COURSE DESCRIPTION CARD - SYLLABUS

Course name
Multidimensional Data Visualization

Course

Field of study
Computing

Year/semester
2/3

Area of study (specialization)
Artificial Intelligence

Profile of study
general academic

Level of study
Second-cycle studies

Course offered in
Polish

Form of study
full-time

Requirements
elective

Number of hours

Lecture
15

Laboratory classes
15

Other (e.g. online)

Tutorials

Projects/seminars

Number of credit points

2

Lecturers

Responsible for the course/lecturer:

dr hab. inż. Robert Susmaga
e-mail: robert.susmaga@cs.put.poznan.pl
tel.: 616652934
Faculty of Computing and Telecommunications
Piotrowo 2, 60-965 Poznan

Responsible for the course/lecturer:

Prerequisites

Basic knowledge regarding elements of linear algebra (vectors, matrices, vector/matrix operations) and analytic geometry (analysing and graphing simple functions).
Basic skills regarding designing, implementing and testing computer programs (in a programming language of one's choice) that implement simple processing of static data structures (vectors and matrices) and generate graphs of simple functions.
(recommended) A fair amount of cognitive curiosity and not less perseverance in pursuing the goals of personal development.

Course objective

Detailed objectives of the course include:

1) Sharing knowledge regarding:
a) properties of multidimensional spaces, vector/matrix operations in multidimensional spaces, properties and applications of the Eigenvalue Matrix Decomposition (EVD),
b) selected transformations and visualizations of multidimensional data, including representations in barycentric coordinate systems; selected dedicated methods, in particular: the Principal Components Analysis (PCA), Multidimensional Scaling (MDS).
2) Sharing skills regarding:
a) identifying, formulating and solving basic problems related to transforming and visualizing multi-dimensional data,
b) designing and creating computer programs that successfully implement the presented methods and algorithms.

Course-related learning outcomes

Knowledge

Students:
-- gain advanced detailed knowledge regarding selected IT issues, especially concerning various aspects of data analysis, including dimensionality reduction (PCA), data smoothing (PCA) and visualization (PCA, MDS) [K2st_W3],
-- gain knowledge about development trends and the most important cutting edge achievements in computer science and other selected and related scientific disciplines, in particular: advanced knowledge related to various aspects of miltidimensional data analysis [K2st_W4]

Skills

Students:
-- can use analytical, simulation and experimental methods to formulate and solve engineering problems and simple research problems, in particular those related to working with multidimensional data [K2st_U4]
-- can -- when formulating and solving engineering tasks -- integrate knowledge from different areas of computer science (and if necessary also knowledge from other scientific disciplines) and apply a systemic approach, also taking into account non-technical aspects, like the multi-aspect human-computer communication issues (data visualization) [K2st_U5]
-- are able to assess the suitability and the possibility of using new achievements (methods and tools) and new IT products, especially those used in multidimensional data processing (dimensionality reduction/selection) [K2st_U6]
-- can carry out critical analyses of existing technical solutions, especially those regarding dimensionality reduction, and put forward improvements/enhancements of those [K2st_U8]
-- are able -- in accordance with a given specification, taking into account non-technical aspects - to design a complex IT system, in particular: a multidimensional data visualization utility -- at least in part - using appropriate methods, techniques and tools, including adapting to this purpose existing tools or developing new ones [K2st_U11]

Social competencies

Students:

-- understand that in the field of IT the knowledge and skills quickly become obsolete [K2st_K1],
-- understand the importance of using the latest knowledge in the field of computer science in solving research and practical problems [K2st_K2].

Methods for verifying learning outcomes and assessment criteria

Learning outcomes presented above are verified as follows:
Formative assessment (laboratory classes): evaluation of the solutions to the assigned programming problems (as they arise).
Final assessment:
-- (laboratory classes): evaluation of the solutions to the assigned programming problems,
-- (lectures): evaluation of the results of a written test (45--60 min) with both multiple choice, short answer and (small) computational questions (mostly: micro-problems to be solved in writing).

Programme content

(Lectures)
Introduction: Multidimensional data and multidimensional spaces. The visualization idea, its advantages and disadvantages. Aspects of visualizing low-dimensional data, aspects of visualizing of high-dimensional data.
Barycentric coordinate systems (three-dimensional and four-dimensional). Applications of four-dimensional barycentric coordinate systems: visualization of confirmation measures and classification performance measures.
Vectors and basic vector operations. Multidimensional vector spaces, vector inner products, projections, angles and norms. Orthogonality and orthogonal vectors. Matrices and basic matrix operations. Matrices as data and as operators. Fundamental scalar characteristics of matrices. Orthogonal matrices, inverse matrices. Spectral analysis: eigenvalues and eigenvectors of matrices with applications.
The Eigenvalue Decomposition of matrices (EVD): the formulation, basic properties and basic applications in matrix algebra. The Principal Component Analysis (PCA): the formulation (using EVD on matrices of covariances), basic properties and basic applications in data analysis (dimensionality reduction, data smoothing and visualization). The Multidimensional Scaling (MDS): the formulation (using EVD on matrices of distances), basic properties and basic applications in visualization.
(Optional material) Correspondence Analysis (CA): idea, formulation and and exemplary applications, t-distributed Stochastic Neighbour Embedding (t-SNE): idea, formulation and exemplary applications.

(Laboratory classes)
Introduction to the Python programming language and a selection of its libraries: NumPy and Matplotlib. Simple programs operating on scalar, vector and matrix arguments. Visualization of scalar, vector and matrix data, scatter-plots, barycentric coordinate systems. Vector inner product, orthogonality, and norm.
The Eigenvalue Decomposition of matrices (EVD) and its exemplary applications in matrix algebra.
The Principal Component Analysis (PCA) and its exemplary applications in dimensionality reduction, data smoothing and visualization.
The Multidimensional Scaling (MDS) and its exemplary applications in visualization.

Teaching methods

Lectures: extensive slide show presentations (theoretical elements, explanations, practical examples, exercises), exemplary computations and visualizations.
Laboratory classes: instructions: short slide show presentations (examples, exercises); practice: designing and creating (in a programming language of one's choice) programs to solve the assigned problems (which illustrate the ideas and notions presented during the lectures).

Bibliography

Basic

1. G. Banaszak, W. Gajda: Elementy algebry liniowej część I i II, WNT, Warszawa, 2002

2. J. Koronacki, J. Ćwik: Statystyczne systemy uczące się , WNT, Warszawa, 2005

Additional

1. I.T. Jolliffe: Principal Component Analysis , Springer-Verlag, Nowy Jork, USA, 2002

2. I. Borg, P.J.F. Groenen: Modern Multidimensional Scaling , Springer Science+Business Media, Nowy Jork, USA, 2005

3. M. Greenacre: Correspondence Analysis in Practice , Chapman & Hall/CRC Press, Nowy Jork, USA, 2007

4. H. Dudycz: Wizualizacja danych , Wydawnictwo AE, Wrocław, 1998

Breakdown of average student's workload

	Hours	ECTS
Total workload	50	2,0
Classes requiring direct contact with the teacher	30	1,5
Student's own work (literature studies, preparation for laboratory classes/tutorials, preparation for tests/exams, project preparation)	20	0,5