Introduction to Machine Learning (acronym IML)

Dr. Maria Salamó
Dept. Mathematics and Informatics,
Faculty of Mathematics and Informatics,
University of Barcelona

UNIVERSITAT DE BARCELONA

Brief description

- This course provides an **introduction on Machine Learning**

- It gives an overview of many concepts, techniques and algorithms in machine learning, beginning with topics such as **classification** and **linear regression** and ending up with more recent topics such **support vector machines** and **recommender systems**

Brief description

- The course is divided into three main topics:
  - supervised learning, unsupervised learning, and machine learning theory

- Topics include:
  - (i) **Supervised learning** (linear decision, non linear decision)
  - (ii) **Unsupervised learning** (clustering, factor analysis, visualization)
  - (iii) **Learning theory** (bias/variance theory, empirical risk minimization)

- The course will also draw from numerous case studies and applications, so that you'll also learn how to apply learning algorithms to computer vision, medical informatics, and signal analysis

Summary

Introduction to Machine Learning

Unsupervised Learning

- **Cluster Analysis**
  - K-Means, Fuzzy C-means, EM
- **Factor Analysis**
  - PCA, ICA
- **Visualization**
  - Self Organized Maps (SOM) , Multi-Dimensional Scaling

Supervised Learning

- **Non Linear Decision**
  - Lazy Learning (K-NN, IBL, CBR)
  - Overfitting, model selection and feature selection
  - Kernel Learning
  - Ensemble Learning (Trees, Adaboost)
- **Linear Decision**
  - Perceptron, SVM

Decision Learning Theory

- **Basic concepts of Decision Learning Theory**
  - Bias/Variance, VC dimension, Practical advice of how to use learning algorithms

UNIVERSITAT DE BARCELONA

Contents

Contents

UNIVERSITAT DE BARCELONA

Contents

Contents

9. Model Selection
   a) Introduction to Model evaluation (performance metrics, confusion matrix, ROC curves, etc.)
   b) Model evaluation (hold-out, cross-validation, overfitting and underfitting, bias vs variance, regularization, etc.)
10. Kernel Learning
    a) Statistical learning theory
    b) Support Vector Machines (SVM)

Contents

11. Recommender Systems
    a) Introduction to recommendation techniques
    b) An overview of Collaborative Filtering
    c) An overview of Content-based Filtering
    d) Conversational Recommenders
12. Ensemble Learning
    a) Introduction to ensemble learning
    b) Additive model: Bagging
    c) Additive model: Boosting

Methodology

- The class is divided in two parts
  - Theory (2 hours): introduce the contents of the course
  - Laboratory (1 hour) which includes:
    - Practical exercises related to work deliveries
    - Participatory class where students talk about the readings suggested to go deeper into a subject

      Note: These readings will be included as theory in the final exam

Activities – work deliveries

- Work 1 – (W1)
  - Clustering exercise
- Work 2 – (W2)
  - Factor Analysis exercise
- Work 3 – (W3)
  - Lazy Learning or kNN Recommenders
- Work 4 – (W4)
  - Kernels exercise

This year work is done in groups of three

UNIVERSITAT DE BARCELONA

Evaluation

- The course is divided into two parts:
  - **Exam**: an exam at the end of the term
  - **Work**: Work deliveries during the semester

Mark = a x Exam + b x Work if exam>=3,5 and Work>= 4,5

This year **a** and **b** will be established as: *a = 0,4 and b = 0,6*

Exam = an exam at the end of the term (14th January 2020)

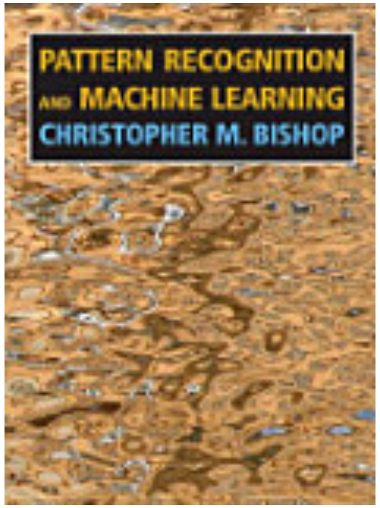
Work = 0,3 x W1 + 0,2 x W2 + 0,3 x W3 + 0,2 x W4

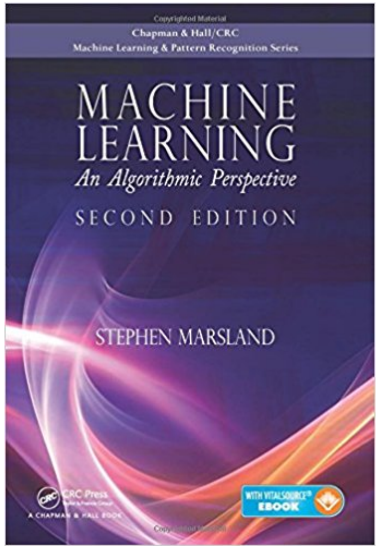
- Some Works may have an exam after the delivery. The mark will be a part of the Wx.

Remarks

- All the information of the course will be in Racó https://**raco.fib.upc.edu/**
- The **schedule** of the course is in another file, look at the notes in racó
- Work is in groups of 3 but the score is individual for each student
- **Late deliveries**: work or projects submitted late will mean the deduction of 1 point per day (out of 10) from the final mark
- **Copy of deliveries**: all the groups involved will obtain 0 points

Bibliography



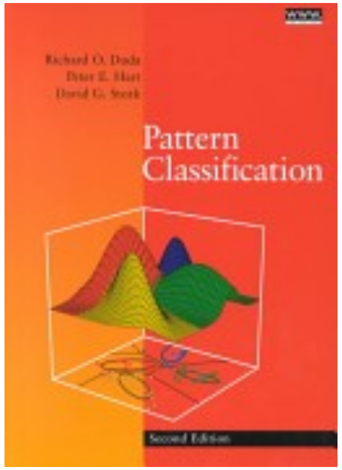


- Bishop, Christopher M., *Pattern Recognition and Machine Learning*. Springer. p. 738. ISBN 978-0-387-31073-2.

- Marsland, Stephen, *Machine Learning: An algorithmic Perspective*, 2nd ed. CRC Press, 2015. ISBN: 978-1-466-58328-3

UNIVERSITAT DE BARCELONA

Bibliography



- Duda, Richard; Hart, Peter; and Stork, David, *Pattern Classification*, 2nd ed. John Wiley&Sons, 2001. ISBN: 978-0-471-05669-0

- Tom Mitchell, *Machine Learning*. McGraw-Hill. ISBN 0-07-042807-7

Tools

- Python

  https://www.python.org/

  http://scikit-learn.org/

WEKA The University of Waikato

- Weka

  http://www.cs.waikato.ac.nz/ml/weka/

- PyCharm

  https://www.jetbrains.com/pycharm/