COURSE DESCRIPTION CARD - SYLLABUS

Course name
Ethics and research

Course

Field of study
Artificial Intelligence

Year/Semester
3/6

Area of study (specialization)

Profile of study
general academic

Level of study
First-cycle studies

Course offered in
English

Form of study
full-time

Requirements
compulsory

Number of hours

Lecture
15

Laboratory classes

Other (e.g. online)
0

Tutorials
0

Projects/seminars
15

Number of credit points

2

Lecturers

Responsible for the course/lecturer:
prof. dr hab. inż. Jerzy Stefanowski

e-mail: jerzy.stefanowski@cs.put.poznan.pl

tel: +48 61 665 2933

Faculty of Computing and Telecommunications
Piotrowo 2, 60-965 Poznań

Responsible for the course/lecturer:

Prerequisites

Prior completion of courses on Artificial Intelligence, Machine Learning, Data Mining, Robotics Fundamentals, Neural Networks along with deep learning. Basic programming skills in Python or Java. Good knowledge of the English language - reading literature with understanding and ability to communicate in a written form and to give a seminar presentation.

Course objective

The purpose of the course is to familiarize students with the ethical aspects associated with the development of artificial intelligence and the basic principles of conducting research projects. In particular, the possible dangers and risks associated with the uncontrolled development of advanced artificial intelligence approaches are presented; the concepts and postulates of so-called trustworthy

artificial intelligence; the issues of data biases and difficulty for learning systems, the impartiality of decision-making to humans, and the basics of methods of explaining intelligent systems. In addition, students will learn the basic concepts associated with the type of research projects, the characteristics of scientific knowledge, the implementation of the standard research process and the ethical aspects associated with it.

Course-related learning outcomes

Knowledge

K1st_W4: knows and understands the basic techniques, methods, algorithms, and tools used for solving computer problems as well as problems in artificial intelligence, including clustering, classification, optimization, and decision support

K1st_W9:knows ethical issues related to the creation and use of computers and, in particular, AI-based systems

Skills

K1st_U1: is able to collect information from the appropriate sources of different natures, perform its critical analysis, interpretation, and synthesis as well as comprehensively justify the formulated opinions

K1st_U4: can efficiently plan and carry out experiments, including computer measurements and simulations, interpret the obtained results and draw conclusions based on the experimental outcomes in the context of data mining problems, machine learning tasks, and decision problems requiring finding an optimal solution or the most preferred subset of alternatives

K1st_U10: can retrieve, analyze and transform different types of data, and carry out data synthesis to knowledge and conclusions useful for solving a variety of decision problems

K1st_U16: can plan and carry out life-long learning, and is aware of the possibilities of MSc studies

Social competences

K1st_K2: is aware of the importance of scientific knowledge and research related to AI in solving practical problems which are essential for the functioning of individuals, firms, organizations as well as the entire society within such example application fields as transport, healthcare, education, home/service robots, public safety, and entertainment

K1st_K3: knows the examples of poorly functioning AI systems, which led to the economic, social, or environmental losses

K1st_K7: is ready to act responsibly in professional life, encourage and promote the suitable behavior patterns, correctly identify and solve the dilemmas related to the work of a computer scientist - a specialist in the field of AI

Methods for verifying learning outcomes and assessment criteria

Learning outcomes presented above are verified as follows:

Lecture: Assessment test is conducted at the last lecture. The students need to answer to several (open

or closed) questions concerning the subjects presented during all lectures. Furthermore the discussion activity of students during lectures is taken into account.

Seminar classes: Students complete homework assignments, in addition, they prepare one presentation during class, which concerns a report on their work with selected research texts. They should also participate in the discussions held in class. The final grade is calculated as the average of the partial grades awarded for each element.

Programme content

An introduction to basic concepts in ethics. Presentation of currently identified risks associated with the applications of Artificial Intelligence (AI) methods, including, in particular, advanced learning systems, robotics and automatic decision-making.

Proposals for so-called responsible artificial intelligence and requirements for trustworthy AI. Analysis of recommendations from the ethical recommendations developed by the high-level expert group on artificial intelligence at the EC . Difficulties of their practical implementation.

An analysis of the quality and complexity distributions of learning examples for ML systems. Possible sources of imperfect data. The concept of data difficulty and complexity. Measures of their experimental evaluation.

Imbalanced distributions in data and methods of dealing with them.

Unfairness in artificial intelligence systems (Fairness-aware ML). Sources and various biases of learning data. Measures. Selected methods for improving data un-biases.

The need to explain the prediction of black box learning systems. Examples of methods.

Objectives of conducting scientific research. Types of research projects. Scientific knowledge and criteria for its evaluation. Formulation of the research problem, hypotheses and research objectives. The course of a standard research proces.

Reproducibility of AI methods and results of its experimental evaluation.

Ethical aspects of conducting research.

Teaching methods

Lecture: slide show presentations illustrated with examples and practical assignments. An active discussion with students.

Seminat classes: Work with scientific texts, discussion of selected issues, report-essay writing and presentation of their results. Attempts to activate students in independent and critical thinking. Encouraging various forms of their intelectual curiosity and their own looking for additional materials.

Bibliography

Basic

S. Matthew Liao, Ethics of Artificial Intelligence and Robotics. Oxford Univ. Press 2020.

Jan Recker, Scientific Research in Information Systems: A Beginner's Guide. Springer 2021.

M.Ryan, V. Stahl: Artificial intelligence ethics guidelines for developers and users: clarifying their content and normative implications, JICES, 2021.

K.Siau, W.Wang: Artificial Intelligence: Ethics of AI and Ethical AI. JDM 2020

Additional

B.Stahl.D.Schroeder, R.Rodrigues, Ethics of Artificial Intelligence. Springer 2022.

Wayne C. Booth, Gregory G. Colomb, Joseph M. Williams, Joseph Bizup, The craft of research. The University of Chicago Press, 2016.

Breakdown of average student's workload

	Hours	ECTS
Total workload	55	2,0
Classes requiring direct contact with the teacher	30	1,0
Student's own work (literature studies, preparation for seminar classes, writing a report and creating a presentation; preparation for the assessment test, addfitional project preparation) [1]	25	1,0

[1] delete or add other activities as appropriate