

# Manuele Macchia

Via Luigi Capriolo, 11 – 10139 Torino, Italia

manuelemacchia.com ◊ github.com/manuelemacchia ◊ macchiamanuele@gmail.com ◊ +39 334 2931307

Neolaureato in Data Science and Engineering con background in ingegneria informatica. Mindset analitico, attento ai dettagli e spirito creativo. Interessato ad applicare machine learning per risolvere problemi del mondo reale.

## Formazione

### Politecnico di Torino

Ottobre 2019 – Dicembre 2021 · Torino, Italia

Laurea Magistrale in Data Science and Engineering

Corsi: Machine learning and Deep learning, Statistical Methods in Data Science, Distributed architectures for big data processing and analytics, Data Ethics and Protection. Voto: 110/110 con lode.

### Politecnico di Bari

Settembre 2015 – Ottobre 2019 · Bari, Italia

Laurea Triennale in Ingegneria Informatica e dell'Automazione

Percorso Sistemi e Applicazioni Informatiche. Voto: 107/110.

### Universidad de Las Palmas de Gran Canaria

Gennaio – Giugno 2018 · Las Palmas, Spagna

Ingegneria Informatica, Programma Erasmus+

## Esperienza

### AIKO: Autonomous Space Missions

Marzo – Maggio 2021 · Torino, Italia

Deep Learning Intern

- Replicazione risultati Temporal Fusion Transformer per multi-horizon forecasting di serie storiche in TensorFlow.
- Sviluppo di una dashboard interattiva con Matplotlib in Jupyter per visualizzare e spiegare le predizioni dei modelli.
- Riduzione peso del modello del 90% mantenendo una quantile loss su test set comparabile.
- Adattamento modello per forecasting di dati satellitari proprietari, ottenendo 0.027 P50 risk sui dati di test.

## Progetti

### Explainable AI in deep active learning per image classification · Tesi laurea magistrale

Maggio – Dicembre 2021

- Progettazione ed implementazione di una libreria per deep active learning in Python e TensorFlow.
- Sviluppo di nuove query strategy basate su explainable AI per migliorare la selezione dei sample da annotare, ottenendo in media +0.7% di accuracy su un subset di ImageNet annotando il 50% dei dati disponibili.
- Ottimizzazione di explainable AI engine per CNN, riducendo il tempo di esecuzione del 90%.

### Predizione incendi boschivi con indici meteorologici · [Report](#) · [Codice](#)

Agosto – Settembre 2021

- Analisi e preprocessing di dataset di incendi boschivi per predire l'estensione degli incendi da misurazioni meteo.
- Sviluppo, validazione e analisi della qualità di modelli lineari e non lineari di classificazione e regressione in R.

### Incremental learning in image classification · [Paper](#) · [Codice](#)

Maggio – Luglio 2020

- Implementazione baseline di incremental learning su base di paper scientifici riproducendone i risultati in PyTorch.
- Progettazione ed implementazione di esperimenti per valutare l'efficacia di modifiche alle baseline, aumentando l'average incremental accuracy del 1.2%.
- Identificazione e sviluppo di nuovi approcci per risolvere problemi aperti nello stato dell'arte.

### Sentiment analysis di recensioni di hotel · [Articolo](#) · [Codice](#)

Dicembre 2019 – Febbraio 2020

- Preprocessing e pulizia di un dataset di recensioni di hotel in lingua italiana estratto da TripAdvisor.
- Sviluppo di una pipeline di machine learning con pandas e scikit-learn per predire il sentiment di nuove recensioni con accuratezza elevata (96.4% F1-score).

## Associazioni

### MALTO: Machine Learning at PoliTO

Novembre 2020 – Presente

- Founding member di team studentesco con l'obiettivo di partecipare a competizioni internazionali di data science, proponendo approcci all'avanguardia per risolvere problemi di machine learning e deep learning.
- Partecipazione ad ACM RecSys Challenge 2021. Sviluppo di recommender system per predire l'engagement di tweet. Analisi esplorativa e preprocessing di un dataset di 1 miliardo di sample fornito da Twitter. Sviluppo pipeline con feature store per allenamento distribuito su cluster di un modello XGBoost con PySpark e H2O. Costruzione di un engagement graph ed estrazione di feature con GraphFrames.