ANX-PR/CL/001-01
LEARNING GUIDE

SUBJECT

93000941 - Predictive And Descriptive Learning

DEGREE PROGRAMME

09AT - Master Universitario En Teoria De La Señal Y Comunicaciones

ACADEMIC YEAR & SEMESTER

2019/20 - Semester 1

Index

Learning guide

1. Description

1.1. Subject details

Name of the subject	93000941 - Predictive And Descriptive Learning
No of credits	6 ECTS
Type	Optional
Academic year ot the programme	First year
Semester of tuition	Semester 1
Tuition period	September-January
Tuition languages	English
Degree programme	09AT - Master Universitario En Teoria De La SeÑal Y Comunicaciones
Centre	09 - Escuela Tecnica Superior de Ingenieros de Telecomunicacion
Academic year	2019-20

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

The subject - recommended (passed), are not defined.

3.2. Other recommended learning outcomes

- It is mandatory to follow this course simultaneously with the subject Machine Learning Lab

- 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

CB06 - 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

CB07 - 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

CB09 - 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

CB10 - Que los estudiantes posean las habilidades de aprendizaje que les permitan continuar estudiando de un modo que habrá de ser en gran medida autodirigido o autónomo

CE01 - Analizar y aplicar técnicas para el diseño y desarrollo avanzado de equipos y sistemas, basándose en la teoría de la señal y las comunicaciones, en un entorno internacional

CE02 - Evaluar y sintetizar los resultados de un trabajo en equipo en proyectos relacionados con la teoría de la señal y las comunicaciones, en un entorno internacional.

CE03 - Valorar y contrastar la utilización de las diferentes técnicas disponibles para la resolución de problemas reales dentro del área de teoría de la señal y comunicaciones.

CT01 - Capacidad para comprender los contenidos de clases magistrales, conferencias y seminarios en lengua inglesa

CT03 - Capacidad para adoptar soluciones creativas que satisfagan adecuadamente las diferentes necesidades planteadas

CT04 - Capacidad para trabajar de forma efectiva como individuo, organizando y planificando su propio trabajo, de forma independiente o como miembro de un equipo

CT05 - 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

RA34 - Capability to develop and evaluate machine-learning techniques and to design big data learning systems

* 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 traditional Machine Learning models to Deep Learning. 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 relevant traditional predictive or supervised techniques: 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. Then the course addresses traditional descriptive or unsupervised techniques: principal components analysis and clustering methods (k-means and hierarchical clustering). From this basic background the course presents the recent and very powerful Deep Learning models: students learn from the basics of Neural Networks to the most common architectures of Feed-Forward Networks, Convolutional Networks and Recurrent Neural Networks.

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. Linear Regression
    2.1. Simple and Multiple Linear Regression
    2.2. Linear Regression and Distributed Machine Learning Principles
    2.3. Interpreting Regression Coefficients
    2.4. Model Selection and Qualitative Predictors
    2.5. Interactions and Nonlinearity
    2.6. Comparison of Linear Regression with KNN
3. Classification
    3.1. Logistic Regression
    3.2. Bayes classifier and Linear Discriminant Analysis
    3.3. Classification error analysis
    3.4. Quadratic Discriminant Analysis
    3.5. K-Nearest Neighbors
    3.6. A Comparison of Classification Methods: Logistic Regression, LDA, QDA and KNN
4. Resampling methods
    4.1. Cross-validation
    4.2. Bootstrap
5. Linear Model Selection and Regularization
    5.1. Feature selection
    5.2. Optimal Model selection
    5.3. Regularization
    5.4. Dimension Reduction
    5.5. High-Dimensional Data

6. Moving Beyond Linearity
   6.1. Generalized Linear Models and Generalized Additive Models
7. Tree-Based Methods
   7.1. Decision trees
   7.2. Bagging
   7.3. Random Forests
   7.4. Boosting
8. Support Vector Machines
   8.1. Maximal Margin Classifier
   8.2. Support Vector Classifiers
   8.3. Kernels and Support Vector Machines
   8.4. Relationship to Logistic Regression
9. Descriptive Learning
   9.1. Supervised vs Unsupervised learning
   9.2. Principal Components Analysis
   9.3. Clustering Methods
   9.4. K-means
   9.5. Hierarchical Clustering
   9.6. Practical Issues in Clustering
10. Introduction to Deep Learning
   10.1. Simple Neural Networks models
   10.2. Feed-forward Networks
   10.3. Convolutional Networks
   10.4. Recurrent Networks
   10.5. Introduction to advanced Deep Learning models: autoencoders, Generative Adversarial Networks (GANs)

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: 02:00 Lecture **Activities Chapter 2** Duration: 02:00 Lecture			
2	**Activities Chapter 3** Duration: 04:00 Lecture			
3	**Activities Chapter 4** Duration: 02:00 Lecture **Activities Chapter 5** Duration: 02:00 Lecture			
4	**Activities Chapter 6** Duration: 01:00 Lecture **Activities Chapter 7** Duration: 03:00 Lecture			
5	**Activities Chapter 8** Duration: 04:00 Lecture			
6	**Activities Chapter 9** Duration: 04:00 Lecture			
7	**Activities Chapter 10 (10.1 , 10.2)** Duration: 04:00 Lecture			
8	**Activities Chapter 10 (10.3)** Duration: 04:00 Lecture			
9	**Activities Use Case Review** Duration: 02:00 Problem-solving class			**Evaluation: Machine Learning use case** Individual presentation Continuous assessment Duration: 02:00
10	**Activities Chapter 10 (10.4)** Duration: 02:00 Lecture			**Evaluation: Machine Learning use case (continuation)** Individual presentation Continuous assessment Duration: 02:00

11	**Activities Chapter 10 (10.4)** Duration: 02:00 Lecture **Activities Chapter 10 (10.5)** Duration: 02:00 Lecture			
12	**Activities Chapter 10 (10.5)** Duration: 04:00 Lecture			
13	**Activities Deep Learning Review** Duration: 04:00 Problem-solving class			
14	**Activities Final Project discussions** Duration: 04:00 Problem-solving class			
15				
16				
17				**Final project evaluation** Group presentation Continuous assessment Duration: 00:15 **Evaluation: Machine Learning use case** Individual presentation Final examination Duration: 02:00 **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: Machine Learning use case	Individual presentation	Face-to-face	02:00	40%	/ 10	CB09 CT01 CT03 CT04 CT05
10	Evaluation: Machine Learning use case (continuation)	Individual presentation	Face-to-face	02:00	%	/ 10	
17	Final project evaluation	Group presentation	Face-to-face	00:15	60%	/ 10	CB09 CT01 CT03 CT04 CT05

7.1.2. Final examination

Week	Description	Modality	Type	Duration	Weight	Minimum grade	Evaluated skills
17	Evaluation: Machine Learning use case	Individual presentation	Face-to-face	02:00	40%	/ 10	CB09 CT01 CT03 CT04 CT05
17	Final project evaluation	Group presentation	Face-to-face	00:15	60%	/ 10	CB09 CT01 CT03 CT04 CT05

7.1.3. Referred (re-sit) examination

Description	Modality	Type	Duration	Weight	Minimum grade	Evaluated skills
Evaluation: Machine Learning use case	Individual presentation	Face-to-face	00:10	40%	/ 10	CB09 CT01 CT03 CT04 CT05
Final project evaluation	Group presentation	Face-to-face	00:15	60%	/ 10	CB09 CT01 CT03 CT04 CT05

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 machine learning models will be made by mid-semester

(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, and in particular deep learning architectures, 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 machine 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.