Alberto Ferrarin

Summary

Data Scientist with expertise in predictive modeling, machine learning, and data-driven optimization. Experienced in designing and deploying AI solutions across diverse domains, with a focus on accuracy, robustness, and interpretability. Skilled in collaborating with cross-functional teams to transform complex datasets into actionable insights and drive data-informed decision-making.

Experience

Istituto Nazionale dei Tumori, Data Scientist

- June 2024 – present
- Milano, IT
- Working on multiple data science and machine learning projects in the healthcare domain, including the EU-funded I3LUNG project on predictive modeling for immunotherapy response.
- Designed and implemented predictive models (traditional ML and neural networks) for large-scale clinical datasets (2,000+ patients), focusing on accuracy, robustness, and interpretability.
- Developed fusion models for multimodal data, integrating clinical, radiological, and pathological information, with support for missing modalities to enhance predictive performance.
- Streamlined data preprocessing and curation pipelines to ensure high-quality, consistent inputs for machine learning workflows.
- Presented results to cross-functional teams and international stakeholders at annual project meetings.

Politecnico di Milano, Fellowship Researcher

- June 2023 – May 2025
- Milano, IT
- Collaborating on the I3LUNG project, in continuity with the role at Istituto Nazionale dei Tumori.

ML cube, Data Engineer

- July 2022 – May 2023
- Milan, IT
- Designed and deployed a distributed electronic data capture platform supporting secure, remote access to clinical data.
- Built automated tools for processing, analyzing, and reporting clinical datasets, improving efficiency and data quality.
- Developed backend services and APIs to streamline data management and integration with research systems.

- Collaborated with clinicians and AI researchers to align technical development with medical research goals.

Publications

Multimodal generative AI jointly learns pathology and clinical data to synthesize a multinational lung cancer cohort

- 2025
- HM Hieromnimon, V Miskovic, M Sacco, A Ferrarin, et al.

Aiding data retrieval in clinical trials with large language models: The APOLLO 11 Consortium in advanced lung cancer patients

- 2025
- F Corso, L Mazzeo, V Peppoloni, et al.

Multimodal AI using host, tumor, and ghost biomarker for predicting immunotherapy efficacy in NSCLC

- 2025
- V Miscovic, M Brambilla, L Mazzeo, R Ferrara, C Silvestri, A Ferrarin, et al.

AI multimodal tool highlights host and ghost biomarkers for guiding immunotherapy decisions in clinical practice

- 2025
- V Miskovic, L Mazzeo, R Ferrara, M Brambilla, C Silvestri, A Ferrarin, et al.

AI-based prediction of long-term survival in NSCLC patients treated with immunotherapy

- 2025
- V Miskovic, L Mazzeo, C Silvestri, A Ferrarin, et al.

Integrating radiomics and real-world data to predict immune-checkpoint inhibitors efficacy in advanced Non-Small Cell Lung Cancer

- 2024
- L Provenzano, M Favali, L Mazzeo, et al.

Open Problem: Leveraging Reinforcement Learning to Enhance Decision-Making in Oncology Treatments

- 2024
- A Zec, A Prejal, A Ferrarin, et al.

MetaLung: Towards a secure architecture for lung cancer patient care on the metaverse

- 2023
- M Zanitti, M Ferens, A Ferrarin, et al.

Education

Politecnico di Milano, MS in Computer Science and Engineering

- Sept 2020 – May 2023
- Graduated with a score of 105/110
- Focus on machine learning, data science, and AI model development.

Politecnico di Milano, BS in Computer Engineering

- Sept 2016 – Sept 2020
- Graduated with a score of 87/110

Skills

- Languages: Python, Java, C, SQL, TypeScript
- AI/ML Frameworks: scikit-learn, TensorFlow
- Data Processing: Pandas, NumPy, SQL
- Deployment & DevOps: Docker, AWS, GitHub Actions
- API & Web Development: FastAPI, Angular, PostgreSQL
- LLMs & Generative AI: Experience integrating and fine-tuning large language models via APIs, LoRA, prompt engineering, and LangChain for LLM-powered applications
- Experiment Tracking: Neptune.ai
- Collaboration & Version Control: Git, GitHub, experienced in working in cross-disciplinary teams