UNIWERSYTET IM. ADAMA MICKIEWICZA W POZNANIU

Introduction to Topological Data Analysis

Educational subject description sheet

Basic information

Study programme Matematyka **Speciality** - **Organizational unit** Faculty of Mathematics and Computer Sciences **Study level** First-cycle programme **Study form** Full-time **Education profile** General academic	**Didactic cycle** 2025/26 **Subject code** 06MATS.1800000PW.16812.25 **Lecture languages** English **Course type** Elective **Block** Elective subjects

Subject coordinator	Bartosz Naskręcki, Łukasz Michalak, Piotr Mizerka
Lecturer	Bartosz Naskręcki, Łukasz Michalak, Piotr Mizerka

Period	**Activities and hours**	**Number of ECTS points**
Przedmioty do wyboru - semestr letni	• Laboratories: 60, Graded credit	5

Goals

Code	Goal
C1	Familiarize students with basic topological concepts relevant to data analysis, including topological spaces, simplicial complexes and persistent homology.
C2	Provide knowledge of how to represent and process geometric data and methods of topological data visualization.
C3	To make students aware of the role of topological invariants in data analysis and their applications in various scientific fields.
C4	To teach students to use Python programming libraries dedicated to topological data analysis and their implementation in practical problems.

Subject learning outcomes

Code	Outcomes in terms of	Learning outcomes	Examination methods
Knowledge – Student:			
W1	knows and understands basic topological concepts and structures, such as topological spaces, simplicial complexes and persistent homologies, and their applications in data analysis.	MAT_K1_W02, MAT_K1_W05	Written work
W2	knows and understands the importance of metrics and topology in data analysis and their impact on methods of classification, dimensionality reduction and data visualization.	MAT_K1_W02, MAT_K1_W05, MAT_K1_W07	Essay
W3	knows and understands algorithmic methods for constructing complexes from data and basic tools for topological calculations.	MAT_K1_W02, MAT_K1_W07	Written work
W4	knows and understands the concept of homotopy, homology and their computational aspects, including persistence diagrams and Euler characteristic curves.	MAT_K1_W02, MAT_K1_W05, MAT_K1_W07	Written work
W5	knows and understands the basic methods of topological data visualization, including mapper and ball mapper algorithms, and their application in the exploration of multidimensional data sets.	MAT_K1_W02, MAT_K1_W07	Essay
W6	knows and understands the role of topological methods in various fields of science and technology, including the analysis of biological structures, social networks and image processing.	MAT_K1_W01, MAT_K1_W02	Essay
Skills – Student:			
U1	is able to use mathematical formalism to define topological concepts, formulate theorems and build mathematical models in data analysis.	MAT_K1_U01, MAT_K1_U05	Written work
U2	is able to construct logical arguments in the context of topological data analysis, with clear identification of assumptions and conclusions, and apply various methods of proof in topological analysis.	MAT_K1_U02, MAT_K1_U05	Essay
U3	is able to use the tools of linear algebra and homology theory to analyze data structure and geometric representation.	MAT_K1_U04, MAT_K1_U05	Project

Code	Outcomes in terms of	Learning outcomes	Examination methods
U4	is able to implement computational topology algorithms, including algorithms for constructing complexes and calculating persistent homology, in Python using available libraries.	MAT_K1_U07, MAT_K1_U08	Project
U5	is able to apply topological data visualization methods, including the mapper algorithm and ball mapper, to analyze the geometric structures of data sets.	MAT_K1_U05, MAT_K1_U07, MAT_K1_U08	Project
U6	is able to recognize and solve data analysis problems, especially those that require the use of topological methods for dimensionality reduction and classification.	MAT_K1_U08, MAT_K1_U09	Project
U7	is able to independently acquire knowledge of topological data analysis and develop their programming and analytical skills in this area.	MAT_K1_U09	Essay
U8	is able to work individually and in a team on topological data analysis projects, effectively communicating their results.	MAT_K1_U10, MAT_K1_U12	Project
Social competences – Student:			
K1	is ready to recognize the limitations of his own knowledge in the field of topological data analysis and understands the need for further education in this field through independent exploration of the literature and critical analysis of new methods.	MAT_K1_K01	Essay, Written work
K2	is ready to work in a team on projects in the field of topological data analysis, demonstrating responsibility, entrepreneurship and ability to cooperate in an interdisciplinary environment.	MAT_K1_K02	Project
K3	is ready to independently form opinions on the application of topological methods in data analysis and to explain their implications to those not involved in the field.	MAT_K1_K03	Essay
K4	is ready to adhere to the principles of professional ethics and intellectual honesty in the analysis of data, interpretation of results and presentation of his research.	MAT_K1_K04	Project, Written work

Study content

No.	Course content	Subject learning outcomes	Activities
1.	**Fundamentals of topology and metrics in data analysis** Mathematical and computational foundations of topology. Metrics and similarity measures. Topology as a tool in data analysis. Metrics suitable for high dimensional data and the curse of dimensionality problem.	W1, W2, U1, K1	Laboratories
2.	**Data analysis methods and their applications** Overview of data analysis methods: differences between qualitative, descriptive and statistical methods. Clustering methods. Dimensionality reduction and variable selection in the context of high dimensionality data.	W2, U6, U7, K1	Laboratories

No.	Course content	Subject learning outcomes	Activities
3.	**Shape approximation in data analysis** Cellular complexes as a tool for representing complex spaces. Discussion of symplectic complexes, singular complexes, point cloud based complexes, cube complexes, and regular CW complexes. Effective data structures for storing these complexes.	W3, U3, U4, K1	Laboratories
4.	**Complexes from data and their importance in network theory** Methods for constructing complexes from trees, graphs and other structures (both abstract and embedded in Euclidean spaces). Topological invariants: homeomorphism, homotopy groups and their computational limitations.	W3, W4, U2, U3, U6, K1	Laboratories
5.	**Persistent homology and cycle analysis in data** Chains and cycles as a generalization of paths and cycles in graphs. Persistentna homology and cohomology (with coefficients in $\mathbb{Z}_2$). Algorithm of reduction of the complex edge matrix. Persistence diagrams, their interpretation and comparison methods.	W4, U2, U3, U4, U6, K1	Laboratories
6.	**Multi-parameter persistent homology theory** Motivation and limitations of multi-parameter persistent homology theory. Euler characteristic curves and profiles. Statistical consistency tests in topological analysis.	W4, W6, U2, U3, U6, K1	Laboratories
7.	**Discrete Morse theory and its applications** Introduction to discrete Morse theory (DMT). Relating DMT to filtering and persistent homology. Iterated Morse complexes as a tool for calculating persistent homology.	W3, W4, U3, U6, K1	Laboratories
8.	**Topological methods of data visualization** Reeb graphs, cover complexes and mapper algorithms. Interpretation of graphs of relations (A, f(A)), where A is a subset sampled from a high-dimensional space. Discussion of the standard mapper and Ball Mapper algorithm, as well as ClusterGraph.	W5, U5, U6, K3	Laboratories
9.	**Homotopy groups and their application in data analysis** Computational homotopy groups and their relationship to group representation. Discrete Morse theory as a tool for simplifying homotopy representations.	W4, W6, U2, U3, U6, K1	Laboratories
10.	**Geometric estimators in data analysis** Estimation of Riemannian metrics on manifolds. Determination of dimensionality, curvature and range from point clouds.	W3, U3, U6, K1	Laboratories
11.	**Dynamic systems and their relationship to topology** Application of topological methods in the analysis of dynamical systems. Wazewski's theorem, Conley's index and their use in stability analysis of dynamical systems.	W6, U6, K1	Laboratories
12.	**Applications of topological data analysis in various fields** Analysis of brain structure, classification of neuronal shapes, classification of materials. Lung structure in chronic obstructive pulmonary disease (COPD). Applications in economics and political science, analysis of financial markets.	W6, U6, U8, K2, K3, K4	Laboratories

Additional information

Activities	Teaching and learning methods and activities
Laboratories	Classes method, Laboratory method, Project method, Work in groups

Activities	Credit conditions
Laboratories	The credit from the lecture part will be given on the basis of a project which you are supposed to select till the end of first month. The presentation of your projects will be held on during classes on a fixed day. You can get maximal 100 points for the project. In order to get the maximal number of points you are required to: • Present your project (up to 100 points.) • Pass the Exercises/LAB sessions with a positive mark (details given below). During the exercises and LAB sessions to Invitation to Topological Data Analysis you can obtain 110 points (110%). In order to get the maximal number of points, you are required to: • Present the mini project (up to 50 points) • Solve homeworks (up to 50 points): you will be given five sets of exercises. In order to get the perfect score on that part, approximately three sets are required to be solved. • Active participation in the classes (up to 10 points, maximal one point per a 90 min meeting). The participation in exercises/LABs is compulsory and it is allowed to have three justified absences. Grade scale: 50% dst (3), 60% dst+ (3.5), 70% db (4), 80% db+ (4.5), 90% bdb (5).

Literature

Obligatory

1. Herbert Edelsbrunner and John Harer, Computational Topology: An Introduction, AMS 2011. The book provides a comprehensive introduction to computational topology, combining mathematical and algorithmic aspects, making it ideal for students and researchers interested in this rapidly developing field.
2. Paweł Dłotko, Applied and Computational Topology Tutorial, 2018. (https://arxiv.org/abs/1807.08607) This tutorial offers a practical introduction to topology applications in data analysis, illustrated with numerous examples using the GUDHI library.
3. Gunnar Carlsson, Topology and Data, Bulletin of the American Mathematical Society, 2009. The article presents the application of topology in data analysis, highlighting how topological tools can reveal hidden structures in complex data sets.

Optional

1. K. Mischaikow, T. Kaczynski, M. Mrozek, Computational Homology, Springer 2004. The book focuses on computational homology, offering both theoretical foundations and practical algorithms, making it a valuable resource for those interested in topological data analysis.
2. Biblioteka GUDHI. (https://gudhi.inria.fr/) GUDHI is an open-source C++ library with a Python interface, offering advanced data structures and algorithms for constructing symplectic complexes and calculating persistent homology, providing an invaluable tool in topological data analysis.

Calculation of ECTS points

Activities	Activity hours*
Laboratories	60

Preparation for classes	20
Preparation of a project	25
Preparation for the assessment	20
Reading the indicated literature	25
Student workload	**Hours** 150
Number of ECTS points	**ECTS** 5

* academic hour = 45 minutes

Efekty uczenia się dla kierunku

Kod	Treść
MAT_K1_K01	The graduate is ready to uznania ograniczenia własnej wiedzy i zrozumienia potrzeby dalszego kształcenia
MAT_K1_K02	The graduate is ready to precyzyjnego formułowania pytań służących pogłębieniu zrozumienia danego zagadnienia
MAT_K1_K03	The graduate is ready to wyjaśniania znaczenia matematyki i jej osiągnięć
MAT_K1_K04	The graduate is ready to samodzielnego wyszukiwać informacje w literaturze i bazach danych
MAT_K1_U01	The graduate can przedstawiać treści matematyczne w mowie i w piśmie, formułować twierdzenia i definicje
MAT_K1_U02	The graduate can objaśniać, interpretować złożone wypowiedzi z użyciem matematycznej notacji i języka oraz formułować problemy w postaci symbolicznej, ułatwiającej ich analizę i rozwiązanie
MAT_K1_U04	The graduate can posługiwać się narzędziami i aparatem logiki matematycznej, teorii mnogości z uwzględnieniem algebry zbiorów, rachunku kwantyfikatorów, relacji porządkujących i relacji równoważności w poznanych działach matematyki oraz w innych dziedzinach wiedzy
MAT_K1_U05	The graduate can tworzyć nowe obiekty drogą standardowych konstrukcji, zwłaszcza przestrzeni ilorazowych i produktów kartezjańskich
MAT_K1_U07	The graduate can definiować, interpretować, opisywać i wyjaśniać zależności funkcyjne, wyrażone w postaci wzorów, tabel, wykresów, schematów i stosować je w zagadnieniach praktycznych
MAT_K1_U08	The graduate can posługiwać się narzędziami i aparatem teorii funkcji rzeczywistych jednej i wielu zmiennych oraz funkcji zmiennej zespolonej z uwzględnieniem rachunku granic, pochodnych i całek, a także stosować je w poznanych działach matematyki oraz w innych dziedzinach wiedzy
MAT_K1_U09	The graduate can rozwiązywać podstawowe typy równań różniczkowych i ich układy
MAT_K1_U10	The graduate can posługiwać się narzędziami i aparatem teorii liczb, algebry liniowej i abstrakcyjnej, z uwzględnieniem klasycznych struktur algebraicznych, takich jak grupy, pierścienie i ciała, oraz geometrii i topologii
MAT_K1_U12	The graduate can posługiwać się narzędziami i aparatem matematyki dyskretnej, teorii algorytmów i metod numerycznych, ze szczególnym uwzględnieniem związków z informatyką
MAT_K1_W01	The graduate knows and understands cywilizacyjne znaczenie matematyki i jej zastosowań
MAT_K1_W02	The graduate knows and understands rolę i znaczenie dowodu w matematyce, a także istotność założeń
MAT_K1_W05	The graduate knows and understands podstawowe pojęcia i metody logiki matematycznej oraz teorii mnogości, a także podstawowe pojęcia, reguły i twierdzenia analizy matematycznej, w tym rachunku różniczkowego i całkowego funkcji rzeczywistych jednej i wielu zmiennych, algebry liniowej i abstrakcyjnej, geometrii i topologii oraz matematyki dyskretnej
MAT_K1_W07	The graduate knows and understands podstawowe pojęcia z zakresu etyczno-prawnych aspektów ochrony własności intelektualnej, pracy naukowej i dydaktycznej