NIKITA BAKLAZHENKO

MACHINE LEARNING ENGINEER

Barcelona, Spain · +34 663 686 669 · · linkedin.com/in/nikita-baklazhenko

EXPERIENCE

Expert Data Scientist   Sep 2023 – Present
Novartis   *Barcelona, Spain*

– Product owner of 2 ML products covering €5B+ in yearly costs across 20+ countries, driving roadmap and stakeholder alignment.
– Built automated cloud forecasting pipelines running monthly without intervention — now the central baseline for analysts and senior finance managers.
– Achieved significant accuracy improvements over human forecasts, saving 150+ analyst hours/month.
– Architected CI/CD pipelines for Dockerized ML projects; built a model registry, custom Databricks connector, and internal Python libraries.
– Led LLM proof-of-concept initiatives — evaluating API-based and local deployments, collaborating with leadership on production strategy.
– Delivered training sessions for analysts and cross-functional teams on ML tooling and forecasting workflows.

Python, Jenkins, AWS, SQL, Docker, Power BI, Streamlit, Snowflake, Databricks

Data Scientist   Mar 2023 – Jul 2023
Citibeats   *Barcelona, Spain*

– Large-scale data scraping, NLP, and unsupervised ML on unstructured text to analyze social trends.
– Worked with Twitter API, Hugging Face transformers, and Llama models for text classification and generation.

Teaching Assistant   2019 – 2022
National Research University Higher School of Economics   *Saint Petersburg, Russia*

– Led seminars, facilitated Q&A sessions, and graded student papers in economics and statistics on a part-year basis.

Consultant   Jan 2021 – Jun 2021
PricewaterhouseCoopers (PwC)   *Saint Petersburg, Russia*

– Performed transfer pricing analysis: benchmarking financials, assessing performance metrics, and examining global corporate structures.

SKILLS

GenAI & LLMs: Agentic AI, RAG, Multi-Agent Orchestration, Prompt Engineering, Fine-Tuning (LoRA), LangChain, OpenAI Agents SDK, MCP, Pydantic, Langfuse, Ollama

ML & Data Science: PyTorch, TensorFlow, Scikit-learn, CatBoost, XGBoost, PySpark, NLP, Transformers, Time Series Forecasting, Anomaly Detection, A/B Testing, Causal Inference, SHAP/LIME, Feature Engineering

MLOps & Cloud: AWS (SageMaker, Lambda, Step Functions, EC2, S3), GCP (Cloud Run), Docker, Jenkins, CI/CD, Terraform, MLflow, Git

Data & Backend: Python, SQL, Snowflake, Databricks, Pandas, Polars, NumPy, SciPy, DuckDB, FastAPI, React, Power BI (DAX, Power Query), Streamlit

Languages: English (Fluent), Russian (Native), Spanish (Intermediate)

SIDE PROJECTS

PolicyArena.dev — Agent-Based Simulation Engine | pip install policy-arena
Simulation engine comparing rule-based, RL, and LLM-powered agents across 15+ game-theoretic scenarios. Live-streamed simulations, per-agent analytics, multi-run comparisons, and data export. Open-source on PyPI. Supports OpenAI, Anthropic, Gemini, DeepSeek, and Ollama.

TripGen.org — Agentic AI Travel Planner
10+ parallelized agents with MCP tool integration and conversational chat generating day-by-day itineraries with activities, photos, and maps. Python + FastAPI on GCP, React, Langfuse, Pydantic.

Hoodwise.net — AI-Powered Neighborhood Finder
Scores neighborhoods on safety, affordability, commute, and lifestyle worldwide. Interactive maps, long-term vs. short-term modes, SEO-optimized city pages. FastAPI on GCP, React, Langfuse, Pydantic.

EDUCATION

Master's Degree in Data Science   2022 – 2023
Barcelona School of Economics   *Barcelona, Spain*
Thesis: Early Prediction of Preeclampsia from cf-RNA with ML

Bachelor's Degree in Economics   2018 – 2022
National Research University Higher School of Economics   *Saint Petersburg, Russia*
Thesis: Statistical analysis of excess mortality factors in Russia during COVID-19

Erasmus Program in Economics   2020 – 2021
University College London (UCL)   *London, United Kingdom*