ANNEX 2.9

DEGREE PROGRAM DIDACTIC REGULATIONS

AUTONOMOUS VEHICLE ENGINEERING (MOVE)

LM-33

School: Scuola Politecnica e delle Scienze di Base

Department: Industrial Engineering

Didactic Regulations in force since the academic year 2024-2025

Course: MACHINE LEARNING AND BIG DATA	Teaching Language: English
SSD (Subject Areas): ING-INF/05	**CREDITS:** 9
Course year: I	**Type of Educational Activity:** B

Teaching Methods:

In-person

Contents extracted from the SSD declaratory consistent with the training objectives of the course:

Data Mining and Machine Learning. Knowledge representation: Trees, Rules, Clusters. Basic Machine Learning methods: Statistical Modelling, Linear Models, Instance-based learning, Clustering. Performance Evaluation: Cross-Validation, Cost-sensitive classification, ROC curves. Advanced Machine Learning: Decision Trees, Support Vector Machines, MLP. Data transformation: attribute selection, PCA. Deep Learning: training and performance evaluation of Deep Networks, Convolutional Neural Networks. Introduction to database systems. Data model for Big Data. NoSQL database: Key-value-Column-family, Graph database systems. Introduction to Big Data systems (BD): definition of a BD system. The Hadoop ecosystem. Yarn. Pig. Hive. Giraph. Spark. Introduction to Big Data Analytics (BDA): BDA Lifecycle: knowledge discovery in the database, data preparation, model planning, model building, data visualization. Examples of commercial and open-source Tools: Oracle, IBM Business Analytics, Microsoft Power BI, Microsoft Azure. AWS. SAP Hana.

Objectives:

The student must know the main Machine Learning (ML) algorithms and must demonstrate the ability to choose the most suitable ML algorithm to solve a specific problem, based on the requirements of the problem itself. The student must also know the techniques to be used for properly evaluating the performance of ML algorithms. The students must know the main Big Data frameworks in order to acquire, model, share, analyze and visualize large amount of information. The student must also demonstrate that he/she is able to choose the most suitable framework to deal with different tasks. The student must demonstrate to be able to solve real problems by using Machine Learning techniques. The student must also demonstrate that he/she can properly evaluate the performance of a machine-learning based system. The student must demonstrate to be able to manage, model and analyze large amount of data through different Big Data frameworks for dealing with different tasks, also evaluating the performance of the designed architecture.

Propaedeuticities:

None

Is a propaedeuticity for:

None

Types of examinations and other tests:
Oral examination and project discussion.