Manuele Macchia

Via Luigi Capriolo, 11 – 10139 Turin, Italy
manuelemacchia.com ◇ github.com/manuelemacchia ◇ macchiamanuele@gmail.com ◇ +39 334 2931307

Fresh graduate in Data Science and Engineering with a background in computer science. Detail-oriented, analytical mindset and creative spirit. Interested in computer vision, natural language processing, big data analytics.

Education

Politecnico di Torino October 2019 – December 2021 · Turin, Italy
Master of Science in Data Science and Engineering
Courses: Machine learning and Deep learning, Statistical Methods in Data Science, Distributed architectures for big data processing and analytics, Data Ethics and Protection. Final grade: 110/110 with honors (con lode).

Politecnico di Bari September 2015 – October 2019 · Bari, Italy
Bachelor of Science in Computer Science and Automation Engineering
Information systems and applications track. Final grade: 107/110.

Universidad de Las Palmas de Gran Canaria January – June 2018 · Las Palmas, Spain
Erasmus+ programme with scholarship, Degree in Computer Engineering

Experience

AIKO: Autonomous Space Missions March – May 2021 · Turin, Italy
Deep Learning Intern
○ Replicated state-of-the-art performance of Temporal Fusion Transformer for multi-horizon time series forecasting.
○ Set up an interactive dashboard via Matplotlib in Jupyter to visualize and explain model predictions.
○ Reduced baseline model weight by 90% while retaining comparable quantile loss on hold-out set.
○ Adapted model to forecast proprietary satellite sensor measurements, achieving 0.027 P50 risk on test data.

Projects

Explainable AI in deep active learning for image classification · Master's thesis May – December 2021
○ Designed and implemented a deep active learning library from scratch with Python and TensorFlow.
○ Developed and evaluated novel query strategies based on explainable AI to improve the selection of samples to annotate, obtaining +0.7% average accuracy on an ImageNet subset by labeling 50% of data.
○ Optimized state-of-the-art explainable AI engine for CNNs, reducing image processing time by 90%.

Predicting the area burned by forest fires with weather data · Report · Code August – September 2021
○ Analyzed and processed a forest fires dataset to predict burned area using weather measurements.
○ Developed and evaluated linear and non-linear predictive models for classification and regression in R.

Incremental learning in image classification · Paper · Code May – July 2020
○ Implemented incremental learning baselines from scratch in PyTorch based on scientific literature.
○ Designed and implemented experiments to study the effect of variations to state-of-the-art approaches, increasing the average incremental accuracy by 1.2% with respect to the baseline.
○ Identified and developed novel approaches to overcome limitations of baselines.

Sentiment analysis of hotel reviews · Article · Code December 2019 – February 2020
○ Processed, cleaned and analyzed a natural language dataset of hotel reviews in Italian scraped from TripAdvisor.
○ Developed a machine learning pipeline based on scikit-learn to predict the sentiment contained in unseen reviews with high accuracy (96.4% F1-score).

Associations

MALTO: Machine Learning at PoliTO November 2020 – Now
○ Founding member of student team with the goal to take part in international data science competitions, achieving high performance on machine learning problems and proposing novel cutting-edge approaches.
○ Participated in ACM RecSys Challenge 2021. Developed a recommender system to predict tweet engagement. Performed exploratory data analysis and preprocessing of a billion-sample dataset provided by Twitter. Developed a pipeline with feature store for distributed training of an XGBoost model on a cluster using PySpark and H2O. Built an engagement graph to extract new features using GraphFrames.

I authorize the processing of personal data contained in my resume based on art. 13 GDPR 679/16.