ANX-PR/CL/001-01

LEARNING GUIDE

SUBJECT

93001070 - Predictive and descriptive learning

DEGREE PROGRAMME

09AQ - Master Universitario En Ingenieria De Telecomunicacion

ACADEMIC YEAR & SEMESTER

2018/19 - Semester 1

Index

Learning guide

1. Description

1.1. Subject details

Name of the subject	93001070 - Predictive and descriptive learning
No of credits	6 ECTS
Type	Optional
Academic year ot the programme	Second year
Semester of tuition	Semester 3
Tuition period	September-January
Tuition languages	English
Degree programme	09AQ - Master universitario en ingenieria de telecomunicacion
Centre	09 - Escuela Tecnica Superior de Ingenieros de Telecomunicacion
Academic year	2018-19

2. Faculty

2.1. Faculty members with subject teaching role

Name and surname	Office/Room	Email	Tutoring hours *
Eduardo Lopez Gonzalo (Subject coordinator)	C-330	eduardo.lopez@upm.es	Sin horario. Appointment arranged by email
Luis Alfonso Hernandez Gomez	C-330	luisalfonso.hernandez@upm.es	Sin horario. Appointment arranged by email

* The tutoring schedule is indicative and subject to possible changes. Please check tutoring times with the faculty member in charge.

3. Prior knowledge recommended to take the subject

3.1. Recommended (passed) subjects

El plan de estudios Master Universitario en Ingenieria de Telecomunicacion no tiene definidas asignaturas previas recomendadas para esta asignatura.

3.2. Other recommended learning outcomes

- Previous exposure to a programming language, such as MATLAB, R or Python

- Elementary course in Statistics

4. Skills and learning outcomes *

4.1. Skills to be learned

CG1 - Poseer y comprender conocimientos que aporten una base u oportunidad de ser originales en el desarrollo y/o aplicación de ideas, a menudo en un contexto de investigación.

CG2 - Que los estudiantes sepan aplicar los conocimientos adquiridos y su capacidad de resolución de problemas en entornos nuevos o poco conocidos dentro de contextos más amplios (o multidisciplinares) relacionados con su área de estudio.

CG4 - Que los estudiantes sepan comunicar sus conclusiones ?y los conocimientos y razones últimas que las sustentan? a públicos especializados y no especializados de un modo claro y sin ambigüedades.

CT1 - Capacidad para comprender los contenidos de clases magistrales, conferencias y seminarios en lengua inglesa.

CT2 - Capacidad para dinamizar y liderar equipos de trabajo multidisciplinares.

CT3 - Capacidad para adoptar soluciones creativas que satisfagan adecuadamente las diferentes necesidades planteadas.

CT4 - Capacidad para trabajar de forma efectiva como individuo, organizando y planificando su propio trabajo, de forma independiente o como miembro de un equipo.

CT5 - Capacidad para gestionar la información, identificando las fuentes necesarias, los principales tipos de documentos técnicos y científicos, de una manera adecuada y eficiente.

4.2. Learning outcomes

RA305 - Capability to design, develop and evaluate machine-learning techniques for a wide range of application areas

* The Learning Guides should reflect the Skills and Learning Outcomes in the same way as indicated in the Degree Verification Memory. For this reason, they have not been translated into English and appear in Spanish.

5. Brief description of the subject and syllabus

5.1. Brief description of the subject

This course covers the concepts and principles of a large variety of Machine Learning methods: from Deep Learning to more traditional Machine Learning models. The course introduces main principles in Machine Learning: supervised, unsupervised and reinforcement learning, though main emphasis is on predictive and descriptive learning as reinforcement learning is covered in a subsequent course. Methodological issues such as model assessment and selection, and overfitting are discussed.

The course starts introducing the most recent Deep Learning models: from the basics of Neural Networks to the most common architectures of Feed-Forward Networks, Convolutional Networks and Recurrent Neural Networks. Then different traditional predictive or supervised techniques are reviewed: as different types of regression, generalized linear models, k-nearest neighbor classifier, classification and regression trees, ensemble methods (Bagging, Random Forests and Boosting) and kernel methods and Support Vector Machines. The course also addresses traditional descriptive or unsupervised techniques: principal components analysis and clustering methods (k-means and hierarchical clustering). The course places special emphasis on presenting and discussing each technique through the analysis of practical use cases. Complementary use cases and experiments for large-scale scenarios are addressed in the Machine Learning Lab.

This course covers the principles and methodology for the design, evaluation and selection of a large variety of

Machine Learning methods for supervised and unsupervised learning.

The students will understand the fundamentals and important topics in statistical machine learning. This outcome represents a fundamental ingredient in the training of a modern data scientist providing a solid base for its use on a wide range of applications in science and industry. In particular students will understand the ideas behind the most used and widely applicable techniques for regression, classification and clustering. Through several examples and use cases, students will also learn how important is to accurately assess the performance of a model. They will also acquire solid criteria on what could be best model for a given data and task. By the end of the course, students should be able to:

- Understand the fundamentals of the most used models and techniques for predictive and descriptive learning.
- Design a proper methodology for accurately assessing and gaining knowledge from the use of each one of the particular machine learning techniques.
- Know the strengths and weaknesses of the various approaches in order to choose the best models for a given data and application scenario.

5.2. Syllabus

1. Introduction to Machine Learning

    1.1. What is statistical learning?

    1.2. Types of Machine Learning

    1.3. Assessing Model Accuracy

2. Introduction to Deep Learning

    2.1. Simple Neural Networks models

    2.2. Feed-forward Networks

    2.3. Convolutional Networks

    2.4. Recurrent Networks

    2.5. Introduction to advanced Deep Learning models: autoencoders, Generative Adversarial Networks (GANs)

3. Linear Regression

    3.1. Simple and Multiple Linear Regression

3.2. Linear Regression and Distributed Machine Learning Principles

3.3. Interpreting Regression Coefficients

3.4. Model Selection and Qualitative Predictors

3.5. Interactions and Nonlinearity

3.6. Comparison of Linear Regression with KNN

4. Classification

4.1. Logistic Regression

4.2. Bayes classifier and Linear Discriminant Analysis

4.3. Classification error analysis

4.4. Quadratic Discriminant Analysis

4.5. K-Nearest Neighbors

4.6. A Comparison of Classification Methods: Logistic Regression, LDA, QDA and KNN

5. Resampling methods

5.1. Cross-validation

5.2. Bootstrap

6. Linear Model Selection and Regularization

6.1. Feature selection

6.2. Optimal Model selection

6.3. Regularization

6.4. Dimension Reduction

6.5. High-Dimensional Data

7. Moving Beyond Linearity

7.1. Generalized Linear Models and Generalized Additive Models

8. Tree-Based Methods

8.1. Decision trees

8.2. Bagging

8.3. Random Forests

8.4. Boosting

9. Support Vector Machines

9.1. Maximal Margin Classifier

9.2. Support Vector Classifiers

9.3. Kernels and Support Vector Machines

9.4. Relationship to Logistic Regression

10. Descriptive Learning

10.1. Supervised vs Unsupervised learning

10.2. Principal Components Analysis

10.3. Clustering Methods

10.4. K-means

10.5. Hierarchical Clustering

10.6. Practical Issues in Clustering

6. Schedule

6.1. Subject schedule*

Week	Face-to-face classroom activities	Face-to-face laboratory activities	Other face-to-face activities	Assessment activities
1	**Activities Chapter 1** Duration: 04:00 Lecture			
2	**Activities Chapter 2 (2.1 y 2.2)** Duration: 04:00 Lecture			
3	**Activities Chapter 2 (2.3)** Duration: 04:00 Lecture			
4	**Activities Chapter 2 (2.4)** Duration: 04:00 Lecture			
5	**Activities Chapter 2 (2.5)** Duration: 04:00 Lecture			
6	**Activities Chapter 2 (review)** Duration: 02:00 Problem-solving class **Activities Chapter 3** Duration: 02:00 Lecture			
7	**Activities Chapter 3** Duration: 02:00 Lecture **Activities Chapter 4** Duration: 02:00 Lecture			
8	**Activities Chapter 4** Duration: 02:00 Lecture **Activities Chapter 5** Duration: 02:00 Lecture			
9	**Activities Chapter 6** Duration: 02:00 Lecture			**Evaluation: Deep Learning models** Individual presentation Continuous assessment Duration: 00:10
10	**Activities Chapter 7** Duration: 02:00 Lecture **Activities Chapter 8** Duration: 02:00 Lecture			

11	**Activities Chapter 9** Duration: 04:00 Lecture			
12	**Activities Chapter 9** Duration: 04:00 Lecture			
13	**Activities Chapter 10** Duration: 04:00 Lecture			
14	**Activities Chapter 10** Duration: 04:00 Lecture			
15				
16				
17				**Final project evaluation** Group presentation Continuous assessment Duration: 00:15 **Evaluation: Deep Learning Models** Individual presentation Final examination Duration: 00:10 **Final project evaluation** Group presentation Final examination Duration: 00:15

The independent study hours are training activities during which students should spend time on individual study or individual assignments.

Depending on the programme study plan, total values will be calculated according to the ECTS credit unit as 26/27 hours of student face-to-face contact and independent study time.

* The subject schedule is based on a previous theorical planning of the subject plan and might go to through experience some unexpected changes along throughout the academic year.

7. Activities and assessment criteria

7.1. Assessment activities

7.1.1. Continuous assessment

Week	Description	Modality	Type	Duration	Weight	Minimum grade	Evaluated skills
9	Evaluation: Deep Learning models	Individual presentation	Face-to-face	00:10	40%	/ 10	CT1 CG1 CT4 CG4 CT3 CG2 CT5
17	Final project evaluation	Group presentation	Face-to-face	00:15	60%	/ 10	CT1 CG1 CT4 CG4 CT3 CT2 CG2 CT5

7.1.2. Final examination

Week	Description	Modality	Type	Duration	Weight	Minimum grade	Evaluated skills
17	Evaluation: Deep Learning Models	Individual presentation	Face-to-face	00:10	40%	/ 10	CT1 CG1 CT4 CG4 CT3 CG2 CT5
17	Final project evaluation	Group presentation	Face-to-face	00:15	60%	/ 10	CT1 CG1 CT4 CG4 CT3 CT2 CG2 CT5

7.1.3. Referred (re-sit) examination

Description	Modality	Type	Duration	Weight	Minimum grade	Evaluated skills
Evaluation: Deep Learning models	Individual presentation	Face-to-face	00:10	40%	/ 10	CT1 CG1 CT4 CG4 CT3 CG2 CT5
Final project evaluation	Group presentation	Face-to-face	00:15	60%	/ 10	CT1 CT4 CG4 CT3 CT2 CG2 CT5 CG1

7.2. Assessment criteria

Students will be qualified through continuous evaluation by default. According to the Normativa de Evaluación del Aprendizaje de la Universidad Politécnica de Madrid, students willing to renounce to continuous evaluation must complete the Moodle task entitled "Renounce to continuous evaluation" before the fourth week of the semester (deadline will be announced in Moodle).

Evaluation will assess if students have acquired all the competences of the subject. Thus, evaluation through final assessment will be carried out considering all the evaluation techniques used in continuous evaluation (EX, ET, TG, etc.), and will be celebrated in the exam period approved by Junta de Escuela for the current academic semester and year. Evaluation activities that assess learning outcomes that cannot be evaluated through a single exam can be carried out along the semester.

Extraordinary examination will be carried out exclusively by the final assessment method.

Continuous assessment will consist of:

- Individual presentations to demonstrate skills in knowing the basics of deep learning models will be made by midsemester

(40% of final grade).

- A final collaborative project will be developed to be evaluated by the end of the semester. Evaluation will be focused on the theoretical knowledge and criteria needed to design, select, and evaluate different machine learning models in practical applications.

(final project assessment will represent 60% of the final grade).

Final assessment:

Those students that have renounced to continuous evaluation should address a final examination including both individual presentations to demonstrate theoretical knowledge on deep learning models (40% of final grade) and their final collaborative project (60% of the final grade).

Extraordinary examination:

Extraordinary examination consists of an individual presentations to demonstrate theoretical knowledge on deep learning models (40% of final grade) and a final collaborative project (60% of the final grade).

8. Teaching resources

8.1. Teaching resources for the subject

Name	Type	Notes
An Introduction to Statistical Learning	Bibliography	James, Gareth, Daniela Witten, Trevor Hastie, and Robert Tibshirani. An introduction to statistical learning. Vol. 112. New York: springer, 2013.
Machine learning: a probabilistic perspective	Bibliography	Kevin P. Machine learning: a probabilistic perspective. MIT press, 2012
The Elements of Statistical Learning Data Mining, Inference, and Prediction,	Bibliography	Hastie, Trevor, Tibshirani, Robert and Friedman, Jerome. The Elements of Statistical Learning Data Mining, Inference, and Prediction, Second Edition. Springer Series in Statistics, 2009
Deep learning	Bibliography	Goodfellow, I., Bengio, Y., Courville, A., &Bengio, Y. (2016). Deep learning. Cambridge: MIT press.
Neural Networks and Deep Learning	Web resource	http://neuralnetworksanddeeplearning.com/index.html
Scaling up machine learning: Parallel and distributed approaches.	Bibliography	Bekkerman, Ron, Mikhail Bilenko, and John Langford, eds. Scaling up machine learning: Parallel and distributed approaches. Cambridge University Press, 2011
Pattern recognition and machine learning (information science and statistics).	Bibliography	Christopher M. Bishop. Pattern Recognition and Machine Learning (Information Science and Statistics), 2006.