

Felipe Operti

Turin, Italy

✉ operti.felipe@gmail.com • 🌐 opertifelipe.github.io • 🔗 opertifelipe
in felipeoperti



As a Full Stack Data Scientist, I have an extensive expertise in Natural Language Processing (NLP) and Deep Learning. My skills include designing, developing, and deploying advanced artificial intelligence solutions that drive business outcomes. I have worked extensively with Azure cloud infrastructure and AWS, utilizing technologies such as Docker, Kubernetes, and FastAPI. I hold a Ph.D. in Physics of Complex Systems, with research focused on computational analysis for socio-economic sciences. My skill set includes natural language processing, generative ai, deep learning, machine learning, software engineering, and data analysis. I have experience in various domains, including insurance, banking, technology, and retail. My background includes leading end-to-end projects with teams composed of data engineers, data scientists, and machine learning engineers. I am committed to delivering innovative solutions that drive business outcomes and am excited to collaborate with organizations that share my passion for artificial intelligence and technology.

EMPLOYMENT

Data Science Specialist
Reale Mutua

Turin, Italy
From Jul 2022

As a Data Scientist Specialist at Reale Mutua, my primary focus is on working with unstructured data using Natural Language Processing (NLP) and Deep Learning models. I work on the Azure cloud infrastructure, leveraging its capabilities to design, develop, and deploy advanced artificial intelligence solutions.

Technologies: Generative AI, Azure, AzureML, LangGraph, LangChain, Python, Docker, Pytorch, HuggingFace, Spacy, etc.

Full Stack Data Scientist
Self-employed

Turin, Italy
From Jan 2021

FraleAnalitica was born from a personal project with a friend, aiming to apply machine learning and deep learning techniques in the fields of agriculture, particularly beekeeping. In FraleAnalitica, I serve as a technology consultant and freelance Full Stack Data Scientist. My primary responsibilities include developing, deploying, and maintaining AI services for agricultural and beekeeping applications.

For more information:

- <https://opertifelipe.github.io/projects/honai/>
- <https://www.fraleanalitica.it/>

Technologies: Azure, MLOps, Machine Learning, Deep Learning, FastAPI, Docker, Python, etc.

Data Scientist | Machine Learning Engineer
Inda - Interviewweb S.r.l. - Zucchetti Group

Turin, Italy
Mar 2021 - Jul 2022

Data scientist at INDA (Interviewweb S.r.l. solution, a Zucchetti company) working principally with Natural Language Processing (NLP) and Deep Learning. Responsible for the development, deployment, and maintenance of machine learning solutions.

Technologies: AWS EKS, AWS ECS, AWS Lambda, AWS Sagemaker, Python, Docker, Kubernetes, FastAPI, ElasticSearch, MongoDB, etc.

Data Scientist Specialist Samsung SDS	São Paulo, São Paulo, Brazil Nov 2020 - Mar 2021
Data Scientist Specialist in CRM area at Samsung SDS. Responsible for the development and deployment of propensity models, churn models, sentimental analysis, and RFV models.	
Technologies: Data Lake, HDFS, Brightics, Adobe Campaign, Python, Scala, Docker, Spark, and PySpark.	
Data Scientist Specialist ViaVarejo	São Paulo, São Paulo, Brazil Jul 2020 - Oct 2020
Responsible for the development of machine learning models for price optimization and retail assortment.	
Technologies: Data Lake, Cloud, DataBricks, Azure, Python, Scala, Jenkins, Docker, Spark, and PySpark.	
Data Scientist Consultant Santander Bank	São Paulo, São Paulo, Brazil Aug 2019 - Jul 2020
Data Scientist in Santander DataLab. Data Lab is an innovative area whose aim is to develop large cross-areas end-to-end projects. Data Lab works as a consultancy society within the bank. In the lab, the interaction between product owners, data engineers, data scientists, and machine learning engineers is strongly encouraged. My responsibilities there are: from one side to develop machine learning models (as data scientist consultant). Due to the heterogeneity of the projects, also the applied machine learning and deep learning models are heterogeneous. On the other side, I am also responsible for checking the entire pipeline (from the ETL, the development of the models, and finally the deployment).	
Technologies: Data Lake, Hadoop, Cloudera, Python, Scala, SQL, Hive, Impala, Spark, PySpark, and Oracle.	
Senior Data Scientist Santander Bank	São Paulo, São Paulo, Brazil Dec 2018 - Aug 2019
Data scientist at Santander Bank in Customer Relationship Management (CRM). My position there was of senior data scientist and my responsibility was the development of machine learning models for different purposes (propensity and churn models for different products among others). Such models were typically machine learning supervised algorithms such as Random Forest, Gradient Boosting, Logistic Regression, and Artificial Neural Networks as well as unsupervised algorithms such as hierarchical, k-means, and DBSCAN. Furthermore, we worked with NLP algorithms for sentimental analysis of reclamations data.	
Technologies: Data Lake, Hadoop, Cloudera, python, R, Scala, SQL, Java, ShellScript, Hive, Impala, Spark, PySpark, and Oracle.	

EDUCATION

PhD in Physics, Complex Systems Federal University of Ceará	Fortaleza, Brazil Feb 2015 - Dec 2018
Thesis: Computational analysis for socio-economic sciences.	
Visiting PhD in Physics, Complex Systems City College of New York	New York City, United States Apr 2018 – Jun 2018
Visiting PhD in Physics, Complex Systems University La Sapienza	Rome, Italy Apr 2017 – Sep 2017
MSc in Physics, Complex Systems Federal University of Ceará	Fortaleza, Brazil Jul 2014 – Jan 2015
Thesis: Interpolation strategy based on Dynamic Time Warping.	
BSc in Physics University of Turin	Turin, Italy Sep 2009 – Oct 2012
Thesis: Monte Carlo Simulation of the radiation distribution emitted by a CT scan in the field of the radioprotection.	

SKILLS

Machine Learning and Deep Learning: LangGraph, LangChain, Pytorch, Unsloth, HuggingFace, OpenAI, Azure Cognitive Services, Spacy, etc.

MLOps: FastAPI, Kubernetes, Docker, Git, Typer, Pydantic, Poetry, Uv, Github Actions, Azure Pipelines, etc.

Cloud: Azure, AzureML, Azure DevOps, Azure Container App, AWS, AWS ECS, AWS EKS, AWS ElasticBeanstalk, AWS Sagemaker, etc.

Data Analysis: Python (Jupyter Notebook, Pandas, NumPy, StatsModels, etc), Scala, SQL, Spark, PySpark, Hadoop, Hive, Impala, Excel, etc.

Data Visualization: Matplotlib, Seaborn, Plotly, Streamlit, Grafana, Gephi, QGIS, etc.

Databases: MySQL, PostgreSQL, HDFS, ElasticSearch, and MongoDB.

Other computer skills: Java, C++, (learning) Rust, and Bash.

Operating systems: Linux (Arch Linux, Debian, and Ubuntu), MacOS, and Windows.

Languages: Italian (native), Portuguese (fluent), and English (fluent).

PUBLICATIONS

- Felipe G. Operti et al. Dynamics of racial segregation and gentrification in New York City. *Frontiers in Physics*. 2022.
- Felipe G. Operti et al. Dynamics in the Fitness-Income plane: Brazilian states vs World countries. *PlosOne*. 2018.
- Felipe G. Operti et al. The light pollution as a surrogate for urban population of the US cities. *Physica A*. 2018.