



# Course guide

## 295601 - AB - Biostatistical Learning

Last modified: 02/06/2022

**Unit in charge:** Barcelona East School of Engineering  
**Teaching unit:** 749 - MAT - Department of Mathematics.

**Degree:** BACHELOR'S DEGREE IN BIOMEDICAL ENGINEERING (Syllabus 2009). (Optional subject).

**Academic year:** 2022    **ECTS Credits:** 6.0    **Languages:** Spanish

### LECTURER

---

**Coordinating lecturer:** JOSE JULIAN RODELLAR BENEDE

**Others:** Segon quadrimestre:  
ANDREA MILENA ACEVEDO LIPES - M11  
FRANCESC POZO MONTERO - M11  
JOSE JULIAN RODELLAR BENEDE - M11  
YOLANDA VIDAL SEGUI - M11

### PRIOR SKILLS

---

Programming course. Basic statistical concepts and tools.

### REQUIREMENTS

---

None

### TEACHING METHODOLOGY

---

The course is developed through four hours per week: half correspond to theoretical lectures and half to laboratory sessions, supervised works and a final exam.

The teaching activities are distributed as follows:

- Theory lectures: 20%
- Computer sessions and works: 20%
- Autonomous learning: 60%

### LEARNING OBJECTIVES OF THE SUBJECT

---

Objectives proposed for the course:

- Understand the basic theory of machine learning
- Formulate statistical learning problems related to different biomedical applications
- Understand a wide range of statistical learning algorithms along with their advantages and limitations
- Implement statistical learning algorithms to solve biomedical problems with moderate complexity
- Compare the performance of several techniques and recommend those that better fit to the proposed problems



## STUDY LOAD

Type	Hours	Percentage
Hours large group	30,0	20.00
Hours small group	30,0	20.00
Guided activities	90,0	60.00

**Total learning time:** 150 h

## CONTENTS

### 1. Introduction

**Description:**

What is statistical learning. Regression and classification problems. Training and validation. Performance measures. Bayes classifier.

Bias/variance tradeoff. Software: Python.

**Specific objectives:**

.

**Related activities:**

Theory lectures 1 and 2

Lab 1: Introduction to Python, Numpy and Pandas

**Full-or-part-time:** 12h

Theory classes: 4h

Laboratory classes: 2h

Self study : 6h

### 2. Linear regression

**Description:**

Linear regression model. Least squares. Statistical significance.

**Specific objectives:**

.

**Related activities:**

Theory lecture 3

Lab 2: linear regression with Python

**Full-or-part-time:** 8h

Theory classes: 2h

Laboratory classes: 2h

Self study : 4h

### 3. Classification

**Description:**

Logistic regression. Linear discriminant analysis. Bayes theorem for classification. Confusion matrix. Quadratic discriminant analysis.

**Specific objectives:**

.

**Related activities:**

Theory lectures 4 and 5

Lab 3: logistic regression and linear discriminant analysis

Work 1

**Full-or-part-time:** 21h

Theory classes: 4h

Laboratory classes: 2h

Guided activities: 2h

Self study : 13h

### 4. Re-sampling methods

**Description:**

Cross validation. Bootstrap

**Specific objectives:**

.

**Related activities:**

Theory lecture 6

Lab 6: Cross validations and bootstrap

**Full-or-part-time:** 8h

Theory classes: 2h

Laboratory classes: 2h

Self study : 4h

### 5. Linear model selection and regularization

**Description:**

Variable and predictor selection.

Dimension reduction. Principal component analysis. Partial least squares.

**Specific objectives:**

.

**Related activities:**

Theory lectures 7 and 8

Lab 5: regression and nearest neighbours method

Lab 6: dimension reduction

**Full-or-part-time:** 16h

Theory classes: 4h

Laboratory classes: 4h

Self study : 8h

## 6. Tree-based methods

### Description:

Decision trees. Regression and classification. Pruning. Bagging. Random forest. Boosting. Multiple classifiers.

### Specific objectives:

.

### Related activities:

Theory lectures 9 and 10  
Lab 7: Tree methods for classification and regression  
Work 2

### Full-or-part-time: 21h

Theory classes: 4h  
Laboratory classes: 2h  
Guided activities: 2h  
Self study : 13h

## 7. Support vector machines

### Description:

Maximal margin classifier. Support vector machines (SVM). Classification of more than two classes.

### Related activities:

Theory lectures 11 and 12  
Lab 8: Applications of SVM  
Work 3

### Full-or-part-time: 20h

Theory classes: 4h  
Laboratory classes: 2h  
Guided activities: 2h  
Self study : 12h

## 8. Neural networks

### Description:

Architecture. Training. Error propagation.

### Specific objectives:

.

### Related activities:

Theory lectures 13 and 14  
Lab 9: Implementation of neural networks  
Work 4

### Full-or-part-time: 20h

Theory classes: 4h  
Laboratory classes: 2h  
Guided activities: 2h  
Self study : 12h



## 9. Non-supervised learning

### Description:

Clustering methods. K-means. Hierarchical clustering.

### Specific objectives:

.

### Related activities:

Theory lecture 15  
Lab 10: Applications of clustering methods  
Exam

### Full-or-part-time: 24h

Theory classes: 16h  
Laboratory classes: 2h  
Guided activities: 2h  
Self study : 4h

## GRADING SYSTEM

Continuous evaluation along the course by means of practical works. By the end of the course the student will pass a final complementary exam.

## EXAMINATION RULES.

.

## BIBLIOGRAPHY

### Basic:

- Bishop, Christopher M. Pattern recognition and machine learning. Springer, 2006. ISBN 9780387310732.
- Géron, Aurélien. Hands-on machine learning with scikit-learn & tensorflow : concepts, tools, and techniques to build intelligent systems [on line]. Sebastopol, CA: O'Reilly Media, Inc, 2017 [Consultation: 15/04/2020]. Available on: <https://ebookcentral.proquest.com/lib/upcatalunya-ebooks/detail.action?docID=4822582>. ISBN 9781491962268.
- James, G.; Witten, D.; Hastie, T.; Tibshirani, R. An introduction to statistical learning with applications in R. Springer, 2013. ISBN 9781461471370.
- Raschka, Sebastian. Python machine learning : machine learning and deep learning with Python, scikit-learn, and TensorFlow [on line]. 2nd ed. Birmingham, UK: Packt Publishing Ltd, 2017 [Consultation: 14/04/2020]. Available on: <https://ebookcentral.proquest.com/lib/upcatalunya-ebooks/detail.action?docID=5050960>. ISBN 9781787126022].

### Complementary:

- Hastie, Trevor; Tibshirani, Robert; Friedman, Jerome. The Elements of statistical learning : data mining, inference, and prediction [on line]. 2nd ed. New York, NY: Springer Series in Statistics, 2001 [Consultation: 27/08/2018]. Available on: <http://dx.doi.org/10.1007/978-0-387-84858-7>. ISBN 9780387848587.

## RESOURCES

### Other resources:

Materials available in ATENEA by the instructors