



Roben Bhatti

✉ robenbhatti@gmail.com  linkedin.com/roben-bhatti  github.com/bhroben
☎ +39 331 211 6804 📍 Brescia, Italia 🇮🇹 Cittadinanza italiana

Esperienza

Agenzia Spaziale Tedesca (DLR)

Ott 2024 - Luglio 2025

Data Scientist Intern

Brema, Germania

- Sviluppato e implementato modelli bayesiani per quantificare l'incertezza nei coefficienti aerodinamici di veicoli spaziali riutilizzabili.
- Sviluppato una pipeline **CI/CD** con **Gitlab** e **Docker**, automatizzando il processo di release garantendo affidabilità.
- Collaborato in un team **Agile** ottenendo risultati basati sui dati, portando a termine la mia Tesi Magistrale e a un articolo **co-autore** per una conferenza (IAC 2025) (in preparazione).
- Ridotto dei bottleneck computazionali nei modelli grazie all'implementazione di Gaussian Processes Sparsi, con una **riduzione del 30%** dei tempi di addestramento e un'analisi dell'incertezza più scalabile.

Formazione

Università di Padova

Ott 2022 - Lug 2025

Laurea Magistrale in Fisica "Physics of Data" (Votazione 110/110)

Padova, Italia

- **Corsi rilevanti:** Metodi Matematici e Numerici, Deep Learning e Reti Neurali, Statistica Avanzata per l'Analisi Fisica, Teoria dell'Informazione e Inferenza, Relatività Generale.

Università di Padova

Ott 2019 - Ott 2022

Laurea Triennale in Astronomia (Votazione 95/110)

Padova, Italia

- **Corsi rilevanti:** Analisi Matematica Avanzata, Statistica, Meccanica Analitica, Fisica Quantistica, Relatività Ristretta.

Progetti

Elaborazione di Dati di Fisica delle Particelle in Streaming con Spark e Kafka

- Progettazione di una pipeline distribuita in grado di elaborare oltre **10 GB** di dati al giorno, simulando lo streaming in tempo reale di un rivelatore di particelle tramite **Kafka** e **Spark**.
- Deploy dell'ambiente su cluster cloud con **AWS S3** per lo storage, Kafka per la messaggistica, e Spark per l'elaborazione distribuita.
- Creazione di una dashboard interattiva in tempo reale con Bokeh per la **visualizzazione dei dati**.

Riconoscimento di Posizioni di Scacchi con Transformers (DETR) in PyTorch

- Sviluppo di una pipeline di computer vision per digitalizzare partite fisiche di scacchi, con una precisione dell'**87.5%** nel riconoscere pezzi e posizioni da un'unica immagine.
- Ottimizzazione del modello DETR (DEtection TRansformer) tramite **transfer learning**, riducendo significativamente i tempi di addestramento e migliorando le prestazioni su dataset personalizzati.
- Sviluppo di un modulo di **post-processing** per convertire automaticamente le rilevazioni grezze in notazione standard (FEN), rendendo subito compatibile lo stato della scacchiera con motori scacchistici e strumenti di analisi.

Pipeline di Machine Learning End-to-End per il Profiling di Clienti Assicurativi (In corso)

- Implementazione di un'architettura "Medallion" per garantire qualità e scalabilità su un dataset di **20K** polizze assicurative.
- Sviluppo di un modello in due fasi per prevedere gli importi dei sinistri: classificazione della probabilità di sinistro, seguita da una regressione con Random Forest per stimare l'ammontare nei casi ad alta probabilità.

Competenze Tecniche

Linguaggi: Python, R, SQL

Tecnologie: Docker, Git, CI/CD, Kafka, Spark, Pytorch, SciPy, NumPy, Pandas, Scikit-learn, Tableau.

Lingue

Inglese: C1, **Italiano:** Madrelingua