

# Domagoj Korais

*Data scientist, background in sviluppo software e fisica.*

email LinkedIn GitHub

## Posizione lavorativa corrente

Data scientist presso Sinergise e Planet Labs.

Sept - 2020 - now

Attualmente lavoro in full remote da Trieste, Planet Labs ha la sede a San Francisco e varie succursali in Europa. (Sinergise è stata acquisita da Planet Labs a settembre 2023, le mie mansioni sono rimaste le stesse).

### Attività:

Sviluppo di soluzioni basate sull'applicazione di algoritmi di machine learning alle osservazioni satellitari, con un focus particolare nell'ambito del monitoraggio delle attività agricole.

Analisi e classificazione di serie temporali.

Comunicazione dei risultati ai clienti e al pubblico.

Gestione di progetti, dall'onboarding del cliente alla consegna finale dei risultati, scrittura assieme al reparto commerciale delle specifiche tecniche, meeting con i clienti.

Organizzazione di journal club per mantenere il team aggiornato sullo stato dell'arte.

Stack tecnologico: Python e il relativo stack di data science (Pandas, Numpy, Pyarrow, pytorch, LightGBM), cloud computing: AWS per l'archiviazione dei dati (bucket S3) e il cloud computing (lavorando su instance EC2, anche multi GPU con training di modelli di deep learning in parallelo), Git come sistema di versionamento, docker per la virtualizzazione delle macchine e la riproducibilità dei risultati, metodologia Agile e Gitlab per organizzare il lavoro. Ray e Dask per l'esecuzione di pipeline. File parquet per le analisi batch con IO parallelo.

## Precedenti esperienze lavorative

Junior software developer presso Cybertec, Trieste.

Oct 2018-Sept-2020

### Attività:

Sviluppo principalmente lato server dell'applicazione Web APS (Advanced Planning and Scheduling) dell'azienda, utilizzando Typescript, Node.js, SQL, Backbone.js, e C++.

Implementazione di nuove funzionalità in base alle esigenze del cliente.

Utilizzo di Selenium Web Driver per test automatizzati basati su browser.

Durante la tesi di laurea magistrale abbiamo sviluppato un sistema basato su regole al fine di rendere più competitivo il prodotto aziendale. L'idea principale è quella di tradurre la conoscenza di dominio in regole, espresse utilizzando variabili linguistiche, e quindi mettere a punto la mappatura tra variabili linguistiche e valori quantitativi utilizzando l'apprendimento automatico.

## Educazione

**Laurea magistrale in Data Science and Scientific Computing** Sep 2017 - March 2020

Voto: 110/110.

Università: Università degli studi di Trieste, Trieste (Italia).

Tesi: "Un sistema di supporto alle decisioni basato su regole di apprendimento attivo per la pianificazione della produzione."

Conoscenze acquisite: - Statistica, statistica bayesiana, Machine learning, visualizzazione dei dati, computer vision, intelligenza artificiale, basi di dati relazionali.

Linguaggi di programmazione: C++, Python, R. - Introduzione ai concetti di HPC utilizzando MPI, OpenMp e CUDA su un cluster.

Linux: bash.

Versioning: git.

Concetti e metodi di sviluppo agile (sviluppo test driven, integrazione continua) in Java.

### **Erasmus+ traineeship in Slovenia**

Apr 2016 - Sep 2016

Tirocinio presso il Karst research institute situato a Postumia. La mia ricerca si è concentrata sullo studio della relazione tra le condizioni meteorologiche esterne e interne, utilizzando i dati meteorologici disponibili e i data logger in situ.

### **Laurea in fisica**

Sep 2011 - March 2016

Voto: 96/110.

Università: Università degli studi di Trieste, Trieste (Italia).

Tesi: "Rilievo gravimetrico della grotta "Grotta Impossibile".

Pubblicazione tesi (non peer reviewed): "Atti e Memorie della Commissione Grotte "E. Boegan""

### **Scientific high school degree**

Sep 2006 - Jul 2011

Liceo scientifico Paolo Ruffini, Viterbo (Italy).