

Laboratorio di web scraping

Laurea triennale in Informatica

6 crediti – 2° semestre

Titolare del corso: Laura Ricci

Il focus del corso è l'analisi di dati ottenuti dal web e da sistemi distribuiti, sia mediante API che mediante web scraping. Come use case ci focalizzeremo sull'ecosistema delle criptomonete e, più in generale, delle blockchain.

Il programma preliminare è il seguente:

Strumenti per l'analisi dei dati

- richiami di statistica di base
- graph models: random graphs, scale free networks, small worlds
- proprietà e caratteristiche di un grafo (degree, centrality, clustering coefficient,..)
- community detection e cluster algorithms su grafi

Data Model

- struttura delle transazioni e blocchi di diversi protocolli di blockchain
 - Bitcoin, Ethereum, Steemit,...
- grafi di transazioni
- ecosistema: storico di transazioni di exchange, reti di comunicazione P2P, external data gathering (e.g. twitter), cross chain swaps

Laboratorio

- breve introduzione a Python
- uso di API per il reperimento di dati dal web
- strumenti di web scraping
- Pandas: Numpy e matplotlib
- librerie per l'analisi di grafi (NetworkX, ...)
- strumenti per la visualizzazione di grafi
- analisi di dati
 - transazioni
 - exchange services
 - smart contracts
 - NFT, Web-3

Modalità di esame: Progetto+ Orale

Materiale didattico: Slide+ Pandas Notebooks