

CORSO: Topics in machine learning

DOCENTI: Gianluca Rossi (Ph.D., 2002)
EMAIL: gianluca.rossi@uniroma2.eu
PAGINE WEB: <https://github.com/glucatv>

DESCRIZIONE DEL CORSO

Il corso affronterà problemi di Classificazione, Regressione e Clustering. Partendo dal problema si esploreranno diversi modelli e se ne confronteranno i risultati. I problemi analizzati per lo più verranno dal campo dell'analisi geo-spaziale, quindi si vedrà in tale contesto come usare i modelli già noti e verranno introdotti modelli più specifici per il particolare campo applicativo.

OBIETTIVI DI APPRENDIMENTO

- ✓ Abilità ad utilizzare e manipolare dati spaziali con gli strumenti opportuni
- ✓ Comprensione dei modelli utilizzati
- ✓ Abilità nell'utilizzare i modelli conosciuti nella soluzione di nuovi problemi.

METODOLOGIA

Lezioni teoriche ed esercitazioni, per le quali si farà l'uso del linguaggio Python e delle librerie pandas, sklearn, geopandas, shapely, matplotlib, rasterio, gdal.

VALUTAZIONE

Progetto individuale: 80%;
Orale 20%

PROGRAMMA

1. Introduzione all'analisi dei dati geo-spaziali
 1. Sistemi di coordinate
 2. Dati vettoriali e dati raster
 3. Acquisizione e manipolazione di dati spaziali
2. Introduzione a Python
 1. Tipi di dato, operatori, strutture dati primarie
 2. Le librerie matplotlib, numpy, pandas, sklearn, geopandas, rasterio, gdal
3. Problemi di Classificazione
 1. Classificazione di utenti in base ai loro spostamenti
 1. Logistic regression
 2. Alberi di decisione
4. Problemi di Regressione
 1. Stimare il prezzo dell'affitto di una casa
 1. Linear regression
 2. Decision tree regression
 2. Stima delle quote altimetriche e confronto con i metodi di interpolazione lineare e gaussiano
5. Problemi di Clustering
 1. Analisi di traiettorie per l'individuazione di punti di interesse
 1. Il modello DBSCAN

2. Il modello OPTICS

LIBRI DI TESTO

M. Bishop "*Pattern Recognition and Machine Learning*", Springer 2006