Syllabus
Course Program

Data-Driven Approaches in Modeling

Specialty
113 – Applied Mathematics

Institute
Institute of Computer Modeling, Applied Physics and Mathematics

Educational program
Computer and Mathematical Modeling

Department
Mathematical Modeling and Intelligent Computing in Engineering (161)

Level of education
Master's level (1 year 4 months

Course type
Special (professional), Elective

Semester
2

Language of instruction
English

Lecturers and course developers

Oleksii Vodka (responsible lecturer)

Oleksii.Vodka@khpi.edu.ua
PhD, docent, Head of Mathematical modelling and intellectual computing in engineering department

General information, number of publications, main courses, etc.

More about the lecturer on the department's website

Nataliia Fomenko (assistant)

Nataliia.Fomenko@infiz.khpi.edu.ua
Assistant lecturer

More about the lecturer on the department's website

General information

Summary

Data-Driven Modeling offers a thorough study of the use of data to build predictive and informative models in a variety of scientific and engineering disciplines. This course is designed to provide students with the theoretical framework and practical skills needed to extract valuable information from data, build robust models, and solve problems effectively.

Course objectives and goals

Aims: the course aims to teach students modern methods of analyzing and modeling complex physical phenomena using linear and nonlinear regression, Dynamic Mode Decomposition, physically informed neural networks, and deep learning tools. The discipline is focused on developing skills in processing and

analyzing large amounts of data, as well as understanding the physical principles underlying various phenomena.
Objectives:
1. To master the methods of creating models based on data sets.
2. Gain knowledge of advanced methods such as symbolic regression, physically informed neural networks, deep learning
3. Improve skills in solving real-world problems.
4. Gain the ability to interpret and communicate results effectively by explaining the inner workings of models, translating results into meaningful conclusions, and presenting them convincingly to a diverse audience.

Format of classes

Lectures, laboratory classes, consultations, self-study. Final control in the form of an exam.

Competencies

GC1. Ability to generate new ideas (creativity) and non-standard approaches to their implementation.
GC3. Ability to master modern knowledge, formulate and solve problems.
PC3. Ability to develop methods and algorithms for the construction, research and software implementation of mathematical models in engineering, physics, biology, medicine and other fields and to analyse them.
PC4. Ability to develop and research mathematical and computer models, conduct computational experiments and solve formalised problems using specialised software.
PC7. Ability to design and develop software to solve formalised problems, including systems with large amounts of data.
PC8. Ability to formalise and build data or knowledge models, obtain relevant knowledge from large amounts of data, choose data mining methods to solve problems.

Learning outcomes

LO5. Justify and, if necessary, develop new algorithms and software tools for solving scientific and applied problems, apply, modify and investigate analytical and computational methods for solving them.
LO7. Solve computer modelling problems by using and developing modern software tools, in particular, distributed, parallel and cloud programming methods.
LO8. Develop and implement algorithms for solving applied problems, system and application software of information systems and technologies.

Student workload

The total volume of the course is 120 hours (4 ECTS credits): lectures - 32 hours, laboratory classes - 16 hours, self-study - 72 hours.

Course prerequisites

Basic understanding of linear algebra, mathematical statistics and probability theory. Experience in Python programming (familiarity with such libraries as NumPy, pandas and scikit-learn is a plus).

Features of the course, teaching and learning methods, and technologies

Interdisciplinary nature: a combination of knowledge from computer science, applied mathematics, physics and engineering.
Practical orientation: emphasis is placed on the application of theoretical knowledge to solve real-world problems in various fields.
The use of modern technologies: laboratory work and independent work involve the use of Python, popular libraries (NumPy, pandas, scikit-learn, TensorFlow/PyTorch) and advanced methods (DMD, PINN, Deep learning).
Development of research skills: students learn how to analyze data, formulate research questions, generate new ideas, and present research results

Program of the course

Topics of the lectures

Topic 1: Introduction
Overview of modeling approaches The importance and application of data-driven modeling.
Topic 2. Regression and model selection
Topic 3. Nonlinear regression
Topic 4. Logistic regression
Topic 5. Genetic programming
Topic 6. Symbolic regression
Topics 7-8. Dimensionality reduction. Dynamic Mode Decomposition
Topic 9. Clustering algorithms
Topic 10. Fundamentals of physically informed neural networks (PINN)
Topic 11: Implementation and training methods (PINN)
Topic 12: Application of physically informed neural networks
Topics 13-14. Deep learning
Topics 15-16. RL for physical systems

Topics of the workshops

Topics of the laboratory classes

Topic 1: Application of linear regression
Topic 2. Comparison of regression methods
Topic 3. Symbolic regression for modeling nonlinear behavior of materials
Topics 4-5. Analysis of fluid dynamics using Dynamic Mode Decomposition
Application of DMD to analyze a data set of time series of velocity measurements in a turbulent flow and to identify its dominant dynamic modes.
Topic 5. Definition of a simple PINN architecture. The loss function. Solving a partial differential equation
Implementation of the basic PINN architecture using TensorFlow or PyTorch. Analysis of the impact of the choice of loss function on model performance. Comparison of the results of solving the PDE with analytical solutions.
Topic 6. Application of physically informed neural networks
Simulation of hydrodynamics. Application of PINN for modeling fluid dynamics in a controlled environment. Checking the results for compliance with the established principles of fluid dynamics. Modeling of mechanical behavior. Prediction of deformations and/or stresses in a material under specified conditions. Comparison with the results obtained by the FEM.
Topic 7-8. Deep learning for time series forecasting
Using convolutional LSTMs to analyze wind speed sensor data, production data, and weather forecasts to predict wind turbine power.

Self-study

A final task on one of the laboratory topics.

Course materials and recommended reading

[1] J. N. Kutz, Data-driven modeling & scientific computation: Methods for complex systems & big data. London, England: Oxford University Press, 2013.
[2] S. L. Brunton and J. N. Kutz, Data-driven science and engineering: Machine learning, dynamical systems, and control, 2nd ed. Cambridge, England: Cambridge University Press, 2022.
[3] K. P. Murphy, Machine Learning: A Probabilistic Perspective. London, England: MIT Press, 2012.
[4] J. VanderPlas, Python Data Science Handbook. Sebastopol, CA: O'Reilly Media, 2016. (https://jakevdp.github.io/PythonDataScienceHandbook/)

Other support materials: links to tutorials or instructions for installing certain packages/libraries are provided during the course.

Assessment and grading

Criteria for assessment of student performance, and the final score structure

Each laboratory work (8) is assessed at 10 points (maximum for all - 80). Self-study is assessed at 10 points. Answer to the exam - 10 points.

Grading scale

Total points	National	ECTS
90–100	Excellent	A
82–89	Good	B
75–81	Good	C
64–74	Satisfactory	D
60–63	Satisfactory	E
35–59	Unsatisfactory (requires additional learning)	FX
1–34	Unsatisfactory (requires repetition of the course)	F

Norms of academic integrity and course policy

The student must adhere to the Code of Ethics of Academic Relations and Integrity of NTU "KhPI": to demonstrate discipline, good manners, kindness, honesty, and responsibility. Conflict situations should be openly discussed in academic groups with a lecturer, and if it is impossible to resolve the conflict, they should be brought to the attention of the Institute's management.
Regulatory and legal documents related to the implementation of the principles of academic integrity at NTU "KhPI" are available on the website: http://blogs.kpi.kharkov.ua/v2/nv/akademichna-dobrochesnist/

Approval

Approved by	Date	
	Date August 30, 2023	Head of the department Oleksii VODKA
	Date August 30, 2023	Guarantor of the educational and professional program (1 year 4 months) Oleksiy LARIN