Anomaly Detection
Introduction

Paul Irofti
Cristian Rusu
Andrei Pătraşcu

Computer Science Department
University of Bucharest

Professors

Course:

- Paul Irofti
- Cristian Rusu
- Andrei Pătraşcu

Lab sessions:

- Andrei Hîji

Course organization

Grading system:

- Lab sessions - 40%
- Research paper presentation - 60%

For the lab sessions:

- attendance is mandatory
- lab professor grades your activity and completion of the lab tasks

Lab organisation

Lab **40 points** – every two weeks

Points

- **minimum 20** points for exam entry

Presence

- **mandatory presence** at all labs
- without attending you can not take the exam
- if you skip one lab you can retake it in the other week

Retakes, reexamination

- lab can only be promoted during the semester
- you can only retake the labs during the first semester during labs hours
- you can not retake the exam before the reexamination or in the second semester

Exam Organization

Details for the presentation:

- individual work
- you will pick a research paper related to AD
- paper choice has to be confirmed with us before proceeding
- code up the idea and test it
- write a report about the method and conclusions
- give a 10 min. presentation

Final Grade

Grading system:

- Lab sessions - 40%
- Research paper presentation - 60%

Passing the course:

- minimum half for the lab sessions
- minimum half on the final research paper presentation

Course references

Main reference: Charu C. Aggarwal. *Outlier Analysis.* Springer (2013).

Secondary referencens:

- Van Loan, Charles F., and G. Golub. *Matrix computations (Johns Hopkins studies in mathematical sciences).* (1996).
- Trevor Hastie, Jerome H. Friedman, and Robert Tibshirani. *Elements of Statistical Learning.* Springer (2008).
- Deisenroth, Marc Peter, A. Aldo Faisal, and Cheng Soon Ong. *Mathematics for machine learning*, Cambridge University Press, 2020
- Charu C. Aggarwal, and Saket Sathe., *Outlier Ensembles.* Springer (2017).

Curricula

A. Basic (linear) algorithms
B. Distance based: OC-SVM, SVDD
C. Tree based: Isolation Forest
D. Statistical algorithms: truncation, LODA
E. Density based: k-NN, LOF
F. Dimensionality reduction: PCA, robust PCA, Autoencoders
H. LLMs
G. Applications: timeseries, graphs, networks