

Machine Learning Engineer Demand Forecasting (f/m/d)

FULL-TIME · WÜRZBURG · HYBRID

Who we are

SOLISTIQ builds an AI inventory planner that brings forecasting, optimization and execution into one product. Built for use across industries, it powers replenishment and production planning for customers in wholesale, retail, manufacturing and food service. At its core is **Q-Foundry**, our own time series foundation model, pre-trained on millions of demand series and billions of data points across industries.

Founded in 2026, SOLISTIQ is a spin-off from the University of Würzburg's Data Driven Decisions group, where Q-Foundry originated as a research project. We are based in Würzburg and are expanding our engineering team to develop the model and product further.

About the role

With us, you won't just fine-tune existing models or build API wrappers: you'll help us build the next generation of Q-Foundry, our own time series foundation model—from designing the architecture and training it from scratch to getting it into production. As one of our first engineering hires, you'll work directly with the founders, make core technical decisions and help build the team along the way.

Your work spans the full model development cycle: understanding forecast errors in production, developing and testing new approaches, improving pre-training and adapting the model to specific industries. You'll take projects from initial experiments through deployment, with the scope to make consequential technical decisions and see how your work improves customers' planning.

What you'll work on

- **Infrastructure:** Build efficient data pipelines for large-scale pre-training and inference, from resource-efficient data loading and validation to reliable distributed processing of millions of demand series throughout week-long training runs.
- **Model architecture:** Improve our time series transformer's efficiency and capabilities. Current areas of interest include sparse covariate handling, text inputs and in-context learning across related products.
- **Pre- and mid-training:** Refine the training setup—scaling approaches, loss functions, optimizers—for training stability and forecast quality. Shape data mixtures and curricula for balanced learning across demand patterns, industries and time scales, including rare patterns and edge cases.
- **Fine-tuning and post-processing:** Develop adaptation strategies for specific industries, customers and applications without losing generalization, and build calibration, bias correction and hierarchical reconciliation into the final forecasts.

Stack: Python, PyTorch, Polars, PostgreSQL, MLflow, Slurm, Docker, Kubernetes, FastAPI.

What we're looking for

- Fluent in Python, PyTorch and SQL. You write production-quality code: tested, maintainable, and used by others.
- You have taken ML models from prototype to production and kept them running: monitoring, debugging, retraining.
- You have trained deep learning models on datasets that don't fit in memory, on multiple GPUs, and you know how data loading, GPU memory and distributed training affect throughput and reliability.

- You can take a model's errors apart: given a customer's data, you say with evidence why the forecast is off and what would fix it.
- You design experiments with a clear hypothesis, a strong baseline and a short turnaround.
- You enjoy taking ownership in a small, ambitious team: setting priorities when the path isn't clear, adapting as you learn, and putting in extra effort when a critical milestone calls for it.
- MSc or PhD in a quantitative field – computer science, mathematics, physics, statistics, engineering, economics – and roughly three years of machine learning or data science, in industry or research; equivalent experience counts.

Nice to have

- Experience in an early-stage startup or building a product or team from the ground up.
- Experience modifying transformer architectures or pre-training models, whether in language, vision, time series or another domain.
- Experience with fine-tuning, adapters or transfer learning to adapt pre-trained models to new tasks.
- Experience with probabilistic modeling, uncertainty estimation or calibration.
- Familiarity with time series forecasting or inventory planning is a plus; you learn the domain on the job.

How we work

You keep the team informed of progress and challenges, discuss technical decisions openly and ask for help early, sharing what you've tried and where you're stuck. We work through problems together and give each other direct feedback.

We work in German and English; either is enough.

What we offer

- A permanent, full-time position with SOLISTIQ GmbH.
- Hybrid work from our office in central Würzburg, with at least two days per week on site.
- Competitive compensation, matched to your experience, with the option of equity participation depending on your scope of responsibility.
- EGYM Wellpass membership.
- An office where people enjoy working together, with great equipment, drinks, snacks and dinner on us.

Apply

Send us your CV and a link to something you have built: a repository, a project, a paper or your thesis. Write to Magnus Maichle, application@solistiq.com.

HOW IT GOES

- 1 **Initial conversation:** 45 minutes with one or two founders to discuss your experience, the role and mutual expectations.
- 2 **Technical discussion:** A focused ML case followed by a discussion with our Chief AI Officer. We share the format and expected preparation time in advance.
- 3 **Meet the team in Würzburg:** Spend a few hours with us, see how we work, and join us for dinner.
- 4 **Offer:** Discuss compensation, responsibilities and a start date.