Politechnika Krakowska
Wydział Informatyki
i Matematyki

Course title	Robustness of machine learning models considering adversarial conditions (Master)
Institute/Division	Faculty of Computer Science and Mathematics/ Department of Computer Science
Course code	F-1.R_ML
Erasmus subject code*	11.4
Number of contact hours**	45 lecture hours (45h)
Course duration	1 semester (Fall or Spring)
ECTS credits	6
Course description (max 100 words)	Rapid machine learning (ML) development has solved many complex problems in various fields. However, ML models are subject to certain types of logical weaknesses due to the inherent limitations of the learning algorithms. Therefore, specific testing techniques and the consideration of additional thread models are required when deciding to use ML. It is essential in areas where decision model failure costs are high. A new branch of machine learning called Adversarial Machine Learning (AML) studies attacks on machine learning algorithms and defences against such attacks. AML techniques enable measuring the model's resistance to adversarial user behaviour (so-called adversarial robustness). The students will learn how to simulate the attack and monitor the performance of the ML model.
Literature	1. ETSI 5G PoC Consortium Steering Committee and Contributors, Artificial Intelligence (AI) in Test Systems, Testing AI Models and ETSI GANA Model's Cognitive Decision Elements (DEs) via a Generic Test Framework for Testing GANA Multi-Layer Autonomics & their AI Algorithms for Closed-Loop Network Automation, 2020, online: https://intwiki.etsi.org/images/ETSI_5G_PoC_White_Paper_No_5.pdf 2. Borovicka, Tomas, et al. "Selecting representative data sets." Advances in data mining knowledge discovery and applications 12 (2012): 43-70. 3. Laskov, Pavel, and Richard Lippmann. "Machine learning in adversarial environments." Machine learning 81 (2010): 115-119.
Course type/organization	• Lectures (15h) • Projects (30h)
Assessment method	Attending lectures and completing the practical projects with the reports.
Prerequisites	• Backgrounds in machine learning, • Advanced practical knowledge of Python, Java
Primary target group	Bachelor degree in computer sciences telecommunication or a similar discipline
Contact person	Joanna Kołodziej (PhD, DsC, Prof.PK)
Remarks	N/A

*please insert one of the following codes:
11.0 Mathematics, Informatics
11.1 Mathematics
11.2 Statistics
11.3 Informatics, Computer Science
11.4 Artificial Intelligence
11.5 Actuarial Science
11.9 Others Mathematics, Informatics

**1 lecture hour=45 minutes