UNIVERSITÀ
DEGLI STUDI DI MILANO-BICOCCA

COURSE SYLLABUS

Machine Learning M

2021-1-F8204B006

Obiettivi

Lo studente apprenderà le tecniche di Machine Learning più efficaci, comprendendo i fondamenti teorici di ogni tecnica e acquisendo il *know-how* per poterle applicare con successo alla risoluzione di problemi pratici. Sarà inoltre fornita una panoramica sulle più innovative soluzioni per l'identificazione del miglior algoritmo di Machine Learning e della sua configurazione ottimale (Automated Machine Learning – AutoML), dato un dataset. Lo strumento di riferimento per il corso sarà R, ma verranno anche presentate alcune soluzioni equivalenti in Python (ad esempio scikit-learn) e Java (ad esempio WEKA, KNIME).

Contenuti sintetici

Concetti basi del Machine Learning: tipologie di dati, *istanze*, *features*, *tasks* e *scenarios*, *parametri* e *iper-parametri*, misure di performance

Tecniche di apprendimento non-supervisionato

Tecniche di apprendimento supervisionato: classificazione e regressione

Modellare non-linearità nei dati: tecniche basate sul concetto di *kernel*

Automated Machine Learning: configurazione automatica di un modello di Machine Learning

Programma esteso

Introduzione

- Machine Learning scenarios & tasks, notazioni utili
- [illegible]

Unsupervised Learning

- Concetti di similarità e distanza
- [illegible]
- [illegible]

Supervised Learning

- Generalità: classificazione e regressione, metriche, tecniche di validazione (hold-out, *k* fold-cross, leave-one-out)
- [illegible]
- [illegible]

Supervised Learning per dati non-lineari

- Non-linearità, VC dimensions, kernel-trick
- [illegible]
- [illegible]
- [illegible]

L’approccio connessionista

- Artificial Neural Networks: paradigma di apprendimento
- [illegible]

Una panoramica su Automated Machine Learning (AutoML)

Esercizi ed esempi pratici

Prerequisiti

Si consiglia la conoscenza di elementi di base di informatica, matematica applicata, probabilità e statistica

Modalità didattica

L'intera attività formativa viene svolta attraverso lezioni in presenza, video-registrate al fine di essere rese d isponibili in formato digitale. Le lezioni riguarderanno sia aspetti teorici che applicazioni pratiche, specificatamente l'utilizzo di librerie software e dati open.

Nel periodo di emergenza Covid-19 le lezioni si svolgeranno [illegible]

Materiale didattico

- Testo di riferimento: Mehryar Mohri, Afshin Rostamizadeh and Ameet Talwalkar (2018). Foundations of Machine Learning.
- Slides e materiale didattico fornito dal docente

Altri tesi suggeriti:

- _____
- Carl Edward Rasmussen and Christopher K. I., Williams (2006). Gaussian Processes for Machine Learning.
- Robert B. Gramacy (2020). Surrogates – Gaussian Processes Modeling, Design, and Optimization for the Applied Statistics.

Periodo di erogazione dell'insegnamento

Second semester

Modalità di verifica del profitto e valutazione

La modalità di verifica prevede le seguenti 2 prove:

- lo svolgimento di un progetto con associata redazione di un rapporto tecnico, stile articolo scientifico,
- un esame orale finalizzato a verificare il grado di comprensione degli argomenti trattati.

_____

Il progetto contribuisce al 60% della valutazione finale, la prova orale al restante 40%.

_____

Orario di ricevimento

Su appuntamento