

Corso di Laurea Magistrale in Fisica		Insegnamento / Course METODI DI APPRENDIMENTO AUTOMATICO PER LA FISICA/ MACHINE LEARNING METHODS FOR PHYSICS		
SSD: INF/01	CFU/Credits: 8	Anno di corso: I/II	Lezione (ore): 40	Esercitazione (ore): 24
<p>Obiettivi formativi: Il corso consentirà ai discenti di conseguire conoscenze e capacità di comprensione nell'ambito dell'intelligenza artificiale con particolare attenzione agli algoritmi di apprendimento automatico. In maniera specifica, i discenti matureranno conoscenze relative alle principali aree dell'apprendimento automatico, come apprendimento supervisionato, non supervisionato e con rinforzo. Inoltre, ai discenti saranno fornite conoscenze di preelaborazione dei dati, come selezione e estrazione delle caratteristiche dei dati, riduzione della dimensionalità dei dati, gestione di dati mancanti, e gestione di dati categorici. I discenti, inoltre, conseguiranno conoscenze legate a tecniche di validazione dei modelli di apprendimento automatico come, ad esempio, il metodo di k-fold cross validation. In tale contesto, i discenti matureranno conoscenze e capacità di comprensione applicate relative all'utilizzo del linguaggio di programmazione Python e alla programmazione distribuita basata su GPU per l'implementazione di algoritmi di apprendimento automatico. Infine, i discenti saranno capaci di valutare in maniera critica le prestazioni di algoritmi di apprendimento automatico, comunicare con elevata abilità informazioni relative a tale dominio di ricerca, e apprendere nuove metodologie in maniera autonoma.</p>		<p>Training objectives: The course will allow learners to gain knowledge and understanding in the field of artificial intelligence with particular attention to machine learning algorithms. Specifically, learners will acquire knowledge related to the main areas of machine learning, such as supervised, unsupervised and reinforcement learning. Furthermore, learners will be provided with knowledge of data preprocessing, such as selection and extraction of features, data dimensionality, handling of missing data, and management of categorical data. Moreover, learners will acquire knowledge related to validation techniques of machine learning models such as, for example, the k-fold cross validation method. In this context, learners will develop knowledge and applied skills related to the use of the Python programming language and GPU-based distributed programming for the implementation of machine learning algorithms. Finally, learners will be able to critically evaluate the performance of machine learning algorithms, communicate information related to this research domain with high skill, and learn new methodologies autonomously.</p>		

Programma sintetico (sillabo):

Il corso si propone come obiettivo principale la presentazione dei principali metodi di apprendimento automatico e la loro implementazione in linguaggio Python. In particolare, saranno presentate le seguenti tematiche:

- Introduzione al machine learning:
 - introduzione all'apprendimento automatico, apprendimento supervisionato, apprendimento non supervisionato, apprendimento per rinforzo;
- algoritmi di apprendimento supervisionato basati sul perceptrone
 - neuroni adattivi lineari
 - regressione logistica
 - macchine a vettori di supporto
 - Metodi Kernel per dati non linearmente separabili
- alberi decisionali e foreste casuali;
- algoritmo di classificazione k-nearest neighbors (knn)
- Data preprocessing
 - Gestione dei missing data
 - Gestione dei categorical data
 - Partizionamento dei dati in training set e test set
 - Standardizzazione e normalizzazione di feature
 - Feature selection
- Riduzione della dimensionalità dei dati
 - Algoritmo Principal component analysis (PCA)
 - Algoritmo Linear discriminant analysis (LDA)
 - Algoritmo Kernel Principal Component Analysis
- Criteri per la valutazione dei modelli di machine learning
 - Metodo K-fold Cross Validation
 - Curve di validazione e apprendimento
 - Approccio di forza bruta per il tuning degli iperparametri
 - Metriche di valutazione dei modelli
- Metodi Ensemble
 - Approccio basato su Majority Vote
 - Approccio basato su Bagging
 - Approccio basato su Boosting
- Predizione di variabili continue
 - Linear regression
- Apprendimento non supervisionato: Clustering analysis
 - Algoritmo k-means
 - Algoritmi di clustering gerarchico
- Reti neurali
 - Single layer neural network
 - Multilayer neural network
 - Deep convolutional neural network e recurrent neural network
- Introduzione al calcolo parallelo basato su GPU in Python

Contents:

The main objective of the course is the presentation of the main methods of machine learning and their implementation in Python language. In particular, the following topics will be presented:

- Introduction to machine learning:
 - Introduction to machine learning concepts, supervised learning, unsupervised learning, reinforcement learning;
- Perceptron-based supervised learning algorithms
 - linear adaptive neurons
 - logistic regression
 - support vector machines
 - Kernel methods for non-linearly separable data
- Decision trees and Random forests;
- K-nearest neighbors (knn) classification algorithm
- Data preprocessing
 - Handling of missing data

- Management of categorical data
- Partitioning of data into training sets and test sets
- Standardization and normalization of features
- Feature selection
- Dimensionality reduction of the data
 - Principal component analysis (PCA) algorithm
 - Linear discriminant analysis (LDA) algorithm
 - Kernel Principal Component Analysis algorithm
- Criteria for evaluating machine learning models
 - K-fold Cross Validation Method
 - Validation and learning curves
 - Brute force approach for hyperparameter tuning
 - Model evaluation metrics
- Ensemble methods
 - Majority Vote based approach
 - Bagging based approach
 - Boosting based approach
- Prediction of continuous variables
 - Linear regression
- Unsupervised learning: Clustering analysis
 - k-means algorithm
 - Hierarchical clustering algorithms
- Neural networks
 - Single layer neural network
 - Multilayer neural network
 - Deep convolutional neural network and recurrent neural network
- Introduction to GPU-based parallel computing in Python

Esami propedeutici / Propaedeutic exams: -

Prerequisiti:

Conoscenze di informatica: architetture degli elaboratori. Programmazione procedurale (preferibilmente in linguaggio Python), costrutti condizionali, costrutti ciclici, funzioni, array e matrici. algoritmi e strutture dati.

Utilizzo di sistemi operativi Unix-like (ad esempio GNU/Linux).

Conoscenze di Matematica: algebra lineare e calcolo differenziale.

Prerequisites:

Computer science knowledge: computer architecture, procedural programming (preferably in language Python), conditional and iterative statements, functions, arrays and matrices. algorithms and data structures.

Use of Unix-like operating systems (for example GNU \ Linux).

Knowledge of Mathematics: linear algebra and differential calculus.

Finalità e modalità di verifica dell'apprendimento

Esame orale e progetto scritto.

Il corso può essere erogato in lingua inglese in presenza di studenti stranieri (es. Erasmus) / The course can be given in English in presence of foreign students (e.g. Erasmus)