

Hasan Hüseyin Yurdagül

SENIOR ML ENGINEER · LLM AGENTS & APPLIED AI

Adana, Türkiye
Relocating to the Netherlands / EU
hhyurdagul@gmail.com
hhyurdagul.com
github.com/hhyurdagul
linkedin.com/in/hhyurdagul

PROFILE

Senior Machine Learning Engineer and hands-on Technical Lead with 5+ years shipping ML to production. Builds generative AI and large language model (LLM) systems end to end: retrieval with offline evaluation, tool-calling agents (LangGraph, MCP), LoRA fine-tuning, vLLM serving, and human-in-the-loop review, deployed as Docker services on Azure. Leads AI R&D for products serving 80+ municipalities, ministries, and utilities, with depth in forecasting, computer vision, and recommenders.

EXPERIENCE

Universal Software (Universal Yazılım)

Istanbul-based · Remote

Technical Lead, AI R&D

Jan 2024 - Present

- Lead the hands-on AI R&D roadmap, architecture reviews, delivery planning, and stakeholder coordination; mentor up to 5 engineers building ML products for municipalities, ministries, and utilities.
- Architected **MdsAI**, a human-in-the-loop support intelligence platform for the FlexCity helpdesk serving 84 municipalities (14 metropolitan, 70 district), built on 132K support events and 29K tickets.
- Built evidence retrieval over past tickets and support history and gated releases on offline evaluation, reaching 97% Recall@5.
- Fine-tuned Gemma 4 (12B, 27B) and Qwen models for Turkish support language and served them on-premises with vLLM for parallel inference: drafts return in under a minute, at near-zero cost per ticket, with no data leaving the premises.
- Put LLM drafts and routing into the workflow with staff approving every reply, cutting median first-response time by 50%; now moving auto-routing and category correction to decision models (Jev, Laya, Clef).
- Co-lead a TÜBİTAK 1832 national R&D grant with hydraulic engineers; designed attention-based seq2seq river-discharge forecasting with event-based backtests and FastAPI serving for 8 gauging stations (Teus).

Senior AI/ML Engineer

Jan 2023 - Dec 2023

- Shipped a real-time Enerjisa field-photo validation pipeline combining YOLOv8 detection with blur, duplication, and empty-frame checks, reducing invalid submissions reaching inventory review by 30%.
- Built predictive-maintenance classifiers for Enerjisa assets and CNC assembly lines, identifying 82% of failures within a 7-day warning window.
- Developed LSTM pipelines for municipal water-demand and industrial gas/electricity forecasting, improving MAPE by 21% relative to seasonal-naive baselines.

AI/ML Engineer

Apr 2021 - Dec 2022

OxyAI

Remote, Türkiye

AI/ML Engineer (Part-time Contract)

May 2021 - Jan 2023

- Built production recommenders for the online art gallery **Collectors** and audiobook app **Dinlebi**, combining multimodal and collaborative-filtering approaches with cold-start onboarding; increased recommendation CTR by 14%, and both systems remained in production for 3+ years.

SELECTED PROJECTS

Muninn - LLM Agent for a Markdown Knowledge Vault

github.com/hhyurdagul/muninn

Open-source Rust agent that reads, searches, and edits a Markdown knowledge vault through a streaming tool-call loop. A seven-crate workspace (~7.5K lines) with a CLI, a TUI, and an MCP server that Claude Code and Codex call every day. Talks to Anthropic natively with prompt caching, and to OpenAI, Gemini, OpenRouter, or any OpenAI-compatible endpoint. Sandboxes file access to the vault, routes edits through an approval step, and journals every change with conflict-aware undo; packaged with Nix.

Rust · Tokio · Ratatui · MCP (JSON-RPC) · Anthropic & OpenAI-compatible APIs · Nix

Fishometry - Single-Image Fish-Length Estimation (MSc thesis)

hhyurdagul.com/projects/fishometry

Estimates fish length from one photo without depth sensors or reference objects: YOLO landmarks, Depth Anything V2, and SAM segmentation feed DINOv2 features into a species-aware ridge regressor. Reached 5.53 cm MAE vs. 8.76 cm for a species-mean baseline (37% lower) on 348 held-out outdoor photos.

Python · PyTorch · DINOv2 · YOLO11 · SAM · Depth Anything V2 · scikit-learn · Polars

Teus - River-Discharge Forecasting

hhyurdagul.com/projects/teus

Built a river-discharge forecasting pipeline for 8 gauging stations: expanding-window backtests on held-out flood events, attention-based seq2seq models compared against persistence and foundation-model baselines, reproducible runs with champion promotion, and FastAPI serving that validates input contracts before hot-swapping checkpoints.

Python · PyTorch · scikit-learn · Polars · FastAPI · TimesFM

CONSULTANCY

Selected Consulting (Part-time)

2020 - Present

Alongside employment · Remote, Türkiye

[Selected consulting work](#)

- **Moka** (fintech): built churn/lifetime-value scoring, price-quote prediction, and financial forecasting behind FastAPI with validated feature contracts, packaged with Docker and deployed on Azure; the merchant-churn model reached 0.93 F1 / 0.95 ROC-AUC in nested cross-validation.
- **Inveon** (e-commerce): built semantic product search and a mixture-of-experts classifier reaching 91% top-1 accuracy across 50+ subcategories, in production since 2023, plus order cancellation/return prediction APIs.
- Also: forecasting and staffing APIs for **Ikas**, a forecasting workbench used across 30+ organizations (**MLStudio**), and AI/R&D team lead at **Smartiks** (2023 - 2026).

SKILLS

- **LLMs & Agents:** retrieval-augmented generation (RAG), embeddings and semantic search, tool-calling agents, LangGraph, MCP servers, LoRA fine-tuning (Gemma, Qwen), vLLM serving, Ollama, Milvus, semantic chunking, prompt caching, evaluation with RAGAS, LangSmith, and LLM-as-judge scoring, human-in-the-loop review
- **Programming & Data:** Python, Rust, SQL, Bash, Polars
- **ML Frameworks:** PyTorch, TensorFlow/Keras, scikit-learn, XGBoost/LightGBM, sentence-transformers, BERTopic
- **ML Systems & Deployment:** FastAPI, Docker, Azure (model deployment), on-premises GPU serving, MLflow, Weights & Biases, CI/CD, model versioning, Kubernetes (working knowledge)
- **Also:** time-series forecasting (LSTM, seq2seq + attention, Transformers), computer vision (YOLO, segmentation, depth estimation, DINOv2), churn/CLTV, recommenders, anomaly detection
- **Languages:** Turkish (native), English (C1; IELTS 7.5)

EDUCATION

Çukurova University

Adana, Türkiye

MSc in Computer Engineering

Aug 2023 - Nov 2026 (expected)

Thesis: *AI-Driven Object Size Prediction in Visual Data Without Depth Information and Reference Points*

BSc in Computer Engineering (English-medium)

Sep 2016 - Aug 2021

PUBLICATIONS

- *Development of Machine Learning-based Revenue Forecasting Models for Electronic Payment and Money Institutions*. IEEE ISMSIT, 2025. [DOI](#).
- *Fault Prediction in Energy Systems: Telemetric Data and Machine Learning Approaches (First Author)*. IEEE ISAS, 2025. [DOI](#).
- *Development of Machine Learning Based Price Quote Prediction Models for Electronic Payment and Money Institutions*. IEEE ISAS, 2025. [DOI](#).
- *Multi-Class Machine Learning Models for Predicting Order Cancellations and Returns in E-Commerce*. IEEE ISAS, 2025. [DOI](#).
- *Forecasting Call Center Arrivals Using Machine Learning*. Osmaniye Korkut Ata University Journal of the Institute of Science and Technology, 2021. [DOI](#).