ANX-PR/CL/001-01

LEARNING GUIDE

SUBJECT

93000943 - Machine Learning Lab

DEGREE PROGRAMME

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

ACADEMIC YEAR & SEMESTER

2023/24 - Semester 1

Index

Learning guide

1. Description

1.1. Subject details

Name of the subject	93000943 - Machine Learning Lab
No of credits	4.5 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	2023-24

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
Juan Ignacio Godino Llorente	C-312	ignacio.godino@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

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

CB08 - Que los estudiantes sean capaces de integrar conocimientos y enfrentarse a la complejidad de formular juicios a partir de una información que, siendo incompleta o limitada, incluya reflexiones sobre las responsabilidades sociales y éticas vinculadas a la aplicación de sus conocimientos y juicios

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

CT06 - Capacidad para emitir juicios sobre implicaciones económicas, administrativas, sociales, éticas y medioambientales ligadas a la aplicación de sus conocimientos

4.2. Learning outcomes

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

RA7 - Capacidad para desarrollar y evaluar técnicas de aprendizaje automático y diseñar sistemas de aprendizaje para datos masivos

* 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, Pytorch, 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, Transformers) 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 practical scenarios. Main 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: Attention Mechanisms, Transformers, Deep Generative Models

9. Reviewing the guidelines to design and develop a Deep Learning project

9.1. Interpretability and Explainability of Machine Learning and Deep Learning models

9.2. Ablation studies

6. Schedule

6.1. Subject schedule*

Week	Classroom activities	Laboratory activities	Distant / On-line	Assessment activities
1	**Introduction to Machine Learning Systems** Duration: 03:00 Lecture			
2		**Linear and Logistic Regression Models** Duration: 03:00 Laboratory assignments		
3		**Linear and Logistic Regression Models** Duration: 03:00 Laboratory assignments		
4		**Resampling methods** Duration: 03:00 Laboratory assignments		
5		**Tree-based models** Duration: 03:00 Laboratory assignments		
6		**Support Vector Machines** Duration: 03:00 Laboratory assignments		
7		**Descriptive Learning** Duration: 03:00 Laboratory assignments		
8		**Python for DataScience and Machine Learning** Duration: 03:00 Laboratory assignments		
9		**Review: Developing Machine Learning models** Duration: 03:00 Laboratory assignments		**Evaluation: Developing Machine Learning models** Individual work Continuous assessment Not Presential Duration: 00:00
10		**Feed-forward Networks in TensorFlow and Keras** Duration: 03:00 Laboratory assignments		
11		**Convolutional Networks for Images and Signals** Duration: 03:00 Laboratory assignments		
12		**RNN for Signals and NLP** Duration: 03:00 Laboratory assignments		

13		**Advanced Deep Learning** Duration: 03:00 Laboratory assignments		
14		**Guidelines to design and develop a Deep Learning project** Duration: 03:00 Cooperative activities		
15				
16				
17				**Evaluation: Developing Deep Learning models** Individual work Continuous assessment and final examination Not Presential Duration: 00:00 **Evaluation: Developing Machine Learning models** Individual work Final examination Not Presential Duration: 00:00

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. Assessment

Week	Description	Modality	Type	Duration	Weight	Minimum grade	Evaluated skills
9	Evaluation: Developing Machine Learning models	Individual work	No Presential	00:00	50%	3.5 / 10	CT05 CB08 CB09 CE02 CT01 CB06 CT03 CB10 CB07
17	Evaluation: Developing Deep Learning models	Individual work	No Presential	00:00	50%	3.5 / 10	CT03 CB10 CT04 CT05 CT06 CB08 CB09 CE02 CT01 CE01 CE03 CB06 CB07

7.1.2. Global examination

Week	Description	Modality	Type	Duration	Weight	Minimum grade	Evaluated skills
17	Evaluation: Developing Deep Learning models	Individual work	No Presential	00:00	50%	3.5 / 10	CT03 CB10 CT04 CT05 CT06 CB08 CB09 CE02 CT01 CE01 CE03 CB06 CB07

17	Evaluation: Developing Machine Learning models	Individual work	No Presential	00:00	50%	3.5 / 10	CT03 CB10 CT05 CB08 CB09 CE02 CT01 CB06 CB07

7.1.3. Referred (re-sit) examination

Description	Modality	Type	Duration	Weight	Minimum grade	Evaluated skills
Evaluation: Developing Machine Learning models	Individual work	Face-to-face	00:00	50%	3.5 / 10	CT05 CB08 CT03 CB10 CB09 CE02 CT01 CB06 CB07
Evaluation: Developing Deep Learning models	Individual work	Face-to-face	00:00	50%	3.5 / 10	CB10 CT04 CT05 CT06 CB08 CB09 CE02 CT01 CE01 CE03 CB06 CB07

7.2. Assessment criteria

Evaluation will assess if students have acquired all the competences of the subject. Thus, evaluation through extraordinary assessment will be carried out considering all the evaluation techniques used in ordinary evaluation (EX, ET, TG, etc.).

Progressive evaluation will be the preferred assessment method as it will be suited to the optimum learning process along the course. It will consist of:

- A Machine Learning Report describing the activities that demonstrate skills in developing Machine Learning models. The evaluation of this Report will represent 50% of final grade. For progressive evaluation, this report must be due by the 9th week. Several course assignments, which will be announced in Moodle, will be planned to review the students' progress through draft versions of their reports so we can give them feedback. We cloud also require students to prepare specific presentations to review their work.
- A Deep Learning Report must be prepared by the end of the course to demonstrate skills in developing Deep Learning models. The evaluation of this Deep Learning Report will represent 50% of final grade. Through several course assignments, announced in Moodle, we will review the students' progress while working on this Report. We cloud require students to attend to specific presentations to review their work.

Deep Learning activities must be developed in working teams, but each team member must individually clearly describe her/his specific activities in the Report.

Challenge-Based Learning

Optionally, students in the course could freely choose to follow a learning process based on undertaking a Challenge in Quantitative Finance.

The Challenge statement will be published at the beginning of the course, including a calendar that will be in accordance with the rest of the course.

Students may request the Challenge as a group or individually. In the latter case, the teachers will be in charge of forming the groups. In both scenarios the groups will be composed by two students, and the students must meet the requirements to be able to develop the group work.

The Challenge topics for the students' groups will be Quantitative Finance on different areas:

1. Quantification of patterns in leading market indicators.
2. Quantification of patterns in market value.

3. Quantification of seasonality in the main market indicators
4. Quantification of intra- and inter-day patterns, volatility, etc.

The development of the challenge will be divided into four phases:

1. Research: study of the challenge statement and research on possible solutions. The students will have to inform themselves and formulate questions that will allow them to understand the dimension of the challenge and to approach a possible solution.
2. Development of the challenge: students will develop in teams small activities leading to identify the most appropriate solution to the problem, all of them proposed by the teacher in view of the previous stages.
3. Verification and validation: the results obtained and the chosen solution will be contrasted in real environments.
4. Elaboration of the report and/or exhibition: the results will be shared through a working report and/or an exhibition, which may be done through a video.

The monitoring of the phases of the activity will be developed in tutorial sessions with the teachers designated for this purpose. The evaluation will be carried out in a coordinated way between the teachers and the participants in the teams. The teachers will carry out a continuous evaluation of the performance and the achievement of the objectives set during the development of the challenge for each student. Likewise, after completing the challenge, students will perform a self-evaluation and a cross evaluation. The weight of the exercise in the grade will be the same as that assigned to the group work.

Global or final evaluation will consist of:

- A Machine Learning Report describing the activities that demonstrate skills in developing Machine Learning models. The evaluation of this Report will represent 50% of final grade and it must be due by the final exam date, although students can submit draft versions before that date can they can receive feedback on their work.
- A Deep Learning Report must be prepared to demonstrate skills in developing Deep Learning models. The evaluation of this Deep Learning Report will represent 50% of final grade and it must be due by the by the final exam date.

Deep Learning activities must be developed in working teams, but each team member must individually clearly describe her/his specific activities in the Report.

For both Reports, students can submit draft versions before the final submission date so they can receive feedback

on their work. Students will be required to attend to specific final presentations to defend their work on both Machine Learning and Deep Learning Reports.

Evaluation through extraordinary assessment will require the same process as the one described before for Global or final evaluation.

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	Géron, Aurélien. Hands-on machine learning with Scikit-Learn, Keras, and TensorFlow: Concepts, tools, and techniques to build intelligent systems. O'Reilly Media, 2nd Edition
Introduction to Statistical Learning	Bibliography	James, Gareth, et al. An introduction to statistical learning. Second Edition (2021) https://hastie.su.domains/ISLR2/ISLRv2_website.pdf
Python for data analysis	Bibliography	McKinney, Wes. Python for data analysis: Data wrangling with Pandas, NumPy, and IPython. " O'Reilly Media, Inc.", 2012.
scikit-learn Machine Learning in Python	Web resource	https://scikit-learn.org/stable/

CRAN Task View: Machine Learning & Statistical Learning	Web resource	https://cran.r-project.org/web/views/MachineLearning.html
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/
Deep learning with Python.	Bibliography	F Chollet. Manning Publications Co., 2017
Andrej Karpathy blog About Hacker's guide to Neural Networks	Web resource	https://karpathy.github.io/
MLLB at Moodle	Web resource	https://moodle.upm.es/titulaciones/oficiales/course/view.php?id=892

9. Other information

9.1. Other information about the subject

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)