

ROBERTA PARISI

Data Scientist

@ roberta.parisi2@gmail.com ☎ +39 3474850757 📍 Rome, Italy
in linkedin.com/in/robertaparis-52a7798a

🔗 github.com/robertaparis

EDUCATION / COURSES

Master Degree in Data Science, 110/110 *With Honors*

Sapienza University

📅 Oct 2016 – Jan 2020 📍 Rome, Italy

- Main Exams: Big Data Analysis and Management, Statistical Learning, Data Mining, Data Monitoring, Bioinformatics, Digital Epidemiology, Social and Behavioural Network Analysis, Bayesian Statistics

Bachelor Degree in Management Statistics

Sapienza University

📅 Oct 2012 – Oct 2016 📍 Rome, Italy

- Main Exams: Descriptive Statistic, Probability, Inferential Statistic, Programming, Optimization and Statistical Models

EXPERIENCE

Data Scientist - ML Specialist

Live Tech s.r.l.

📅 August 2020 - ONGOING 📍 Rome, Italy

- Working in projects as consultant for several companies (Trelleborg, Enel, Sogei, Terna...) and European Projects (i.e.: FAIRE project) using advanced ML/DL methods in order to solve practical industry problems.
- Working to develop new feature for the proprietary framework Loko.ai (i.e. AutoML tool, which select the best model for you, Entity Extractor, Computer Vision component).

Tools: Python (TF, Optuna, Sklearn, PyTorch, Sanic, Sktime, Darts), Git, DS4Biz Studio, MongoDB, Docker

Machine Learning Specialist

Integris SPA - Bridgestone

📅 April 2020 - August 2020 📍 Rome, Italy

Data Scientist Research Intern

New York University

📅 Nov 2018 – February 2020 📍 New York City, USA

- Working on my Master Thesis Projects with a team of PhD candidates in the field of Information retrieval: global ordering of web page (static ranking) exploiting the usage of Learning To Rank algorithms.
- Experience with handling large dataset, using PySpark to clean it and extract from it new information (ClueWeb09-B collections) to use as features for the algorithm.

Data Scientist Intern

Vitrociset S.P.A.

📅 April 2018 – Oct 2018 📍 Rome, Italy

Management of Machine Learning applications and data science project: Worked on a Topic Extraction problem using LDA as topic model algorithm. Worked also as data analyst for travel BI.

Tools: Python (PySpark, Flask, pyLDAvis), Paperspace, Domo, Tableau

ACHIEVEMENTS

- Placed in 3rd position at *SUS³*, Statistics Hackaton held by Findomestic in Florence, Italy. Me and my team used an XGBoost model to predict and select the 10,000 client that would have a higher probability to sign a new loan.

SKILLS

Statistic/Machine Learning

●●●●●●

Linear and Logistic regression, LASSO, RIDGE, Elastic Net, Naive-Bayes Classifier, PCA, Factor Analysis, Clustering methods (KNN, K-means and K-median), Random Forest, Boosting, Neural Network (Perceptron, MLP, RBF, FFN, CNN, RankNet, TCN), SVM, NLP (LDA, NER models), Transfer Learning, Time Series modelling

Python

●●●●●●

Pandas, PySpark, Numpy, Scikit-Learn, Optuna, Sktime, Darts, Matplotlib, Seaborn, NetworkX, TensorFlow, Flask, PyTorch, Scipy, Sanic

R, MySQL, Neo4j

●●●●●●

Tableau, Domo

●●●●●●

Java, Matlab, MongoDB, Cassandra, Qlik, Mfinder, Rust, HTML

●●●●●●

EXAMPLE PROJECTS

Chemical Formula Prescription - Automotive Field

Creation of an optimization system that, starting from the desired chemical-physical properties, gives an accurate prescription of a formula, by choosing among all the trial done the one which allows to minimize the difference between the desired and the expected properties. The optimization is based on a combination of Random Search and Semi-Bayesian methods, performed using the framework Optuna. This system **helps the client** to lower consistently expenses, waste for disposal and needs of highly skilled personal in a time-consuming business process, since they don't need anymore to try on each formula variation.

LoKo AI

LoKo AI is the proprietary framework of LiveTech S.R.L., which is a Low Code platform whose aims is to democratize AI. My personal contribution to this framework is the development of several Machine Learning components, namely blocks, each of which carry out a specific task: OCR, Time Series, Computer Vision and AutoML tool.