

About the Company

“We are passionate about finding easy solutions to complex matters”

The αHead-Research (<https://ahead-research.com/>) team is the AI R&D team of Spindox (<https://www.spindox.it/it/>), an Italian ICT company with over 1,300 employees. αHead boasts a number of highly skilled professionals whose careers range from research to business areas. Born from the experience of ACT Operations Research, αHead-Research is the result of more than 25 years of commitment to the areas of operations research, stochastic simulation, machine learning, and data intelligence that are currently grouped under the umbrella term Artificial Intelligence.

The knowledge gained in a large and heterogeneous set of business processes and experience in solving real-world problems have enabled αHead to effectively design and delivery complex AI solutions for big and international enterprise. Competence and experience are the solid foundations on which we build innovative solutions.

About the internship

Our collaboration with UniPD is now rooted over the years first with ACT Operations Research and now as αHead-Research. We have held more than 10 internships in collaboration with UniPD and as many as 5 of them have resulted in hires within the company. We are looking for candidates who are passionate about the world of AI and want to put what they have learned in their studies into practice on real industrial settings.

The internship can be completely remote, with the possibility of joining us in our Rome office. For sure, choosing to work completely remotely is also fine and does not affect the application.

About the projects

Large Language Models

We have been developing a proprietary framework for years, in order to (i) Utilize advanced language models, whether commercial or open-source, (ii) Deploy LLM-based pipelines on-premises or as a service, with a focus on data privacy and IP protection, (iii) Streamline fine-tuning, in-context learning, and prompt engineering for specific tasks, (iv) Effectively handle memory and knowledge to overcome LLM token and computational limitations, reducing infrastructure costs, (v) Create pipelines with multiple LLM-based agents for collaborative or competitive tasks, (vi) Support multi-modality in input (i.e., by storing images/ videos as vectors) and output (i.e., allowing to generate text and/ or images), including text and images, and (vii) Ensure deterministic LLM outputs to address hallucination issues.

We continuously develop new functionalities, and we are now addressing better techniques for finetuning and reinforcement learning with human feedback.



Timeseries forecasting

Within αHead-Research's Data Intelligence Department, which focuses on machine learning, descriptive statistics, signal processing, novelty detection, and quantitative methods for time series (TS) analysis. The candidate, supported by senior Data Research Engineers, may choose to work on the following topics operating on real industrial datasets:

- Demand curve forecasting;
- Predictive maintenance;
- Macroeconomics forecasting models;
- Self-Organizing Radio Networks.

Vehicle Routing problem addressed with Graph Neural Networks

The problem of finding the optimal delivery route for a set of customers is a famous NP-Hard problem called Vehicle Routing, where finding exact solutions is computationally expensive. The aim of this internship will be to apply a deep learning approach based on Graph Neural Networks, where the attractive aspect of GNNs is their fast computation times which can, potentially enable application to even greater problem sizes than current solvers can handle.

Clustering for time series

Clustering is an unsupervised technique for grouping related instances within a dataset when there is no prior knowledge about classes. In the context of time series, this task aims to identify efficiently homogeneous groups of similar patterns and temporal characteristics that appear in a dataset. Consequently, it is essential to address similarity and scalability challenges, especially in real-world applications such as industrial contexts. Under such considerations, this proposal aims to:

- study suitable clustering approaches for time series and develop a methodology for their comparison, highlighting the computational costs and the most representative partition through appropriate data visualization techniques by considering one or more real case studies.
- investigate the goodness of the final partitions, with particular attention to the degree of compactness and separation among the time series involved.

Automatic weed detection

A weed plant can be described as a plant that is unwanted at a specific location at a given time. The growth of weeds on railroad track beds and embankments raises potential issues: the vegetation endangers the security of workers inspecting the railway infrastructure and may result in poor track drainage, which could cause the embankment to collapse. Usually, weed detection is carried out by a driver of a truck who sprays various amounts of herbicide based on the amount of weed that is observed.

In this internship, the student will study and implement image processing methods for performing automatic weed detection and will work on the definition of an algorithm that indicates the quantity of herbicide to be used in near real time based on the quantity of weed identified. Actual videos obtained from cameras mounted on trucks used by local railway maintenance contractors will be used.

Vehicle Routing Problems with Fuel Consumption and Stochastic Travel Speeds

Conventional vehicle routing problems (VRP) always assume that the vehicle travel speed is fixed or time-dependent on arcs. However, due to the uncertainty of weather, traffic conditions, and other random factors, it is not appropriate to set travel speeds to fixed constants in advance. The aim of this internship is to design a mathematic model for calculating expected fuel consumption and fixed vehicle cost where average speed is assumed to obey normal distribution on each arc which is more realistic than the existing model.

Collaborative city logistics system

The objective of this thesis is to examine an innovative system for collaboratively organizing deliveries for a n-tier hyperconnected city logistics system, mixing math optimization and machine learning. The student will



focus on the tactical planning of services within the first tier of the system, in particular from external areas generally located on the periphery of the city (logistics platforms, distribution centers/urban consolidation, etc.) to the satellites from which goods are distributed to end customers. The distinguishing feature of this model is that we consider a coalition of carriers and logistics operators who share their resources (vehicle fleets and storage capacity) and information flows to provide more efficient services, thereby reducing costs and environmental impact.

Quantum Optimization and Quantum Machine Learning

Due to the inherent speedups of quantum effects, Quantum Computing promises to overcome computational limitations with better and faster solutions for optimization problems. While quantum computing is expected to have a significant impact on various industries, many questions and challenges remain: Which specific optimization problems can be solved? How can this quantum speedup advantage be translated into business impact?

Accordingly, the objective of this thesis is to examine innovative Quantum Optimization methods to solve combinatorial optimization problems arising from real-world applications.

The same scientific questions apply to machine Learning models definition and training, like e.g. quantum Support Vector Machines or quantum clustering.

Contacts

Tommaso.colombo@spindex.it

Stefano.leonardi@spindex.it

Alessandro.pinzuti@ahead-research.com

Marco.boresta@ahead-research.com

