Springer, 1996.
3. M. Anthony and P.L. Bartlett, Neural Network Learning: Theoretical Foundations. Cambridge University Press, 1999.
4. V.N. Vapnik: Statistical Learning Theory. Wiley-Interscience, 1998.
5. T. Hastie, R. Tibschirani, J. Friedman: Elements of Statistical Learning. Springer, 2017.
5. N. Cesa-Bianchi and G. Lugosi: Prediction Learning and Games. Cambridge University Press, 2006.
6. M. Kempka, W. Kotłowski, M. K. Warmuth: Adaptive scale-invariant online algorithms for learning linear models. International Conference on Machine Learning (ICML), 2019

Breakdown of average student's workload

	Hours	ECTS
Total workload	75	3,00
Classes requiring direct contact with the teacher	30	1,50
Student's own work (literature studies, preparation for laboratory classes/ tutorials, preparation for tests/exam, project preparation)	45	1,50