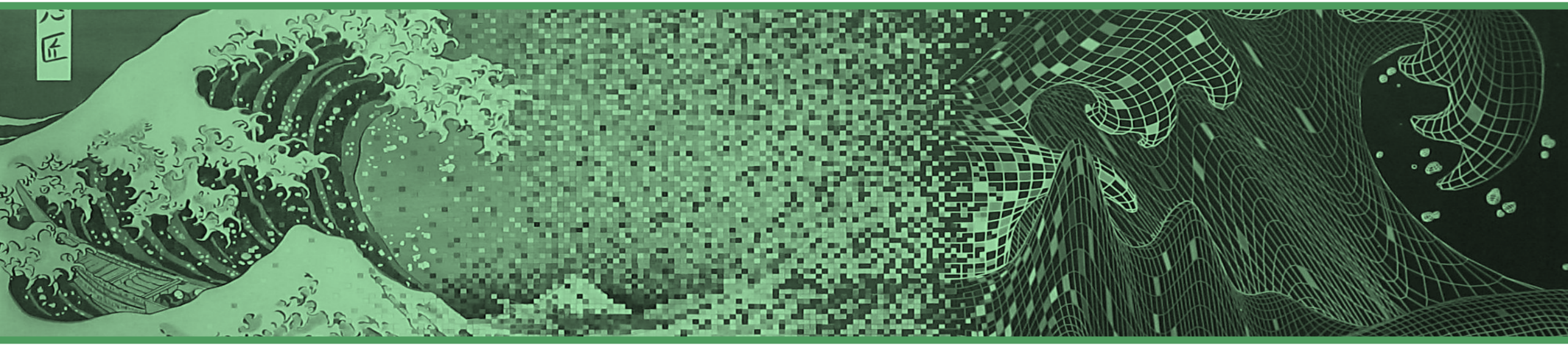
TÉCNICO LISBOA

Course organization

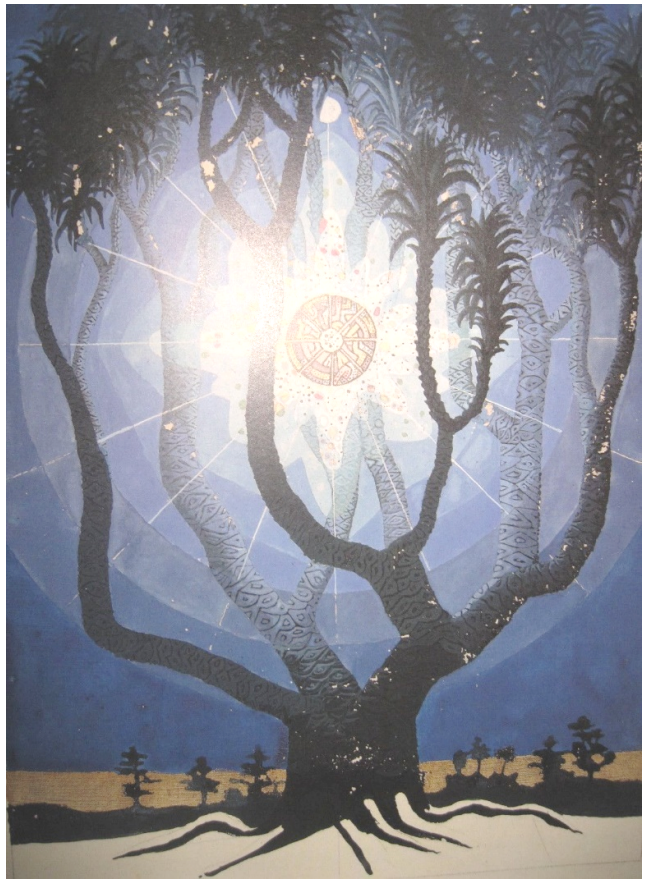
course objectives, assessment and plan



Aprendizagem 2024

Miguel Couceiro miguel.j.couceiro@tecnico.ulisboa.pt
Andreas Wischert andreas.wichert@tecnico.ulisboa.pt

Outline



- Faculty hosts
- Objectives
- Planning
- Bibliography
- Evaluation
- Homeworks
- Schedule
- Lab attendance
- Office hours

Faculty hosts

– **Miguel Couceiro** (*responsible* @ Alameda)
  miguel.j.couceiro@tecnico.ulisboa.pt
  miguel.couceiro@loria.fr (alternative)
  Lectures [PT/EN] @ Alameda
  Office: Room TBC, INESC-ID **or** Room 4.25, Dep. Math

– **Andreas (Andrzej) Wichert** (*responsible* @ Tagus)
  andreas.wichert@tecnico.ulisboa.pt
  Lectures [EN] @ Tagus
  Office: Room N2 5-7, Taguspark

Faculty hosts: *practical classes*

- **Miguel Moreira**
  miguelmatamoreira@tecnico.ulisboa.pt
- **Inês Magessi**
  ines.magessi@tecnico.ulisboa.pt
- **Gonçalo Correia**
  goncalo.m.correia@tecnico.ulisboa.pt
- **Javier de Muller Santa-Maria**
  javier.de.muller@tecnico.ulisboa.pt

Objectives

- Answer real-world data-rich problems: **formalize** and **translate them into ML tasks**
- Understand **major topics on ML**
  - master the **statistical** and mathematical **foundations** behind ML
  - critically compare the behavior of ML approaches
    - **Bayesian learning**
    - **lazy** and **associative learning**
    - **neural network learning**
- Master **supervised ML** approaches, including classification and regression
- Master **unsupervised ML** approaches, including clustering and data transformations
- Robustly **evaluate ML solutions** (loss, statistical significance, generalization ability)

Planning

1. Introduction to ML	I
2. Univariate ML	
3. Associative learning	
4. Evaluation	
5. Bayesian learning	II
6. Lazy learning	
7. Kernel learning	
8. Regression learning	
9. Gradient descent	III
10. Neural network learning	
11. Deep Learning	
12. Clustering	IV
13. Dimensionality reduction	

Theoretical background

- Information theory (2,3)
- Probability theory (5,12)
- Algebra theory (6,7,13)
- Optimization theory (8-11)

Foundations
Supervised learning
Unsupervised learning

Practical classes

- **L1** Univariate statistics, evaluation
- **L2** Associative learning (decision trees)

⇒ **HW1**

- **L3** Bayesian learning
- **L4** Lazy learning
- **L5** Linear regression

⇒ **HW2**

- **L6/7** Gradient descent learning
- **L8/9** Neural network learning

⇒ **HW3**

- **L10** Clustering
- **L11** Dimensionality reduction

⇒ **HW4**

Main literature

- **reference slides** @ course's webpage

- **Machine Learning: A Journey to Deep Learning**
  Andreas Wichert and Luis Sá Couto, World Scientific, 2021
  https://www.worldscientific.com/worldscibooks/10.1142/12201

- **Machine Learning**
  Tom Mitchell, McGraw Hill, 1997
  http://www.cs.cmu.edu/~tom/mlbook.html

- **Data Mining and Machine Learning: Fundamental Concepts and Algorithms**
  Mohammed J. Zaki and Wagner Meira Jr, Cambridge Univ. Press, 2nd Ed., 2020
  https://dataminingbook.info/

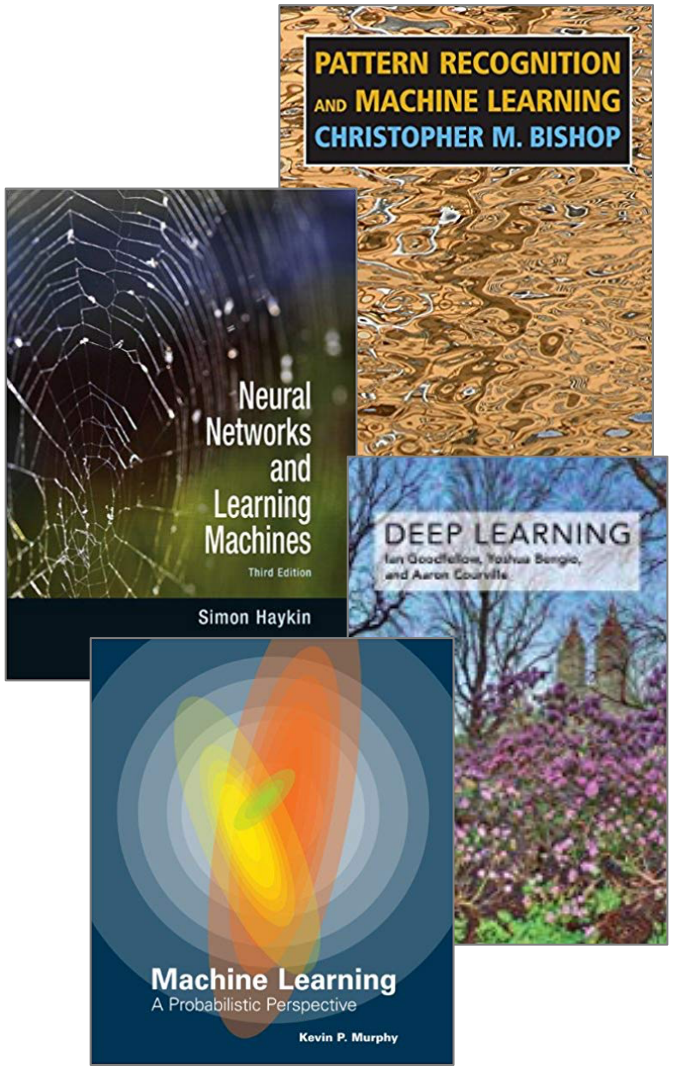
Secondary literature

- **Pattern Recognition and Machine Learning**
  Christopher M. Bishop, Springer 2006
  https://www.microsoft.com/en-us/research/people/cmbishop/#!prml-book

- **Neural Networks and Learning Machines**
  Simon O. Haykin, (3rd Edition), Pearson 2008

- **Deep Learning**
  I. Goodfellow, Y. Bengio, A. Courville, MIT Press 2016
  https://www.deeplearningbook.org

- **Machine Learning: A Probabilistic Perspective**
  K. Murphy, MIT Press 2012



Tertiary literature

- *Software literature*
  - **Hands-On Machine Learning with Scikit-Learn and TensorFlow: Concepts, Tools, and Techniques to Build Intelligent Systems**

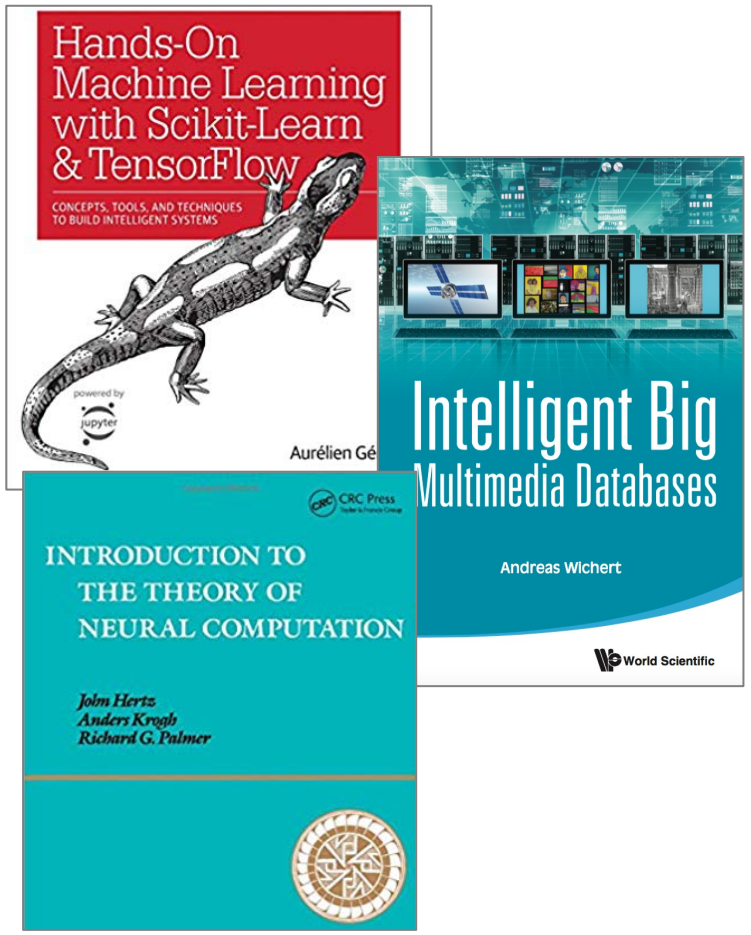
    Aurélien Géron , O'Reilly Media, 2017

    https://github.com/amitanalyste/aurelienGeron

- *Complementary*
  - **Intelligent Big Multimedia Databases**

    A. Wichert, World Scientific, 2015
  - **Introduction to the theory Of Neural Computation**

    J.A. Hertz, A.S. Krogh, R.G. Palmer, Addison-Wesley, 1991



Evaluation

Final grade: Homeworks × 50% + Exam × 50%

Homeworks:

- average(HW1,HW2,HW3,HW4)
- groups of **2** (no exceptions)

Exam: maximum(Exam1,Exam2)

Minimum exam grade: 8v

Working students

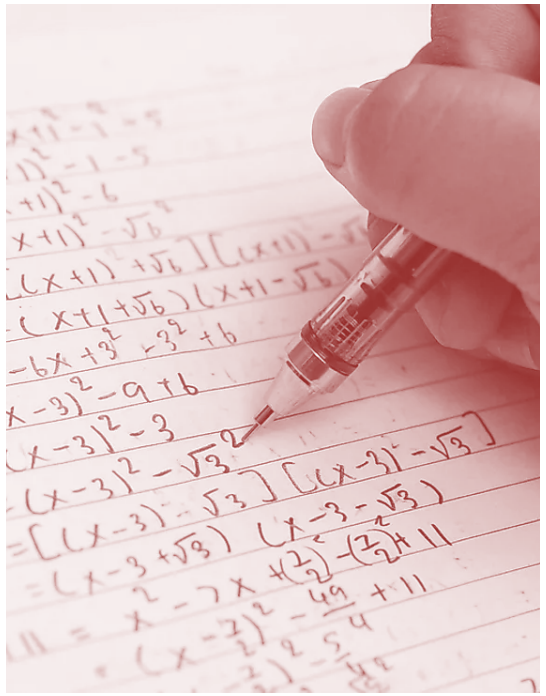
- (Homeworks × 50% + Exam × 50%) **or** 100% Exam
- decision communicated until September 22nd

Special season: 100% Exam

Deadlines

- **Homework deliveries**
  - HW1: Friday, September 27th 23h59
  - HW2: Monday, October 7th 23h59
  - HW3: Friday, October 18th 23h59
  - HW4: Monday, October 28th 23h59
  - *statements* released 10 days before delivery
- **Exams**
  - Exam1: November 4, 2024 (13h-15h)
  - Exam2: February 6, 2025 (10h30-12h30)
  - Special session: July 24, 2025 (13h-15h)

Homeworks



- groups of **2** (*no exceptions*)
- statements at the course webpage two weeks prior delivery
- always consult webpage **FAQ** before posting questions
- three major **components**
  - **practical** pen-and-paper **quest**
  - small **computational quest**
  - **critical thinking**
    - excelling answers are clear, precise and succinct, and establish scientific hypothesis grounded on empirical evidence

Final exam

– **extensive** number of **contents** to prepare

  – requiring a good mastery: *critical analysis* and *practical calculus*

– comprehensive exam with a **2h duration**

– you are allowed to bring a **single** A4 **page with notes** (e.g. formulas)

– carefully follow theoretical classes to minimize risks

– objective assessment: quantitative result answers, true-or-false

Merit award:
NTT Data in Learning

- Participation is not mandatory (no impact on course grade!)
- HW groups (2 students per group!)
- Groups will be ranked according to

  (i) the performance on a challenge in a HW, and

  (ii) the group oral presentation of results
- Requirements:

  1. successful HW assignments and
  2. a grade of at least 10 in course exam(s)
- Prize: 2500€ to the winning group

Schedule @ Alameda

Horas	Segunda		Terça				Quarta	Quinta				Sexta	
08:00	T1							T1					
08:30		P3							P4			P8	P9
09:00				P6	P4	P5							
09:30										P10			
10:00	T2												
10:30				P8	P9	P7							
11:00													
11:30			T2										
12:00									P3	P7		P6	P10
12:30											P5		
13:00													
13:30													
14:00													

provisory schedule

Planning

	TEÓRICAS	PRÁTICAS	deadlines	
09-Sep	1 Introduction (ML, IA) 2		Lab reg: Set 23, 15h	
	3 Associative learning 4			
16-Sep	5 Evaluation 6	1 Foundations		
	7 Bayesian learning 8	2 Decision trees / Evaluation		
23-Sep	9 Lazy and kernel learning 10	3 Naïve Bayes	Ficha 1 (Sep 27)	Prob, DTrees, Eval
	11 Regression 12	4 kNN / Evaluation		
30-Sep	13 Perceptron/SGD 14	5 Regression		
	15 Neural network learning 16	6 Perceptron/GD		
07-Oct	17 Neural network learning 18	7 GD/Neural networks	Ficha 2 (Oct 7)	NBayes, kNN
	19 DL and Large language models 20	8 Neural networks		
14-Oct	21 Clustering 22	9 Neural networks/Clustering	Ficha 3 (Oct 18)	MLPs, regression
	23 Dimensionality reduction 24	10 Clustering		
21-Oct	25 Wrap-up 26	11 PCA		
	27 28	12 Wrap-up		
28-Oct			Ficha 4 (Oct 28)	clustering, dim red

Office hours

- Explore your questions during **practical classes** and **lectures**
- **Check webpage for latest information on office hours**
  - *Miguel Couceiro* office hours
    - Monday 12h, Tuesday 9h30, Thursday 10h (end of lectures)
    - additional office hours available via appointment
  - office hours of other faculty hosts on the webpage
- **Golden rule**
  - general e-mail contacts with subject preceded by **[Aprendizagem]**
  - homework-related questions preceded by **[HW Aprendizagem]**

Thank You

miguel.j.couceiro@tecnico.ulisboa.pt

andreas.wichert@tecnico.ulisboa.pt