MATTEO MERLO

Data Scientist

matteo.merlo.995@gmail.com | +39 3497391525 | matteomerlo95 | Turin, Italy
MatteoM95 | matteom95.github.io/ | 0009-0002-8008-5093 | Italian

EXPERIENCE

AI Applied Researcher - Master's Degree Thesis

Links Foundation

Oct 2022 – Aug 2023 | Turin, Italy

- Developed an **extensive dataset** (500+ images) of past wildfires, using Sentinel-2 satellite imagery.
- Successfully proposed a **multitask learning semantic segmentation approach** for wildfire delineation and burn severity estimation. Tested and evaluated several state-of-the-art semantic segmentation models.
- Achieved robust results with **F1 score** over **92** for delineation and **RMSE** scores lower than **0.9** for severity estimates. Work published.

Data Engineer - Internship

Synapta s.r.l

Mar 2019 – May 2019 | Turin, Italy

- Development of **ETL pipelines**, including procedures to manipulate data in PostgreSQL environment.
- Development of an unstructured data web scraper in Python.

Software Engineer - Internship

Consoft Sistemi s.p.a.

Apr 2018 – Nov 2018 | Turin, Italy

- Implemented firmware on Arduino board in C++, testing various **sensors**.
- Tested **LoRaWAN communication protocol** as a solution in an IOT environment in Python.

EXTRACURRICULAR EXPERIENCE

Member Area IT

Icarus PoliTO

Oct 2016 – July 2020 | Turin, Italy

Icarus is a PoliTO **students team** focused on UAV airplane and rocket design. My primary contributions were:

- Designing and developing the ground station control and parachute system of the rocket on Arduino/STM32 Nucleo board in C++.
- Designed from scratch a **flight route planner** through clouds using algorithms such as Dijkstra and A* in C++, Java and C#. [Repository]

PUBLICATIONS

Conference Paper

- E. Arnaudo, L. Barco, M. Merlo, and C. Rossi, "Robust burned area delineation through multitask learning," 2023. arXiv: 2309.08368.

Dataset

- E. Arnaudo, L. Barco, M. Merlo, and C. Rossi, "Wildfires cems dataset," 2023.

EDUCATION

M.S. in Data Science and Engineering

Politecnico di Torino

Oct 2020 – July 2023

- Graduated with 92/110 (German: 1.9)
- Thesis: Multitask segmentation from satellite imagery for burned area delineation and severity estimation.

B.S. in Computer Engineering

Politecnico di Torino

Oct 2015 – July 2020

- Graduated with 95/110 (German: 1.9)

SKILLS

Code Languages:
Python, C++, Java, C, C#, JavaScript, R.

Machine Learning/Deep Learning:
Pytorch, Tensorflow, Keras, CUDA, Numpy, Pandas, Scikit-learn.

Databases and Big Data:
SQL, NoSQL, ETL, Pyspark, MapReduce.

Soft Skills:
Teamwork, Flexibility, Patience, Deep focus, Persistence, Consistency.

CERTIFICATES

- DeepLearning.AI - Deep Learning
- DeepLearning.AI - Generative AI with LLM
- IELTS (2017) - Overall band 7.0
- EF - Deutsch Kurszertifikat - A2

ACHIEVEMENTS

Best Paper Award
Conference ECML PKDD 2023 - MACLEAN

LANGUAGES

Italian
English
German

HOBBIES AND INTEREST

Chess | Hiking | Space Exploration | Reading scientific journal | Formula 1

CURRICULAR PROJECTS

Check out my Github for more cool projects:

Real-time Domain Adaptation in Semantic Segmentation

Project on computer vision focusing on image processing for real-time applications within the realm of autonomous driving solutions. By using a domain adaptation in combination with a style transfer techniques, it is possible to overcome the challenge of annotating large datasets for semantic segmentation.

➥ *Skill used*: Python, PyTorch, Torchvision, NumPy, TensorBoard, CUDA

[Repository] [Paper]

Default of Credit Card Clients Dataset Analysis

The project involved an in-depth data analysis utilizing advanced Machine Learning techniques, including SMOTE and PCA in preprocessing, followed by model training using Logistic Regression, SVM and Random Forest classifiers. Achieved a F1 score of 0.53 combining different preprocessing methods together.

➥ *Skill used*: Python, Scikit-Learn, Pandas, SMOTE, PCA, SVM, Random Forest, Logistic Regression

[Repository] [Notebook]

Smart Home Surveillance System

The indoor video surveillance system is designed to detect human intrusion through the integration of sound and visual recordings. This system operates entirely on Edge Computing taking advantages of TensorFlow Lite libraries, running on a Raspberry Pi 4.

➥ *Skill used*: Python, MQTT, TensorFlow, Speech Recognition, OpenCV

[Repository] [Paper]

Twitter Sentiment Analisys

In this project is proposed a study on a dataset of tweets using machine learning techniques to conduct sentiment analysis. The objective is to predict the sentiment associated with a tweet based on its text content. Achieved a F1 score of 0.85 using a Tf-idf Vectorizer.

➥ *Skill used*: Python, Scikit-Learn, NumPy, Pandas, Grid Search

[Repository] [Paper]

REFEREES

Prof. Garza, Paolo
Politecnico di Torino
paolo.garza@polito.it

Prof. Caputo, Barbara
Politecnico di Torino
barbara.caputo@polito.it