

Faculty: Science and Technology, and Engineering

Course: Artificial Intelligence and Data Science

Programme: Study Abroad in Engineering

Semester: 1 – Fall

ECTS credits: 3

Duration: 22,5 hours

Language of instruction: English

Instructor: Miquel Camprodon Masnou

Course Description

This course covers the fundamentals of Data Science with emphasis on its application on real world problems. The main part of the course consists of covering all phases of the Data Science workflow including theory and practical programming. The last part of the course is dedicated to Artificial Intelligence with the study of classic learning algorithms and fundamentals of neural networks.

Prerequisites

First course of engineering studies or similar: fundamentals of math, statistics, and programming.

Attendance and punctuality policies

Attendance is mandatory for all classes. Any presentation or activity missed due to student absences can only be rescheduled in cases of certified medical or family emergencies. If a student misses more than three classes in any course half a letter grade will be deducted from the final grade for each additional absence. Seven absences in 6 ECTS courses or four absences in 3 ECTS courses will result in a Fail grade. Notice that there is a minimum of 80% attendance.

Students will be marked ABSENT from any class if they arrive more than 20 minutes late. Students will not be permitted to enter the class unless the professor specifically accepts it. Even if the instructor allows students to join the class, they will still be marked as absent for that lesson.

Absences can be justified in the following cases: Death of a first-degree relative, serious illness of the student/ first-degree relative or obligation to attend legal affairs. In all these cases a document or receipt must be sent via e-mail to studyabroad@uvic.cat adding your professor in copy. Important! In case of injury/ illness of the student, a medical document issued in Vic* needs to be provided.

*Medical documents accepted: physical doctor's notes which contain the hospital's stamp and signature in handwriting OR digital doctor's notes which contain the doctor's valid digital signature (a digital signature is valid when it shows the authentication of the person who signs and prevents the pdf to be modified after being signed).

Learning outcomes

By the end of the course, students should be able to:

- Know the Data Science workflow
- Know the main methodologies applied to each step of the Data Science workflow
- Work in a real Data Science environment
- Be fluent programming in Python
- Understand the typologies of Machine Learning problems: supervised learning, with classification and regression, and unsupervised learning
- Understand basic Artificial Intelligence techniques used on Game Algorithms, Evolutionary Algorithms, Constraint Programming
- Know the fundamentals of Neural Networks

Method of presentation

The course is structured into 13 lessons of 1,5 hours each divided into two parts, theory and programming.

Theory: magistral lessons given by the professor complemented with exercises and problems for students.

Programming: coding and programming in a Data Science environment (Databricks) with a Programming Language (Python) with a problem-solving approach. The course will have an automatic grading system.

Case analysis

Programming lessons will be based on specific datasets focusing on real-world problems.

Required work and assessment methods

Required work: every week students will have at least one programming assignment to be delivered. Weekly study of theory will be necessary.

Assessment methods: there will be a final exam and assignments during the course.

Final grade will be calculated as follows.

- 40% Programming Assignments
 - 13 Programming Assignments. One of them will account for double.
- 60% Final Exam
 - 40% Theory Final Exam
 - 20% Programming Final Exam

A minimum grade of 4 in the Programming Assignments is necessary to pass the course.

There will be a Retake Exam to recover the Final Exam.

Contents

Part 1: Data Science

Introduction to AI and DS

Data Science Workflow

Introduction to SQL

Advanced SQL

Data Exploration and Visualization

Data Cleaning and Preprocessing

Introduction to Machine Learning

Supervised Learning: Regression and Classification

Model Evaluation and Validation

Unsupervised Learning: Clustering and Dimensionality Reduction

Magistral Class by Pau Agulló: Analytics in Retail

Part 2: Artificial Intelligence

Game Algorithms and Learning Algorithms

Evolutionary Algorithms and Constraint Programming

Introduction to Neural Networks

Recommended reading

BASIC

Python for Data Analysis – Wes McKinney – O’Reilly

An Introduction to Statistical Learning – Gareth James - Springer

The Elements of Statistical Learning – Friedman - Springer

COMPLEMENTARY

Mining of Massive Datasets - Leskovec

Sites:

Databricks Community Edition

<https://community.cloud.databricks.com>