Course syllabus

Course title	Introduction to machine learning
Instructor(s)	dr Julian Zubek
Contact details	j.zubek@uw.edu.pl
Affiliation	Faculty of Psychology, University of Warsaw
Course format	lecture, class
Number of hours	60 hours
Number of ECTS credits	**6 ECTS credits**
Brief course description	This course provides an overview of machine learning concepts and algorithms. It focuses mostly on techniques related to classification and regression, such as nearest neighbors methods, generalized linear models, tree-based methods, feed-forward neural networks. Simple clustering techniques (k-means clustering, hierarchical clustering) are also introduced. Lecture covers main principles behind different algorithms, model evaluation strategies and basics of statistical learning theory. Connections with topics known from cognitive modeling (e.g., categorization models, signal detection theory) or statistics (e.g., sampling, probability density estimation, logistic regression) are made. During laboratory classes students learn practical applications of the introduced methods using libraries from Python ecosystem (scikit-learn, XGBoost, PyTorch).
Full course description	Machine learning is a somewhat eclectic field drawing from artificial intelligence, statistics, optimization and data mining. It is pragmatically oriented as it focuses on solving practical problems using various heuristics rather than looking for ‘correct’ solutions. The center of interest is building predictive models based on available data. Machine learning provides general schemes for framing problems (e.g., classification, regression, clustering) and a set of diverse techniques for handling them. The course is designed to provide an overview of machine learning techniques and foster a specific way of thinking about problems. In accordance with a motto TIMTOWTDI (There is more than one way to do it), multiple ways of approaching similar tasks will be discussed. Material covered will seek a balance between simple algorithms illustrating particular concepts, models interesting from cognitive science perspective, and modern techniques with practical applications. Lectures will cover necessary theoretical background, including key concepts and some mathematical details. Classes in computer laboratory

	will focus on applications of the introduced algorithms. Generally, ready-made software packages will be used, students will not be required to implement algorithms on their own.
Learning outcomes	Student knows and understands: - basic concepts of machine learning, popular techniques and software libraries (K_W04, K_W08) - role of machine learning in artificial intelligence, cognitive processes modeling, experimental data analysis (K_W01, K_W02) Student is able to: - discuss modeling assumption behind particular techniques (K_U01) - design and implement a procedure of machine learning experiment using existing data (K_U02, K_U03, K_U04, K_U05) - build her understanding of machine learning on her own through reading scientific literature (K_U08, K_K01, K_K02)
Learning activities and teaching methods	Lecture will have a standard format with some degree of interactivity. During class in computer laboratory will experiment with machine learning algorithms on their own. Short homework assignments should be expected.
List of topics/classes and bibliography	1. What is machine learning? Standard machine learning tasks. Different classifier families at a glance. 2. Linear algebra review. Linear regression. Least squares approximation. 3. Working with data. Probabilistic interpretation of classification task. Bayes error. Naive Bayes classifier. Logistic regression. 4. Model complexity and overfitting. Strategies for model evaluation. Signal detection theory. 5. Regularizing models. Lasso and ridge regression. Support vector machines. 6. K-nearest neighbors classifier, kernel density estimation and their relations to categorization models. 7. Decision trees and ensemble methods. Bias-variance error decomposition. 8. Neural networks: representation and learning. Backpropagation of errors. 9. Recurrent neural networks for modeling dynamical processes. 10. Case studies: prediction from structured and unstructured data. 11. Statistical learning theory for supervised learning. What is learnable? 12. Unsupervised learning: clustering. 13. Unsupervised learning: dimensionality reduction. Principal component analysis (PCA) and factor analysis. Applications in psychometry. 14. Unsupervised learning: signal processing. Fourier transform, wavelet transform. Independent component analysis (ICA).

	15. Introduction to reinforcement learning. Modeling operant conditioning. 16. Reinforcement learning in continuous spaces. Autonomous control. Literature: ● Trevor Hastie, Robert Tibshirani and Jerome Friedman (2009). The Elements of Statistical Learning (2nd edition) https://web.stanford.edu/~hastie/ElemStatLearn/ ● Gareth James, Daniela Witten, Trevor Hastie and Robert Tibshirani (2013) An Introduction to Statistical Learning with Applications in R http://www-bcf.usc.edu/~gareth/ISL/index.html ● Ian Goodfellow, Yoshua Bengio and Aaron Courville (2016). Deep Learning https://www.deeplearningbook.org/ ● Michael Nielsen (2015), Neural networks and deep learning http://neuralnetworksanddeeplearning.com/ ● Richard Sutton, Andrew Barto (2018). Reinforcement Learning: An Introduction http://incompleteideas.net/book/the-book-2nd.html
Assessment methods and criteria	**Written exam (70%)** Exam will cover theory presented during lectures. **Assignments (30%)** Small assignments connected with topics of individual classes will be given for students to solve at home individually.
Attendance rules	Attendance at lectures is not obligatory, although it is strongly recommended. Attendance at classes in computer laboratory is obligatory, two unexcused absences are allowed in the semester.
Prerequisites	Good grasp of Python ("Advanced Python for cognitive scientists" course), basic probability theory and statistics.
Academic honesty	Students must respect the principles of academic integrity. Cheating and plagiarism (including copying work from other students, internet or other sources) are serious violations that are punishable and instructors are required to report all cases to the administration.
Remarks	