QaA aprile Python

In questa lezione mensile che andremo ad avere risponderò a tutte le domande che mi sono state poste da voi studenti!

Potete pormele tramite il portale di MasterD oppure contattandomi direttamente



Python ha una vasta gamma di librerie, alcune delle quali sono molto popolari come NumPy, Pandas e Matplotlib, ma ci sono anche molte gemme meno conosciute che possono essere molto utili a seconda delle tue esigenze.

Ecco alcune librerie Python meno conosciute ma potenzialmente molto utili:



Arrow - Una libreria per la gestione di date e orari che offre un'interfaccia più intuitiva rispetto al modulo datetime standard di Python.

Boltons - Una raccolta di oltre 200 strumenti e utility Python "durevoli" che non si trovano nei moduli della libreria standard.

Dataset - Semplifica l'accesso ai database SQL, fornendo una API basata su Python che permette di evitare query SQL scritte a mano per le operazioni comuni.



Delegator.py - Consente di eseguire comandi del subprocesso in modo più semplice e leggibile, ideale per script di automazione.

Tenacity - Una libreria per la rielaborazione di operazioni fallite, ideale per scenari di programmazione di rete o cloud dove possono verificarsi errori temporanei.

Typer - Utilizzata per la creazione di applicazioni CLI, basata su Python 3.6+ type hints, Typer permette di costruire interfacce a linea di comando in modo veloce e con poco codice.



Rich - Permette di migliorare l'output nei terminali con formattazione avanzata, tabelle colorate, barre di progresso, ecc.

Pydantic - Una libreria per la validazione dei dati e la gestione delle impostazioni basata sui type hints di Python, rendendo il codice più robusto e mantenibile.

FastAPI - Anche se sta guadagnando rapidamente popolarità, vale la pena menzionarla per lo sviluppo di API con documentazione automatica e basata sui moderni standard Python.



Dash - Una libreria per la creazione di applicazioni web analitiche interattive in puro Python (senza necessità di JavaScript).

Django: Un framework per lo sviluppo web ad alto livello che incoraggia lo sviluppo rapido e pulito, con un'enfasi sulla riusabilità e sulla "pluggabilità" dei componenti. È uno dei framework web più popolari per Python.

Flask: Un micro-framework per applicazioni web. A differenza di Django, che richiede l'uso di specifici strumenti e librerie, Flask lascia molta più libertà allo sviluppatore, permettendo di utilizzare gli strumenti che meglio si adattano alle loro esigenze.



TensorFlow: Una libreria per l'apprendimento automatico sviluppata da Google, che permette la creazione di modelli di deep learning e machine learning. È uno dei toolkit più famosi nel campo dell'intelligenza artificiale.

PyTorch: Simile a TensorFlow, è una libreria per l'apprendimento automatico sviluppata da Facebook che offre una maggiore flessibilità e velocità nello sviluppo di modelli complessi.



Scikit-learn: Una libreria per il machine learning che fornisce strumenti semplici ed efficienti per l'analisi dei dati e il data mining. È la libreria di riferimento per chi inizia con il machine learning.

Jupyter Notebook: Una piattaforma open-source che permette di creare e condividere documenti che contengono codice live, equazioni, visualizzazioni e testo narrativo. È molto usata in data science e nell'insegnamento.

Matplotlib: Una libreria di plotting per Python che fornisce una varietà di formati di output e interfacce per interagire con vari ambienti interattivi. È molto usata per la creazione di grafici e figure in pubblicazioni scientifiche.



- Statsmodels: Permette agli utenti di esplorare dati, stimare modelli statistici, e eseguire test statistici. Un'ottima scelta per l'econometria e l'analisi statistica.
- Plotly: Un'altra libreria di visualizzazione che supporta grafici interattivi e dinamici che possono essere pubblicati online come parte di report o dashboard.
- Bokeh: Simile a Plotly, Bokeh è rivolta alla creazione di grafici interattivi che possono essere facilmente integrati in applicazioni web.

Analisi Statistica

- Descrittiva: Calcolo di statistiche riassuntive come media, mediana, moda, range, varianza, deviazione standard.
- Inferenziale: Test di ipotesi, intervallo di confidenza, analisi di correlazione per inferire le relazioni tra variabili.

Analisi Esplorativa dei Dati (EDA)

 Visualizzazione dei dati: Uso di grafici come istogrammi, box plot, scatter plot per esplorare distribuzioni e relazioni tra variabili.



 Pulizia dei dati: Rimozione di valori nulli o duplicati, correzione di errori di formato, normalizzazione dei dati.

Modellazione Predittiva

- Regressione: Lineare, logistica, polinomiale per prevedere valori continui o categorici.
- Classificazione: Alberi decisionali, foresta casuale, SVM, K-Nearest Neighbors per categorizzare i dati in classi.
- Clustering: K-means, DBSCAN, clustering gerarchico per identificare gruppi naturali di dati non etichettati.

Riduzione della Dimensionalità

- Analisi delle Componenti Principali (PCA): Per ridurre il numero di variabili mantenendo la maggior parte delle informazioni.
- t-SNE: Per la visualizzazione di dati ad alta dimensione in spazi a due o tre dimensioni.
- Feature Selection: Tecniche per selezionare le variabili più informative.

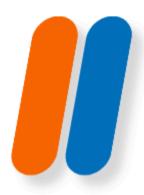


Analisi di Serie Temporali

- Modelli ARIMA: Per analizzare e prevedere dati che variano nel tempo.
- Smoothing e Decomposizione: Per identificare la tendenza e la stagionalità nei dati temporali.

Analisi di Testo (NLP)

- Tokenizzazione: Divisione del testo in parole o frasi.
- TF-IDF: Calcolo della frequenza dei termini per valutare l'importanza delle parole nel contesto.
- Sentiment Analysis: Determinare l'atteggiamento o l'emozione del testo.
- Word Embedding: Tecniche come Word2Vec o GloVe per convertire il testo in forme numeriche che catturano il contesto semantico.



Deep Learning

- Reti Neurali Convoluzionali (CNN): Per analisi di immagini e riconoscimento di pattern.
- Reti Neurali Ricorrenti (RNN): Per dati sequenziali come testo e serie temporali.
- Autoencoders: Per compressione dati e riduzione della dimensionalità.

Visualizzazione Avanzata

 Heatmaps, network graphs, e geographical maps per rappresentare complesse relazioni tra dati in modi intuitivi e interattivi.



