ZAKI RASA

> Data Scientist & Founder

forecasting / full-stack products / ML pipelines / data governance

xakirasa@gmail.com / LinkedIn / zakirasa.pages.dev / forecastmyshop.com / Porto, Portugal

01 PROFESSIONAL SUMMARY

Data Scientist and founder who builds forecasting products end to end: from deep-learning research and production ML pipelines at SISQUAL WFM to **ForecastMyShop**, a forecasting service for independent retailers that I designed, built, and launched as a full-stack product. Spans statistical modelling, data engineering, APIs, web development, and data governance, with a growing focus on RAG and enterprise LLM systems. A clear communicator who works weekly with the Chief of Innovation and the Chief Developer and shares knowledge through workshops.

02 PROFESSIONAL EXPERIENCE

Founder & Full-Stack Engineer / ForecastMyShop (forecastmyshop.com) — 2026 – Present

- Founded and built a forecasting service for independent shops (pharmacies, grocery, fashion, beauty, restaurants, online): owners upload an existing sales file and get per-product forecasts with ranges, order suggestions, and insights.
- Designed the forecasting engine around **Chronos-2**, Amazon's open-source foundation model, alongside classical methods; each forecast is back-tested on the shop's own recent sales, the best method wins, and a simple baseline is kept when it matches (a bias for simplicity).
- Built the ingestion and data-quality layer for messy real-world inputs (Excel, till exports, monthly-column sheets, CSV, copied tables), detecting gaps, spikes, invalid dates, and duplicates before any forecast runs.
- Delivered the full product stack solo: web app with accounts and plans (free, standard, assisted), credit-based usage and billing, Excel/PDF/CSV exports, customer-facing API access, and a **9-language** interface.
- Privacy by design: EU-hosted servers, uploads deleted after 24 hours and results after 72 hours, a data-processing agreement, no training on customer data.

Data Scientist / SISQUAL WFM, Porto, Portugal — Feb 2025 – Present

- Design and maintain FastAPI endpoints for model training, prediction, evaluation, and model management in a production workforce-forecasting pipeline on SQL Server, with AWS SageMaker training and inference and S3 artefact versioning.
- Integrate and benchmark forecasting approaches including PyTorch-based models and AutoGluon, selecting production models on accuracy and latency trade-offs; reduced WMAPE by **18%** through feature engineering and Optuna-based hyperparameter tuning.
- Design and run Python A/B testing methods to compare candidate forecasting models against the production baseline, measuring incremental impact and speeding up model testing and iteration.
- Build ingestion, cleaning, resampling, and feature pipelines with external signals (weather, football fixtures, holidays, events) and automate GPS event-data processing across client deployments.
- Lead the DataHub rollout as an enterprise data catalog (domains, ownership, glossary, lineage, PII classification) and co-design a governed semantic/query layer with MCP- and RAG-based prototypes for enterprise LLM assistants.
- Work side by side with the Chief of Innovation and the Chief Developer in weekly meetings: present results, translate business needs into requirements, and contribute to design decisions and the data/AI roadmap.
- Drive research-led innovation by benchmarking deep-learning forecasting methods before production, and share knowledge through workshops, demos, and documentation for technical and non-technical teams.

Data Analyst / SISQUAL WFM, Porto, Portugal — Aug 2023 – Feb 2025

- Built Python and SQL ETL workflows for forecasting and reporting, and improved preprocessing, data-quality checks, and feature preparation for models.
- Developed automated dashboards and reporting workflows, cutting manual querying and reporting time by roughly **40%** for stakeholder teams.

Computer Systems Engineer / Ministry of Higher Education, Afghanistan — Feb 2019 – Aug 2021

- Modernized the national Higher Education Management Information System (HEMIS), a Laravel/PHP platform with an HTML/CSS/Bootstrap frontend used by **500+** university IT staff, and built its analytics and reporting modules.

03 EDUCATION & SELECTED RESEARCH

// Both degrees fully instructed in English

MSc in Data Science and Engineering / University of Porto (FEUP), Portugal — 2022 – 2025

THESIS *Improving Pharmacy Sales Prediction: A Global Forecasting Approach* (FEUP; dataset confidential)

External Researcher / LIAAD, INESC TEC — Aug 2023 – Oct 2024

- Implemented and benchmarked DeepAR, N-BEATS, and XGBoost global forecasting models against SARIMAX and Holt-Winters baselines, using PySpark to process a confidential real-world dataset of **20M** rows covering 119 community pharmacies.
- Improved prediction accuracy by **9%** versus an autoregressive baseline through feature engineering and Optuna-based hyperparameter tuning.
- Completed academic projects in CNN-based wind-turbine defect detection using ResNet, MobileNet, and EfficientNet; LLaMA-based question answering; and transformer-based sentiment analysis using Spark and CoreNLP.

BSc in Computer Systems Engineering / UET Peshawar, Pakistan — 2014 – 2018

- Fully-funded scholarship awarded by the Prime Minister of Pakistan, covering the full degree for international students.

04 SKILLS

Programming & Data	Python · NumPy · pandas · scikit-learn · PySpark · R · SQL · SQL Server · C# · Matplotlib · Plotly · Power BI · Tableau
Forecasting & Statistics	Chronos-2 · DeepAR · N-BEATS · SARIMAX · Holt-Winters · global forecasting · back-testing · A/B testing · hypothesis testing · statistical inference
Machine Learning	PyTorch · XGBoost · AutoGluon · Optuna · CNNs · transfer learning · transformer/NLP fundamentals
Full-Stack & Product	Web applications · REST APIs · FastAPI · Laravel/PHP · HTML/CSS · authentication & billing · multilingual (i18n) · exports · product & pricing
Data Eng. & MLOps	ETL pipelines · data ingestion · data quality · feature pipelines · MLflow · Weights & Biases · Databricks ML · model registries
Cloud & Infrastructure	AWS SageMaker · S3 · Google Cloud · Vertex AI · Docker · Git · EU hosting · Azure (working exposure)
Data Governance	DataHub · metadata management · data cataloguing · PII/sensitivity classification · business glossary · lineage
Generative AI	Hugging Face · OpenAI API · Anthropic/Claude API · LangChain · LlamaIndex · Ollama · RAG · MCP-based architectures

05 SELECTED ADDITIONAL EXPERIENCE

Volunteer / European Solidarity Corps, EU — 2022 – 2023

- Mentored refugees and migrants and provided translation support (Refugees Welcome, JRS, CPR); co-organized a youth exchange in Bosnia and joined exchanges in France, Portugal and Spain.