ANX-PR/CL/001-01

LEARNING GUIDE

SUBJECT

93001072 - Machine Learning Lab

DEGREE PROGRAMME

09AQ - Master Universitario En Ingenieria De Telecomunicacion

ACADEMIC YEAR & SEMESTER

2021/22 - Semester 1

Index

Learning guide

1. Description

1.1. Subject details

Name of the subject	93001072 - Machine Learning Lab
No of credits	4.5 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	2021-22

2. Faculty

2.1. Faculty members with subject teaching role

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

Mateo Jose Camara Largo	C-301	mateo.camara@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

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

- It is highly recommended to follow this course simultaneously with the subject Predictive and Descriptive Learning unless you have a theoretical background in Machine Learning and Deep Learning

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.

CG5 - 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.

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

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

In this laboratory students will learn how to apply the variety of Machine Learning methods presented in the Predictive and Descriptive Learning course to practical scenarios. Students will practice using scientifically-oriented processing environments and most common programing languages and machine learning libraries (TensorFlow, Keras, Python scikit-learn, ML libraries in R).

Along the course students will address practical problems on the application of the variety of Machine Learning methods presented in the Predictive and Descriptive Learning course. Experimental activities will cover both predictive or supervised learning (from classical linear and logistic regression or random forest and SVM to Deep Learning -Feed-forward, Convolutional Networks, Recurrent Networks) and descriptive or unsupervised prevised learning (principal component analysis, t-SNE and cluster analysis). Several realistic and practical scenarios and use cases will be addressed (as those proposed in Kaggle competition, www.kaggle.com). Students will practice using scientifically-oriented languages and cloud environments, mainly working with Python and R languages. Through all lab activities students will have to gain practice on model accuracy using cross-validation and on how to draw precise conclusions and valuable interpretations from machine learning results and models.

The students will acquire the skill to apply the variety of Machine Learning methods on to practic course outcome will be to consolidate the theoretical study of machine learning techniques along this Master Programme. Through hands-on experience case studies students will learn how to select and accurately assess the performance evaluation of machine learning methods. They will also acquire solid criteria on what could be best model for a given data and task as well to be able to draw precise conclusions and interpretations from experimental results. By the end of the course, students should be able to:

- Understand how to apply the most used models and techniques for predictive and descriptive learning to different real scenarios.

- Design a proper experimental methodology for accurately assessing and gaining knowledge from the use of each one of the different machine learning techniques.

- Work with both scientifically-oriented processing environments and cluster computing frameworks for big data processing that can be used in a wide range of applications in science and industry.

5.2. Syllabus

1. Introduction to Machine Learning Lab
    1.1. Designing a Machine Learning System
    1.2. Introducing Python for DataScience and Machine Learning
2. Linear Regression
    2.1. Developing interpretable Linear Regression models
3. Classification
    3.1. Developing and understanding Logistic Regression models
4. Resampling methods
    4.1. Using Cross-Validation and Bootstrap
5. Tree-Based Methods
    5.1. Decision tress, Bagging, Random Forests and Boosting
6. Support Vector Machines
    6.1. Kernels and Support Vector Machines

7. Descriptive Learning

   7.1. Principal Components Analysis, t-SNE, K-means and Hierarchical Clustering

8. Introduction to Deep Learning

   8.1. Simple Neural Network in TensorFlow (Basic Deep Learning Design Methodology)

   8.2. Feed-Forward Neural Networks (TensorFlow/Keras , PyTorch)

   8.3. Convolutional Networks for Images and Signals (TensorFlow/Keras , PyTorch)

   8.4. Recurrent Neural Networks: Signal and Natural Language Processing use cases (TensorFlow/Keras , PyTorch)

   8.5. Advanced Deep Learning architectures

9. Building Machine Learning Pipelines

6. Schedule

6.1. Subject schedule*

Week	Face-to-face classroom activities	Face-to-face laboratory activities	Distant / On-line	Assessment activities
1	**Introduction to Machine Learning Systems** Duration: 03:00 Lecture		**Introduction to Machine Learning Systems** Duration: 03:00 Lecture	
2		**Linear and Logistic Regression Models** Duration: 03:00 Laboratory assignments	**Linear and Logistic Regression Models** Duration: 03:00 Laboratory assignments	
3		**Linear and Logistic Regression Models** Duration: 03:00 Laboratory assignments	**Linear and Logistic Regression Models** Duration: 03:00 Laboratory assignments	
4		**Resampling methods** Duration: 03:00 Laboratory assignments	**Resampling methods** Duration: 03:00 Laboratory assignments	
5		**Tree-based models** Duration: 03:00 Laboratory assignments	**Tree-based models** Duration: 03:00 Laboratory assignments	
6		**Support Vector Machines** Duration: 03:00 Laboratory assignments	**Support Vector Machines** Duration: 03:00 Laboratory assignments	
7		**Descriptive Learning** Duration: 03:00 Laboratory assignments	**Descriptive Learning** Duration: 03:00 Laboratory assignments	
8		**Python for DataScience and Machine Learning** Duration: 03:00 Laboratory assignments	**Python for DataScience and Machine Learning** Duration: 03:00 Laboratory assignments	
9		**Review: Developing Machine Learning models** Duration: 03:00 Laboratory assignments	**Review: Developing Machine Learning models** Duration: 03:00 Laboratory assignments	**Evaluation: Developing Machine Learning models** Individual presentation Continuous assessment Presential Duration: 00:10
10		**Feed-forward Networks in TensorFlow and Keras** Duration: 03:00 Laboratory assignments	**Feed-forward Networks in TensorFlow and Keras** Duration: 03:00 Laboratory assignments	
11		**Convolutional Networks for Images and Signals** Duration: 03:00 Laboratory assignments	**Convolutional Networks for Images and Signals** Duration: 03:00 Laboratory assignments	
12		**RNN for Signals and NLP** Duration: 03:00 Laboratory assignments	**RNN for Signals and NLP** Duration: 03:00 Laboratory assignments	

13		**Advanced Deep Learning** Duration: 03:00 Laboratory assignments	**Advanced Deep Learning** Duration: 03:00 Laboratory assignments	
14		**Machine Learning Pipelines** Duration: 03:00 Laboratory assignments	**Machine Learning Pipelines** Duration: 03:00 Laboratory assignments	
15				
16				
17				**Final project evaluation** Group presentation Continuous assessment Presential Duration: 00:15 **Evaluation: Developing Machine Learning models** Individual presentation Final examination Presential Duration: 00:10 **Final project evaluation** Group presentation Final examination Presential Duration: 00:15

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 schedule is based on an a priori planning of the subject; it might be modified during the academic year, especially considering the COVID19 evolution.

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: Developing Machine Learning models	Individual presentation	Face-to-face	00:10	40%	/ 10	CG4 CT3 CT4 CT1 CT5 CG2 CG5 CG1
17	Final project evaluation	Group presentation	Face-to-face	00:15	60%	/ 10	CG4 CT3 CT4 CT1 CT5 CG2 CG5 CG1

7.1.2. Final examination

Week	Description	Modality	Type	Duration	Weight	Minimum grade	Evaluated skills
17	Evaluation: Developing Machine Learning models	Individual presentation	Face-to-face	00:10	40%	/ 10	CG4 CT3 CT4 CT1 CT5 CG2 CG5 CG1
17	Final project evaluation	Group presentation	Face-to-face	00:15	60%	/ 10	CG4 CT4 CT1 CT5 CG2 CG5 CG1

7.1.3. Referred (re-sit) examination

Description	Modality	Type	Duration	Weight	Minimum grade	Evaluated skills
Evaluation: Developing Machine Learning models	Individual presentation	Face-to-face	00:10	40%	/ 10	CG4 CT3 CT4 CT1 CT5 CG2 CG5 CG1
Final project evaluation	Group presentation	Face-to-face	00:15	60%	/ 10	CG4 CT3 CT4 CT1 CT5 CG2 CG5 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 developing 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 cover both individual achievements in Machine Learning, and more specifically Deep Learning models, as well as the development of teamwork skills, as this is one of the learning objectives for the course (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 skills in developing machine learning models (40% of final grade) and their final collaborative project covering machine learning and deep learning models (60% of the final grade).

Extraordinary examination:

Extraordinary examination consists of an individual presentations to demonstrate skills in developing machine learning models (40% of final grade) and a final collaborative project focused on machine learning and deep learning models (60% of the final grade).

8. Teaching resources

8.1. Teaching resources for the subject

Name	Type	Notes
Hands-On Machine Learning with Scikit-Learn, Keras, and TensorFlow: Concepts, Tools, and Techniques to Build Intelligent Systems	Bibliography	GGéron, Aurélien. Hands-on machine learning with Scikit-Learn, Keras, and TensorFlow: Concepts, tools, and techniques to build intelligent systems. O'Reilly Media, 2019.

Introduction to Statistical Learning	Bibliography	James, Gareth, et al. An introduction to statistical learning. Vol. 112. New York: springer, 2013
Python for data analysis	Bibliography	McKinney, Wes. Python for data analysis: Data wrangling with Pandas, NumPy, and IPython. " O'Reilly Media, Inc.", 2012.
Theory and R examples	Bibliography	James, Gareth, Daniela Witten, Trevor Hastie, and Robert Tibshirani. An introduction to statistical learning. Vol. 112. New York: springer, 2013.
Keras: the Python deep learning API	Web resource	https://keras.io/ Keras is an open-source neural-network library written in Python
PyTorch Tutorials	Web resource	https://pytorch.org/tutorials/beginne/nlp/pytorch_tutorial.html
Deep learning with Python.	Bibliography	F Chollet. Manning Publications Co., 2017
Deep Learning with Keras	Bibliography	Gulli, Antonio, and Sujit Pal. Deep Learning with Keras. Packt Publishing Ltd, 2017.
MSTC GitHub	Web resource	https://github.com/MasterMSTC
Andrej Karpathy blog About Hacker's guide to Neural Networks	Web resource	https://karpathy.github.io/

9. Other information

9.1. Other information about the subject

For on-line learning activities we will use UPM Moodle platform and tools. Moodle, GutHub and Youtube will be the environments to share specific course materials. Specific communication frameworks such as Skype for Business or Microsoft Teams could be used allowing UPM students to interact with instructors.

The increasing relevance of technological developments based on Machine Learning makes this course an educational activity directed to contribute to Goal 4.4 in Sustainable Development Goals (SDGs) 2030 United Nations Agenda, empowering our students with relevant skills, including technical and vocational skills, for employment, decent jobs and entrepreneurship.

Through approaching practical scenarios in our Lab, students will develop relevant skills and in-depth knowledge on the impact of different Machine Learning techniques on different fields as health, environmental monitoring, smart energy management, or finance. This will help them to become more aware of how technology can contribute to several SDGs goals: end poverty (Goal 1), promote well-being (Goal 2), and promote sustainable management of water, energy, economic growth and industrialization (Goals 5, 6, 7, and 8) as well as to reduce inequality among countries (Goal 10).

Also, due to the relevance of using machine learning to extract value from data in a broad range of economic sectors, the course will also contribute to SDG Goal 17 (Strengthen the means of implementation and revitalize the Global Partnership for Sustainable Development) in particular working on systemic issues on Data monitoring and accountability (17.18 and 17.19)