

Enes Caglar

Data & AI/ML Engineer

Heilbronn, Germany • mail.enes.caglar@gmail.com • [linkedin.com/in/enescaglar](https://www.linkedin.com/in/enescaglar) • github.com/enescaglar • enescaglar.com

Data & AI/ML Engineer, B.Sc. Computer Science and Engineering (Sabancı University), incoming M.Sc. Information Engineering at TUM (Oct 2026). Fast at turning ideas into working prototypes: a first-place AWS Bedrock hackathon build, production BI pipelines, custom FP16 CUDA kernels and RAG systems. Seeking Werkstudent and internship roles in data and AI engineering in Germany.

EDUCATION

Technical University of Munich

M.Sc. in Information Engineering

Heilbronn, Germany

Starting Oct 2026

Sabancı University

B.Sc. in Computer Science and Engineering, GPA: 3.38/4.00

Istanbul, Turkey

Sep 2021 – Jan 2026

SKILLS

Data Engineering: SQL (SQL Server, MySQL), PySpark, Microsoft Fabric, Delta Lake, Power BI / DAX

ML & Deep Learning: PyTorch, Hugging Face Transformers, scikit-learn, Keras/TensorFlow, Pandas, NumPy

GPU & Deployment: CUDA, TensorRT, ONNX Runtime, Nsight Systems

LLM & RAG: LangChain, ChromaDB, Vanna AI, Amazon Bedrock

Languages & Tools: Python, SQL, C++, Git, Streamlit

EXPERIENCE

Vispera

Data & Insights Intern

Istanbul, Turkey

Jun 2025 – Aug 2025

- Led end-to-end migration of Power BI reporting to Microsoft Fabric, cutting a scheduled weekly KPI pipeline **~75% (~45s → ~11s)** via column pruning, predicate pushdown and two-week windowing over four joined Delta tables.
- Validated the Fabric Warehouse schema and built dimensional data models, including a DAX master date table, standardizing time-based filtering across reports.
- Developed a PySpark QC pipeline validating GPS-based audit records via Haversine-distance checks: all **2,208** rows matched source geodistance within **0.2m**, **97%** within **0.1m**, with automated pass/fail reporting to Microsoft Teams.
- Designed a GPS-based fraud-detection pipeline combining **5** statistical anomaly tests with DBSCAN clustering, validated against manual reviewer judgments; found **~15%** of audit records were lost to missing GPS coordinates and proposed a recovery method projected to cut the loss to **~10%**.

Awarded a Vispera Internship Excellence Certificate.

iLAB

Data Science Intern

Istanbul, Turkey

Jan 2024 – Feb 2024

- Engineered Frequency/Monetary features from **320,148** clickstream sessions across **130,346** unique clients for a large Turkish real-estate classifieds platform, with category-aware price tiers for rent vs. sale listings.
- Segmented clients with K-Means on Frequency/Monetary and engagement features (**k=10** via elbow method), visualized with t-SNE and Plotly segment-profile dashboards.

PROJECTS

Wealth Management Text-to-SQL

Aug 2026

Personal Project • github.com/enescaglar/wealth-management-text-to-sql

- Built two natural-language-to-SQL approaches over a SQL Server wealth-management warehouse (advisors, clients, accounts, assets, transactions, portfolios): a Vanna AI model trained on schema/DDL/docs (RAG via ChromaDB), and a zero-shot LangChain SQL chain.
- Measured head to head on a 30-question suite (12 multi-table joins, 8 aggregation/group-by, 5 view-backed): RAG **30/30** execution accuracy vs **25/30** zero-shot baseline, over **8** Faker-seeded tables and **~15,000** rows.

MobileNet Optimization for Edge Deployment

Feb 2025 – Jan 2026

Undergraduate Graduation Project (ENS491/492) • github.com/enescaglar/mobilenet-edge-deployment

- Designed custom FP16 CUDA kernels for depthwise and pointwise convolutions, achieving up to **2.1x** and **2.6x** layer-level speedups over the PyTorch FP16 baseline on an A100.
- Built a PyTorch→ONNX→TensorRT inference pipeline reaching **~0.30 ms** per image at batch 1 on an A100 (**3,342 FPS**), using Nsight Systems profiling to isolate depthwise and pointwise convolution as the dominant cost.

SemEval-2026 Machine-Generated Code Detection

Sep 2025 – Jan 2026

Team course project (CS445, Group 18); Contribution: ModernBERT track • github.com/enescaglar/machine-generated-code-detector

- Owned the ModernBERT track of an 11-system team benchmark (4 lexical/tree baselines, 7 fine-tuned encoders): fine-tuned ModernBERT-base for human- vs. machine-generated code classification, reaching **0.997** validation Macro F1.
- Reported the track on SemEval's hidden test rather than validation alone: **0.997 → 0.28–0.35** Macro F1 under distribution shift, against a TF-IDF+SVM baseline that led the team benchmark at **0.46**, evidence the group's encoder-first assumption did not hold.

AWARDS & HACKATHONS

SOCAR Türkiye × AWS AI Mini Bootcamp, 1st Place

Nov 2025

Four-person team: an ethics-declaration risk agent on Amazon Bedrock and Claude Sonnet, built in one day; later rebuilt independently in Streamlit.

github.com/enescaglar/ethics-risk-agent

LANGUAGES

Turkish (native) • English (C1) • German (A1, in progress)