GDAŃSK UNIVERSITY OF TECHNOLOGY

Subject card

Subject name and code	Ethics in Machine Learning, PG_00064485		
Field of study	Informatics		
Date of commencement of studies	February 2025	Academic year of realisation of subject	2024/2025
Education level	second-cycle studies	Subject group	Optional subject group Specialty subject group Humanistic-social subject group
Mode of study	Full-time studies	Mode of delivery	at the university
Year of study	1	Language of instruction	Polish
Semester of study	1	ECTS credits	2.0
Learning profile	general academic profile	Assessment form	assessment
Conducting unit	Department of Multimedia Systems -> Faculty of Electronics, Telecommunications and Informatics		
Name and surname of lecturer (lecturers)	Subject supervisor	prof. dr hab. inż. Andrzej Czyżewski	
	Teachers	prof. dr hab. inż. Andrzej Czyżewski	

Lesson types and methods of instruction	Lesson type	Lecture	Tutorial	Laboratory	Project	Seminar	SUM
	Number of study hours	15.0	0.0	0.0	0.0	15.0	30
	E-learning hours included: 0.0						

Learning activity and number of study hours	Learning activity	Participation in didactic classes included in study plan	Participation in consultation hours	Self-study	SUM
	Number of study hours	30	4.0	16.0	50

Subject objectives	The goal of the course is to familiarize students with selected ethical principles, such as addressing the problems associated with responsible approaches to learning and to applications of trained decision systems. In addition, the course aims to draw students' attention to such issues as: general ethical principles in IT projects, principles of legal registration and use of sensitive data, privacy, responsible data management issues and ethical use of legally protected intellectual property.

Learning outcomes

Course outcome	Subject outcome	Method of verification
[K7_W11] knows and understands, to an increased extent, the general principles of creation and development of forms of individual entrepreneurship and the economic, legal and other conditions of various types of activities related to the awarded qualification, including the principles of protection of industrial property and copyright law	Student knows the principles of data protection personal information. Knows the basic concepts and principles of protection of intellectual property intellectual property and copyright law.	[SW1] Assessment of factual knowledge [SW2] Assessment of knowledge contained in presentation
[K7_K71] is able to explain the need to apply knowledge from humanistic, social, economic or legal sciences in order to function in a social environment	Works and acts responsibly, considering the possible consequences of his/her actions. Is open to new theories, ideas and attitudes which he/she seeks to learn and understand. He is ready, under their influence. Recognizes and is able to indicate standard procedures of solving dilemmas connected with his/her profession, including his/her familiarity with them of his or her profession, including his or her familiarity with the legal regulations applicable to his or her professional work.	[SK1] Assessment of group work skills [SK4] Assessment of communication skills, including language correctness
[K7_U71] is able to apply knowledge from humanistic, social, economic or legal sciences in order to solve problems	He understands and accepts diversity of attitudes and ethical norms represented by representatives of different environments, cultures and professions. Is aware of his/her own autonomy, which he/she develops and is able to defend without infringing on the autonomy of others.	[SU3] Assessment of ability to use knowledge gained from the subject [SU2] Assessment of ability to analyse information
[K7_W08] knows and understands, to an increased extent, the fundamental dilemmas of modern civilisation, the main development trends of scientific disciplines relevant to the field of education	Has a basic knowledge of the presence of ideas of ethical values and moral norms in the spheres of technical sciences, artificial intelligence, design of solutions and systems affecting the lives of individuals and society, understands the relationship between these spheres.	[SW2] Assessment of knowledge contained in presentation [SW1] Assessment of factual knowledge

Subject contents

Ethics in computer science and engineering projects. Types of licenses in public datasets. Recording sensitive personal data (image, voice) in light of RODO, "Digital inequalities". Unbalanced datasets (dataset bias, methods for dealing with class imbalance, long-tail recognition, "minority" class recognition).Security in AI (uncertainty estimation methods, spurious detections, robustness model). Privacy in AI. Algorithms that provide "privacy" for users providing data (differential privacy, federated learning).Accountability in supervised learning, unsupervised learning, and applications of pre-trained models. Explainability and interpretability of artificial intelligence decisions. Trust in artificial intelligence (trustworthy AI). Responsible data management: data integrity, authenticity, and trustworthiness. Respecting, understanding and assimilating norms in machine learning (theoretical, behavioral and hybrid models).

Prerequisites and co-requisites

Knowledge of building information systems, programming principle, basic machine learning methods.

Assessment methods and criteria

Subject passing criteria	Passing threshold	Percentage of the final grade
Presentation at seminar	100.0%	50.0%
Attendance at lecture	80.0%	50.0%

Recommended reading	Basic literature	L. Royakkers, J. Timmer, L. Kool, and R. van Est, Societal and ethical issues of digitization, Ethics Inf. Technol., vol. 20, no. 2, pp. 127 142, 2018Easton- Calabria and W. L. Allen, Developing ethical approaches to data and civil society: from availability to accessibility, Innovation , vol. 28(1), pp. 52 62, 2015 J. Lodge, The dark side of the moon: Accountability, ethics and new biometrics, in Second generation biometrics: Theethical, legal and social context, Springer, 2012, pp. 305328D. Goroff, J. Polonetsky, and O. Tene, Privacy Protective Research: Facilitating Ethically Responsible Access to Administrative Data, Ann. Am. Acad. Pol. Soc. Sci., vol. 675(1), pp. 4666, 2018. T. Meek, H. Barham, N. Beltaif, A. Kaadoor, and T. Akhter, Managing the ethical and risk implications of rapid advances in artificial intelligence: A literature review, in 2016 PICMET, 2016, pp. 682693.D. Helbing, Societal, Economic, Ethical and Legal Challenges of the Digital Revolution: From Big Data to Deep Learning, Artificial Intelligence, and Manipulative Technologies, in Towards Digital Enlightenment, Springer, 2019, pp. 4772
	Supplementary literature	N. Dorasamy and N. Pomazalová, Social Impact and Social Media Analysis Relating to Big Data, in Data Science and Big Data Computing, Cham: Springer, 2016, pp. 293313. M. Steinmann et al., Embedding Privacy and Ethical Values in Big Data Technology, in Transparency in Social Media, Cham: Springer, 2015, pp. 277301 P. Prinsloo and S. Slade, Big Data, Higher Education and Learning Analytics: Beyond Justice, Towards an Ethics of Care, in Big Data and Learning Analytics in Higher Education, Cham: Springer, 2017, pp. 109124.
	eResources addresses	Adresy na platformie eNauczanie:
Example issues/ example questions/ tasks being completed	The assessment will include answers to questions posed to the student after his/her presentation at the seminar and the student's activity during lecture discussions. The detailed range of questions depends on the topic of lectures and seminar presentations.	
Work placement	Not applicable	

Document generated electronically. Does not require a seal or signature.