

Natnael Masresha Zerihun

Machine Learning & Data Engineer
Helsinki, Finland | [LinkedIn](#) | [GitHub](#) | [Medium](#)

PROFILE

Software engineer with 3+ years of professional experience across machine learning, data engineering, and full-stack development (Python, C#/.NET, TypeScript). Built production ML/LLM pipelines and creative-automation tooling in ad tech, ELT data platforms (Airflow, dbt, Kafka, Spark), and real-time sensor-data systems. M.Sc. in Smart Systems adds a rare hardware dimension, FPGA/VHDL digital design and MEMS sensor characterization, enabling work from device signals to deployed services.

EDUCATION

Aalto University USN BME <i>Erasmus Mundus Joint M.Sc. in Smart Systems Integrated Solutions (SSIs)</i>	Finland Norway Hungary Aug. 2024 – Aug. 2026
Embedded systems, digital design (FPGA/VHDL), MEMS sensor systems, and system-level integration across three universities; thesis in industry on thermal-camera calibration software and algorithms.	
10 Academy <i>Data Engineering ML Engineering Web3</i>	Ethiopia Aug. 2022 – Nov. 2022
Intensive training with weekly end-to-end projects: data pipelines, ML models, REST APIs, and dashboards (see Projects).	
Addis Ababa University <i>B.Sc. in Electrical and Computer Engineering Computer Engineering major</i>	Ethiopia Sept. 2015 – Dec. 2020

EXPERIENCE

Thermal Imaging Platform & Calibration Software (M.Sc. Thesis) <i>IDEAS – Integrated Detector Electronics AS</i>	Jan. 2026 – June 2026 Oslo, Norway
- Built a real-time streaming pipeline for 16-bit thermal video: RTP-over-UDP packet parsing and frame assembly (640×480 @ 10 fps), with TCP and WebSocket forwarder modes for remote camera nodes. - Developed a FastAPI backend with per-client sessions, authenticated forwarders, producer–consumer frame distribution, live histogram/statistics computation, and incremental binary recording with background MP4 encoding. - Implemented calibration and drift-compensation algorithms in NumPy (one/two-point correction, a 95-matrix lookup-table library, temperature-indexed interpolation, scene-based correction) running on the live stream, with image-quality metrics for automatic correction-matrix selection. - Built a React + TypeScript web client: live visualization (Plotly.js), camera control, pixel inspection, recording management, and client-side PDF reporting.	
ML Engineer & Fullstack Developer <i>Adludio Ltd</i>	Jan. 2023 – Sept. 2024 London, UK (remote)
- Developed the "Creative Design Manager Studio", a web platform enabling non-technical users to build production-ready interactive ad creatives. - Reduced creative delivery time by ≈85% (2–3 weeks to 3 days) via reusable abstract game/ad logic and a structured JSON-based ad template adapting to multiple interaction types. - Managed and optimized LLM pipelines (LangChain) for ad-concept generation and automated creative variants. - Applied computer vision for frame metadata extraction, template generation, interaction-type recommendation, brand-compliance checking, and KPI-based creative scoring. - Embedded click, hover, view-time, and interaction trackers into generated creatives.	
Fullstack Developer <i>Upwork Freelancer</i>	Feb. 2022 – Nov. 2022 Remote
Achieved 100% Job Success score delivering 8+ projects (React front-ends, ASP.NET backends, REST APIs, auth flows, admin dashboards).	
Fullstack Developer R&D <i>CNET Software Technologies</i>	March 2021 – Sept. 2022 Addis Ababa, Ethiopia
- Developed APIs, dashboards, and ERP web applications using .NET Core, Blazor, SQL, and Python. - Created a company-wide LMS used by 500+ employees, reducing new-hire training time by ≈30%.	

DATA ENGINEERING & ML PROJECTS

- Traffic Data Warehouse (ELT)** | *Airflow, dbt, PostgreSQL, Redash, Docker* 10 Academy
• Designed a fully dockerized ELT data warehouse for vehicle-trajectory traffic data: Airflow-orchestrated DAGs loading raw data into PostgreSQL, dbt transformation models with hosted documentation, and Redash dashboards for analysts. [GitHub](#) | [Medium](#)
- LLM News-Relevance Scoring & Entity Extraction** | *Cohere, FastAPI, DVC, MLOps* 10 Academy
• Built an LLM evaluation pipeline scoring news items for topic relevance (0–10) and extracting entities from job descriptions via prompt engineering and fine-tuning on Cohere models.
• Served the models through FastAPI endpoints with DVC-versioned datasets (v0–v2 train/test splits) and modular prompt design so LLM providers and topics can be swapped without pipeline changes. [GitHub](#) | [Medium](#)
- Speech-Corpus Collection Pipeline (team project)** | *Kafka, Spark, Airflow, S3* 10 Academy
• Co-built a distributed data-collection platform for Amharic speech: text prompts and audio recordings streamed through Kafka topics, processed with Apache Spark, orchestrated by Airflow, and stored in S3; React recording front-end with FastAPI services. [Medium](#)
- Pharmaceutical Sales Forecasting** | *LSTM, Random Forest, Streamlit, DVC* 10 Academy
• Built an end-to-end ML solution forecasting store sales six weeks ahead from promotions, competition, holidays, and seasonality (Rossmann dataset, 25+ features); deployed an interactive Streamlit dashboard with DVC-versioned data. [GitHub](#) | [Medium](#)
- A/B Testing & Causal Inference for Ads and Logistics** | *Statistics, Bayesian networks* 10 Academy
• Developed an A/B hypothesis-testing framework (classical and sequential testing) measuring brand-awareness lift of ad campaigns. [GitHub](#)
• Applied causal inference (Bayesian networks) and K-means clustering to delivery-fleet location data to identify drivers of unfulfilled orders and answer counterfactual intervention questions. [GitHub](#)
- Wildlife Acoustic Monitoring System** | *PyTorch, Raspberry Pi, TDOA* Aalto, Finland
• Built a solar-powered Raspberry Pi sensor node with a microphone array for real-time elk monitoring; trained a few-shot binary sound classifier (PyTorch, modified YAMNet) on augmented elk-vocalization data and implemented time-difference-of-arrival (TDOA) localization of animal calls.

ELECTRONICS & FPGA (M.Sc.)

- FPGA Frequency-Measurement & Classification System** | *VHDL, ModelSim, Quartus* USN/BME
• Designed a hierarchical VHDL system (parameterized frequency-measurement core, band detectors with temporal filtering, two-stage input synchronizer) for real-time acoustic event classification; verified in ModelSim, synthesized with Intel Quartus Prime (TimeQuest STA) on a Cyclone III FPGA.
- MEMS Design, Fabrication & Test** | *Cleanroom, COMSOL, ANSYS, LabVIEW* USN/BME/Aalto
• Executed a full cleanroom microfabrication flow on Si wafers (photolithography, ICP/KOH etching, PVD); designed MEMS sensors in COMSOL/ANSYS with reduced-order models; automated hardware-in-the-loop DC-DC converter validation with LabVIEW and NI DAQ.

TECHNICAL SKILLS

Languages: Python, C#, TypeScript/JavaScript, SQL, VHDL, C/C++
ML & MLOps: PyTorch, scikit-learn, OpenCV, LangChain, LLM prompt engineering & fine-tuning (Cohere), MLflow, DVC, Streamlit, NumPy/Pandas
Data Engineering: Apache Airflow, dbt, Kafka, Spark, PostgreSQL, Redash, ELT/ETL design, AWS S3, Docker/Docker Compose
Backend & Frontend: FastAPI, ASP.NET Core, Blazor, REST, WebSockets, TCP/UDP socket programming, React, Plotly.js
Electronics & Embedded: FPGA (VHDL, ModelSim, Intel Quartus Prime), Raspberry Pi, Arduino, LabVIEW (NI DAQ, HIL testing), sensor calibration & characterization
Practices: Git, TDD, design patterns, data versioning, technical writing (8 published articles)