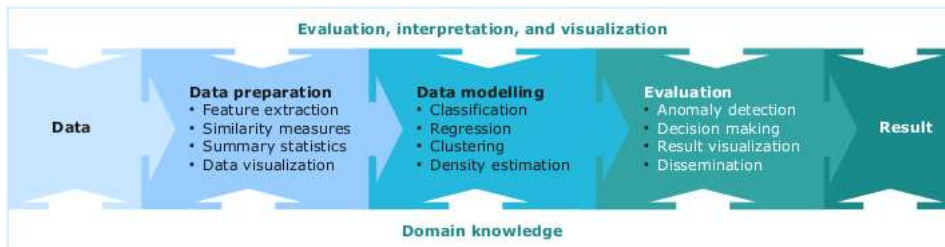




02450 Introduction to Machine Learning and Data Mining

Machine learning and data mining

The course is designed around a data modeling framework shown in the figure. Each lecture/assignment will focus on an aspect of the data modeling framework.



We emphasize the holistic view of modeling in order to motivate and stress the relevance of individual components and building blocks, disseminate the obtained competence (see the course [learning objectives](#)), and make them applicable for a broad spectrum of engineering problems in e.g. biomedical engineering, chemistry, electrical engineering, and informatics.

Resources

Location

The lectures will take place in building 303A auditorium 41 and 42 Tuesdays from 13:00-15:00 followed by exercises in

Python: Building 308 rooms 101, 109, 117, 1st floor area north and south

Matlab: Building 210 rooms 162, 168 as well as building 303A auditorium 42

R: Building 308 room 127

from 15:00-17:00. Please bring a laptop computer for the exercises.

Reading material, lecture slides and exercises

The course will use lecture notes and other freely available material. Lecture notes, slides, course assignment instructions etc. is available at the [DTU Campusnet](#) course page (requires formal enrolment to the course).

Course description

A description of the course can be found at the [DTU Coursebase](#)

Online help and support

Online help and support is available through the [Piazza](#) course platform.

Teacher

∴ **Morten Mørup (MM)**, mmor@dtu.dk

Lecture schedule

No.	Date		Subject	Preparation
1	5 September, 2017	MM	Introduction	C1
Data: Feature extraction, and visualization				
2	12 September, 2017	MM	Data and feature extraction	C2, C3. (P3.1, P2.1, P3.2)
3	19 September, 2017	MM	Measures of similarity and summary statistics	C4. (P4.1, P4.2, P4.3)
4	26 September, 2017	MM	Data Visualization and probability	C5, C6. (P5.1, P5.2, P6.1)
Supervised learning: Classification and regression				
5	3 October, 2017	MM	Decision trees and linear regression (Hand in project 1 before 13:00)	C7, C8. (P8.1, P7.1, P7.2)
6	10 October, 2017	MM	Overfitting and performance evaluation Holiday	C9. (P9.1, P9.2, P9.3)
7	24 October, 2017	MM	Nearest Neighbor, Bayes and Naive Bayes	C10, C11. (P11.1, P11.2, P10.1)
8	31 October, 2017	MM	Artificial Neural Networks and Bias/Variance	C12, C13. (P13.1, P13.2, P13.3)
9	7 November, 2017	MM	AUC and ensemble methods	C14, C15. (P14.1, P14.2, P15.1)
Unsupervised learning: Clustering and density estimation				
10	14 November, 2017	MM	K-means and hierarchical clustering (Hand in project 2 before 13:00)	C16. (P16.1, P16.2, P16.3)
11	21 November, 2017	MM	Mixture models and density estimation	C17, C18. (P18.1, P17.1, P17.2)
12	28 November, 2017	MM	Association mining	C19. (P19.1, P16.2, P16.3)
Recap				
13	5 December, 2017	MM	Recap and discussion of the exam (Hand in project 3 before 13:00)	C1-C19

(**Cx** refers to Chapter **x** of the course notes. **Px.y** refers to problem number **y** in chapter **x** of the course notes. The first listed problem will be that weeks discussion question at the exercises.)



Morten Mørup



Tue Herlau



Mikkel N. Schmidt



Florian Gawrilowicz



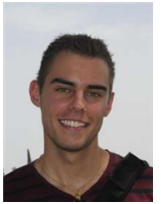
Sayantan Sengupta



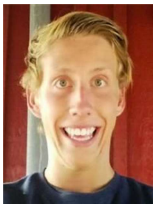
Rasmus M. T. Høegh



Malte K. E. Jensen



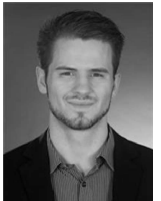
Sebastian Mira



Mads O. Jakobsen



Andreas Munk



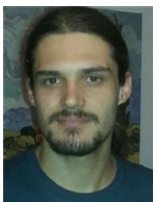
Benjamin Jüttner



Frederik B. Hüttel



Quoc Tien Au



Lorenzo Belgrano



Paolo A. Mesiano

