Syllabus
Course Program

Mathematical Methods of Data Visualization

Specialty
113 – Applied Mathematics

Institute
Institute of Computer Modeling, Applied Physics and Mathematics

Educational program
Computer and Mathematical Modelling

Department
Mathematical Modeling and Intelligent Computing in Engineering (161)

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

Course type
Special (professional), Mandatory

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

Ruslan Babudzhan (assistant)

Ruslan.Babudzhan@infiz.khpi.edu.ua
Assistant lecturer

More about the lecturer on the department's website

General information

Summary

This course covers a wide spectrum of mathematical models of data visualization topics, from foundational concepts to advanced techniques. Beginning with an overview of chart types and potential misrepresentations, it proceeds to explore Python libraries, software tools, and specialized plots. The course delves into advanced domains, discussing the visualization of fields, vectors, tensors, tree structures, and graph representations. It also covers dimension reduction methods, text data visualization using LaTeX, and the principles of effective infographic design. With practical examples and comparisons, the course aims to provide a comprehensive understanding of data visualization for effective communication.

Course objectives and goals

Objectives:
– Understand the importance and levels of data visualization. Chart Proficiency: Master various chart types and recognize potential misrepresentations.
– Gain practical proficiency in Python libraries like Matplotlib, Seaborn, and Plotly.
– Explore advanced visualization methods for fields, vectors, tensors, trees, and graphs.
– Apply both linear and nonlinear dimension reduction methods using Sklearn.
– Learn LaTeX for effective text data visualization.
– Grasp principles for designing effective infographics.

Goals:
– Develop practical skills for accurate and compelling visualizations.
– Foster critical thinking in choosing appropriate visualization methods.
– Apply visualization techniques to real-world scenarios.
– Enhance communication through clear and accurate visualizations.
– Gain adaptability in using diverse visualization tools.
– Understand applications in various domains.
– Develop creative problem-solving skills in data representation.

Format of classes

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

Competencies

GC1. Ability to generate new ideas (creativity) and non- standard approaches to their implementation.
GC5. Ability to conduct professional activities, in particular in the international environment.
GC7. Ability to think abstractly, analyse and synthesise.
GC8. Ability to communicate and carry out professional activities in the state language and the language of the EU countries.
PC1. Ability to solve tasks and problems that can be formalised, require updating and integrating knowledge, in particular in conditions of incomplete information.
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

LO1. Communicate within the scope of professional competences in one of the EU languages.
LO4. Build mathematical models of complex systems and choose methods of their research, implement the built models in software and check their adequacy using computer technologies.
LO8. Develop and implement algorithms for solving applied problems, system and application software of information systems and technologies.
LO10. Develop and apply modern concepts of machine learning and data mining.
LO14. To have the knowledge to mathematically formalise the formulation of scientific and practical problems, to choose a mathematical analytical or numerical method of its solution, which ensures the required accuracy and reliability of the result.
LO15. To be able to carry out mathematical and computer modelling, computational experiment, solve formalised problems using specialised software.

Student workload

The total volume of the course is 90 hours (3 ECTS credits): lectures - 16 hours, laboratory classes - 16 hours, self-study - 58 hours.

Course prerequisites

Basic programming skill with Python language, knowledge algorithm and mathematical concept like scalar, vector, tensor, graph, tree.

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

Not appliable.

Program of the course

Topics of the lectures

Topic 1. Introduction to data visualization.
Levels of visualization. Chart types: Line, Pie, Bar, Histogram, Spaghetti, Map, Steam, Bubble, Gantt, Sunburst, Polar Clock, Radar, Tag cloud, Heat map, Trees, Mind Map. How to lie with charts? Chart Junk.
Topic 2. Software for chart plotting in Python.
Matplotlib, Seaborn, Plotnine(ggplot), Bokeh, pygal, Plotly, geoplotlib, Gleam, missingno, Leather, Altair, Folium. 2D and 3D plots. Log scale. Box plot, Violin plot. Examples.
Topic 3. Visualization of fields, vector, tensors
Introduction and problem statement. Scalar filed. Gradient of scalar field. Divergence and Vorticity. Vector Glyphs. Vector Color Coding. Displacement Plots. Stream Lines, Tensor glyphs, Tensor field lines, Hyperstreamlines.
Topic 4. Tree visualization.
Historical context, first paper of tree visualization (Tutte, Knuth). Application of tree visualization (UML, Biology, Networks, Security, Social Networks). Requirements, conventions to tree visualization. D. Knuth algorithm, Theorem of area of tree laying. Tree Drawing Algorithm: Layered Drawing, Radial Drawing, HV-Drawing, Recursive Winding. Dendrogram. Balloon trees. Hyperbolic Tree. Space-filling methods: Tree map, Beamtree, Icicle.
Topic 5. Graph visualization
Introduction and problem statement. Algorithm of graph drawing: Straight Line Drawing. Orthogonal Drawing. Grid Drawing. Circular Drawing. Polyline Drawing. Upward drawing. Sugiyama method. Force directed graph drawing (mathematical problem statement): Kamada & Kawai, Fruchterman & Reingold, Davidson & Harel.
Software for Graph visualization: NetworkX, Gephi. Examples. Visualization of large graph.
Topic 6. Method of Dimension reduction
Introduction and problem statement. Linear method. PCA and SVD. Examples: Iris, Eigen Faces, MNIST. Discussion. Nonlinear method: MDS, ISOMAP, LLE, Laplacian eigenmaps, SNE, tSNE. Examples. Method comparation. Sklearn library.
Topic 7. Visualization of text data
Introduction to latex, setting up a latex document, typesetting text, handling latex errors, typesetting equations, using latex packages. Structured documents, sections, labels and cross-references, figures and tables in latex, automatic bibliographies with bibtex, useful latex packages and online resources, latex presentations with beamer, drawing in latex with tikz, tracked changes and comments with todo notes.
Topic 8. Infographics
Definition and purpose, Historical context and evolution, Importance in modern communication. Principles of Effective Infographic Design: Simplicity and clarity, Visual hierarchy, Consistency and coherence, Audience consideration. Types of Infographics: Statistical infographics, Informational infographics, Timeline infographics, Comparison infographics, Process infographics, Geographical infographics. Choosing the Right Data for Infographics: Identifying key messages, Color, Typography, and Layout. The psychology of color in communication, Typography best practices, Layout considerations for optimal readability.

Topics of the workshops

Not appliable.

Topics of the laboratory classes

Class 1. Creating charts with matplotlib
Class 2. Visualization scalar, vector and tensor fields
Class 3. Creating custom drawing and animation with matplotlib
Class 4. Visualization of binary tree
Class 5. Visualization graphs with networkx and Gephi
Class 6. Visualization of graph obtained from social network
Class 7. Nonlinear dimension reduction
Class 8. Linear dimension reduction with PCA and SVD
Class 9. Creating reports with Latex (Overleaf)

Self-study

Self-study also includes preparation to laboratory classes and report preparation on passed laboratory classes.

Course materials and recommended reading

1. Handbook of Data Visualization. (2007). Germany: Springer Berlin Heidelberg.
2. Ward, M. O., Grinstein, G., Keim, D. (2015). Interactive Data Visualization: Foundations, Techniques, and Applications, Second Edition. USA: CRC Press.
3. Knaflic, C. N. (2015). Storytelling with Data: A Data Visualization Guide for Business Professionals. Germany: Wiley.
4. Telea, A. (2015). Data Visualization: Principles and Practice, Second Edition. Philippines: Taylor & Francis.
5. Data Visualization: Trends and Challenges Toward Multidisciplinary Perception. (2020). Germany: Springer Nature Singapore.
6. Setlur, V., Cogley, B. (2022). Functional Aesthetics for Data Visualization. USA: Wiley.

Assessment and grading

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

Total mark (100 points) consists of two parts:
1. Test on theory (60 points)
2. Practice (lab) passing (40 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 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