Warsaw University of Technology

COURSE OFFERED IN THE DOCTORAL SCHOOL

Code of the course	4606-ES-000000C-0250	Name of the course	Polish	Statystyka i Teoria Sztucznej Inteligencji i Aplikacja SI w Inżynierii
			English	Statistic and Theory of AI and its application to engineering sciences
Type of the course	Special courses			
Course coordinator	Dr hab. inż. Jordi Mongay Batalla, profesor uczelni			
Implementing unit	Institute of Telecommunication	Scientific discipline / disciplines*	information and communication technology	

Level of education	Doctoral studies	Semester	Summer		
Language of the course	English				
Type of assessment:	Graded credit	Number of hours in a semester	60	ECTS credits	4
Minimum number of participants	10	Maximum number of participants	50	Available for students (BSc, MSc)	Yes/~~No~~

Type of classes		Lecture	Auditory classes	Project classes	Laboratory	Seminar
Number of hours	in a week	2	0	2	0	0
	in a semester	30	0	30	0	0

* does not apply to the Researcher's Workshop

1. Prerequisites

PhD students from the Doctoral School

2. Course objectives

The aim of the course is to understand statistics and theories related to Machine Learning (ML) and Artificial Intelligence (AI). The course deals with the theoretical knowledge of popular AI/ML tools and the principles of mathematics behind them. Thanks to the knowledge gained in this course, the student will be able to understand the possibilities (potential and limitations) of machine learning and deep learning tools as well as explain the success of artificial intelligence in various areas of engineering.

Knowing about machine learning theory helps building ML models for a given problem statement in various fields of engineering. In this course, students will learn about classification or regression statistics, decision making under uncertainty, supervised and unsupervised learning, and reinforcement learning. In addition, students will be able to apply new skills to any engineering problem requiring a machine learning approach. One problem will be presented and solved by students (proposed, defined, implemented and tested) within the framework of projects carried out by students.

Artificial Intelligence aims to make computer systems more "intelligent" to solve complex problems and provide more natural and effective services to people. Artificial Intelligence is a source of innovative IT ideas and techniques and is widely used in many research fields. This course provides a comprehensive introduction to the mathematical methods used in AI, from statistics and statistical modeling, to decision making in uncertainty, to machine learning and deep learning.

The first part of the course will reveal statistical data learning theories, which are a toolkit for understanding data. These tools essentially cover two classes: supervised learning and unsupervised learning. Basically, supervised learning refers to the prediction of outcomes based on one or more inputs. One or more estimators make such a forecast. The choice of the estimator (s) is closely related to the nature of the data. On the other hand, unsupervised learning provides a relationship or pattern in data with no supervised outcome.

The second part of the theory described in the course presents the theory of decision making in AI. This theory discusses how to represent knowledge, including incomplete and uncertain knowledge about the real world; how to logically justify this knowledge using probabilities; how to use these models and inference methods to decide what to do, especially by making plans (Constraint Satisfaction Problems, CSP); and how to reason and make decisions (Multi-criteria decision making) in the face of uncertainty about the world.

The final part deals with Machine Learning, which describes both symbolic and statistical learning methods as well as reinforcement learning, deep learning, and multi-agent learning (game theory) to generate the knowledge required by the reasoning and / or decision components of intelligent agents or systems. Here, the methods and algorithms for providing machine learning will be analyzed and the theory of artificial intelligence will be revealed and analyzed.

Students will use all the knowledge gained in the lesson to develop AI/ML software that will be embedded in the student's engineering profile. Students from all engineering fields are invited to this course. Different student profiles generate a wide variety of different AI/ML applications implemented by students in their projects. The course will include a presentation of a joint discussion of each project.

3. Course content (separate for each type of classes)

Lecture

The content of the lecture covers three main areas: statistics, decision-making theory and machine learning theory.

The first lesson will reveal the history of artificial intelligence and its applications in the present world: interfaces with people (e.g., natural language processing, understanding emotions); machine comprehension (classifying images and videos; games); process management (warehouses, robotics); network management (security, self-management). In addition, the concept of a logical agent was presented as an entity that learns from the environment and introduces changes in the environment. The logical agent can create new rules of operation (learning) taking into account new information received from the environment. The loop consisting of the following elements: 'environment - sensors - agent intelligence - actuators - environment' is the basis of the learning process. In this loop, decision making and learning theory are critical to proper rule reasoning.

The rest of the course will be broken down into three main blocks: Statistics, Decision-Making Principles, and Machine Learning Basics.

Statistics will cover a brief description of the basics of statistics for advanced multi-criteria analysis. Specifically:

- Basics of descriptive statistics (trend, coefficient of variation, skewness, correlation)
- Measures of central tendency, asymmetry and volatility
- Estimators and estimates. Confidence intervals: advanced topics. Statistical inference
- Hypothesis testing
- One-variable analysis (discriminant analysis, Chi square, linear regression, analysis of variance: ANOVA)
- Multiple Variable Analysis

The decision-making principles will introduce statistics into artificial intelligence algorithms, and more specifically, we will deal with the problem of V classification regression, which characterizes the decisions of quantitative or qualitative (categorical) variables. Variants and prejudices are of particular importance in making decisions. Variance refers to the amount by which the function would change if we estimated it with different sets of training data. Basically, when we over-fit a model to a given training data set (the reduction error in the training set is very low, but it is very high in the validation set), then we end up with a model that has more variance because each training data point changes the results of the model (e.g., the weights) in a significant way. On the other hand, bias refers to an error introduced by approximating a real problem that can be extremely complicated by a much simpler model - for example, modeling nonlinear problems with a linear model. Basically, when we over-adjust, the model on a given dataset produces very less variation.

In addition to the classical decision-making process (single variable and multiple variables), the lesson will discuss reference-level decision making, which is a potential tool for solving many different decision problems. It is assumed that the data is known, in particular that the relationship of the data to each of the variables is known.

The fundamentals of the Machine Learning will be studied. Concretely, we will analyze: back propagation method, classifier and regression, PAC theory (probably approximately correct), data value. The following ML tools will be studied:

- Support Vector Machine is a classifier that maximizes the margin. The purpose of the classifier is to find a hyperplane with dimension (n-1) that separates two classes present in n-dimensional space.
- Logistic regression models the probability of a given output response belonging to a specific category. The basic idea is to estimate in such a way that the estimated value and the observed value of the results are as close as possible. An example of one neuron in neural networks.

- The K nearest neighbors The classifier is a lazy learning technique in which a set of training data is represented on the Euclidean hyperplane and test data is assigned to labels based on Euclidean distance metrics. The K nearest neighbors The classifier allows you to learn a complex objective function without losing information, but its cost is high.
- Generalized additive models provide a generalized structure that extends standard multivariate linear regression with the nonlinear function of each variable, while maintaining its additive nature. In this way, all nonlinear functions can be independently calculated and added later.
- Decision trees or trees are useful and simple methods for both regression and classification that segment the predictor space into simple regions. Typically, decision trees are drawn upside down, which means that the leaves are at the bottom of the tree. The points at which the predictor space is divided are called interior nodes, and the leaf nodes or terminal nodes are those that produce predictions. For prediction, we take a top-down approach (in the first point, all observation belongs to only one region), a greedy (best differentiation is made at a given stage) approach known as recursive binary matching. There are tree pruning strategies that overcome the problem of trees by cutting down some branches to make a small sub-tree.
- Packaging is a general purpose method for reducing variance in a statistical learning method. The basic idea is that averaging a set of observations reduces the variance. Therefore, we randomly sample our data multiple times and create a tree for each sample and average all the predictions to get a low variance result.
- The supportive approach is a slow learning statistical method in which classifiers are trained on a modified data set sequentially. In the context of decision trees, each tree is grown based on information from previous trees. Thus, we do not fit into one large tree.

Finally, the application of machine learning algorithms to artificial intelligence will focus on unsupervised and reinforcing learning.

The most commonly used statistical method for unsupervised learning is K-Means Clustering. This method takes "k" random points from the dataset and maps all other points to one of the K regions based on their proximity to the K selected random points. The algorithm then changes random K points to the centroid of the clusters thus formed and repeats the operations until the algorithm notices a slight change in the cluster created after each iteration. Other techniques such as the analysis of the main components in unsupervised learning will also be explained.

Reinforcement learning is a scenario where the agent has to make many decisions before reaching a goal and provides a reward, "+1" or "-1", instead of being notified of how well or bad the path agent has done. Methods such as the Markov decision-making process, Monte Carlo methods and learning time differences will be discussed:

- The Markov Decision Process (MDP) provides a mathematical framework for modeling an agent's decision-making in situations or environments where the outcomes are partially random and partially controlled. In this model, the environment is modeled as a set of states and actions that the agent can perform to control the state of the system. The goal is to control the system in such a way that the total payout of the agent is maximized.
- Monte Carlo methods do not require full knowledge of the environment, unlike MDP. Monte Carlo methods only require experience that is obtained from exemplary sequences of states, actions, and rewards from actual or simulated interaction with the environment. Monte Carlo methods examine space through to the final result of the selected sample sequences and update the estimates accordingly.

Learning time difference is a key topic in reinforcement learning. The time difference is a combination of both Monte Carlo and dynamic programming ideas. As in Monte Carlo, time difference methods can learn directly from raw experience without an environmental dynamics model. Like dynamic programming, time difference methods update estimates partially based on other learned estimates, without waiting for the final result.

Projects
The project aims to develop machine learning software to generate your own ideas and visions, and discuss and share deep understandings. The student will analyze his / her own scientific discipline in terms of

implementing Artificial Intelligence and Machine Learning tools. The result of this analysis will be an understanding to what extent AI and ML are used as valuable tools for the implementation of scientific research and potential new ideas for the PhD student in his research work.

The student will be asked to think through and convey higher-level abstract ideas, including concepts, strategies, principles, and algorithms for artificial intelligence, rather than the technical details of implementation and programming. You will also be encouraged to test and evaluate ideas and algorithms through research-oriented research involving programming and / or exploration.

The software will be developed in any AI adaptive language such as Python, C ++ or any specialized language (generally developed by strong AI technology vendors). Given the diversity of PhD students, our idea is to open up potential emerging platforms as much as possible to engage many different engineering ambitions. The project will be assessed in terms of concept, problem description, algorithm selection and development, as well as final software performance in terms of AI quality. Aspects such as software performance and details of the developed programs will not be taken into account during the qualification.

The stages of the project implementation will include: (1) presenting the problem with a detailed explanation of the importance of AI for its solution; (2) a detailed description of the delivered solution, including algorithms and mechanisms; (3) mathematical analysis of the solution; (4) software development: from choosing the deployment platform and language to identifying and describing libraries; (5) implementation; (6) debugging; (7) testing. All steps will be documented.

The projects will be implemented mainly using any language for Artificial Intelligence software (Python, C ++, etc.).

4. Learning outcomes			
	Learning outcomes description	Reference to the learning outcomes of the WUT DS	Learning outcomes verification methods*
	Knowledge		
K01	Studenci będą wdrażać narzędzia Sztucznej Inteligencji lub Uczenia Maszynowego do własnej dziedziny inżynierii. Chodzi o to, aby projekty przygotowywane przez studentów odpowiadały na aktualne problemy badawcze związane z aktualnymi danymi uzyskanymi z różnych zastosowań inżynierii	SD_W1	Project evaluation
K02	Pierwszym krokiem realizacji projektu będzie analiza impaktu Sztucznej Inteligencji we własniej branży inżynierii. Złożoność wymagana do projektu będzie zależało jak rozwinięta jest Sztuczna Inteligencja (lub Uczenie Maszynowe) we właśniej branży. Jeśli Sztuczna Inteligencja jeszcze nie istnieje w zareprezentowanej przez doktoranta inżynierii, wtedy projekt będzie miał niewiele złożoności.	SD_W2	Project evaluation
K03	Jak powyżej student będzie analizował narzędzia Sztucznej Inteligencji używane we właśniej dyscyplinie i będzie rozumiał zasady matematyczne działające za takie narzędzie używane coraz bardziej w inżynierii.	SD_W3	Project evaluation
	Skills		
S01	Projekt realizowany przez doktoranta będzie zawierał: • cel i przedmiot badań, oraz hipotezę badawczą; • techniki i narzędzia badawcze bazowane na Sztuczną Inteligencję; • wyniki i analiza wyników badań	SD_U1	Project evaluation

S02	Wykłady i ogólnie cały przedmiot będzie skupiony na charakteryzację danych jako wartość dla badań naukowych w reprezentowanej dyscyplinie naukowej.	SD_U2	Project evaluation
S03	Ideą tworzenia projektu jest wykonanie badań na podstawie narzędzi Sztucznej Inteligencji i Uczenia Maszynowego. Takie badania powinny być właściwe dla reprezentowanej dyscypliny naukowej i projekt powinny być zrobiony z standardami publikacji naukowej (i publikować badania przez studenta jeśli to jest możliwe).	SD_U4	Project evaluation
S04	Wykłady oraz projekty będą po angielsku. Student będzie musiał przygotować raport projektu po angielsku. W projekcie student będzie musiał analizować międzynarodową literaturę na temat uczenia maszynowego i sztucznej inteligencji	SD_U6	Project evaluation and lecture attendance
S05	Tak jak omówiono powyżej, badania realizowane w ramach projektu powinny być właściwe dla reprezentowanej dyscypliny naukowej i projekt powinny być zrobiony z standardami publikacji naukowej (i publikować badania przez studenta jeśli to jest możliwe). Projekty będą samodzielne.	SD_U7	Project evaluation
Social competences			
SC01	Analiza zrobiona przez doktoranta odnośnie zużycia narzędzi Sztucznej Inteligencji we własnej dziedzinie inżynierii ma dawać wizję odnośnie nowych kierunków i narzędzi matematycznych używanych w nauce i ulokować lepiej badania studenta na tle nauki międzynarodowej.	SD_K1	Project evaluation

*Allowed learning outcomes verification methods: exam; oral exam; written test; oral test; project evaluation; report evaluation; presentation evaluation; active participation during classes; homework; tests

5. Assessment criteria
The evaluation will be based on the realisation of the individual projects.

6. Literature

Basic references:

[1] Artificial Intelligence: A Guide for Thinking Humans Kindle Edition – Farrar, Straus and Giroux Eds. - by Melanie Mitchell

[2] George Mastorakis, Constandinos X. Mavromoustakis, Jordi Mongay Batalla, Evangelos Pallis, Convergence of Artificial Intelligence and the Internet of Things, Springer Eds. 2020. ISBN 978-3-030-44906-3

[3] Artificial Intelligence: A Modern Approach, Global Edition – Pearson Education Limited - by Stuart Russell, Peter Norvig

[4] Analytics of Life: Making Sense of Artificial Intelligence, Machine Learning and Data Analytics – NLITX Eds. - by Mert Damlapinar

[5] Artificial Intelligence with Python: Your complete guide to building intelligent apps using Python 3.x and TensorFlow 2 – Packt Publishing - by Alberto Artasanchez, Prateek Joshi

Other references:

[1] George Mastorakis, Constandinos X. Mavromoustakis, Jordi Mongay Batalla, Evangelos Pallis, Intelligent Wireless Communications, The Institution of Engineering and Technology Eds. 2021. ISBN 978-1-83953-095-1

[2] Kevin P. Murphy, Machine Learning: A Probabilistic Perspective (Adaptive Computation and Machine Learning series). The MIT Press. ISBN-13: 978-0262018029. 2012

[3] Christopher M. Bishop: Pattern Recognition and Machine Learning (Information Science and Statistics) (Information Science and Statistics series). Springer Nature. ISBN-13: 978-0387310732. 2011

[4] Papers of Special Issue on "Artificial Intelligence for Cloud Based Big Data Analytics" of Big Data Research journal (Elsevier) with Editors: Konstantinos E. Psannis, Yutaka Ishibashi, Jordi Mongay Batalla, Brij Gupta, Muhammad Imran, Byung-Gyu Kim, Ibrar Yaqoob: https://www.journals.elsevier.com/big-data-research/call-for-papers/artificial-intelligence-for-cloud-based-big-data-analytics

7. PhD student's workload necessary to achieve the learning outcomes**		
No.	Description	Number of hours
1	Hours of scheduled instruction given by the academic teacher in the classroom	60 (30 h. lecture + 30 h. project)
2	Hours of consultations with the academic teacher, exams, tests, etc.	10 h.
3	Amount of time devoted to the preparation for classes, preparation of presentations, reports, projects, homework	48 (8 h. literature study + 40 h. project)
4	Amount of time devoted to the preparation for exams, test, assessments	2h. (project presentation)
	Total number of hours	**120**
	ECTS credits	**4**

** 1 ECTS = 25-30 hours of the PhD students work (2 ECTS = 60 hours; 4 ECTS = 110 hours, etc.)