Mert Cihan Bayir

Artificial Intelligence Engineer — Machine Learning — Computer Vision — LLM AI

mertcihanbayir@gmail.com | LinkedIn | GitHub

Summary

Artificial Intelligence Engineer specializing in Machine Learning, Computer Vision, and Large Language Models (LLMs). Experienced in architecting and deploying production-grade AI systems, including end-to-end RAG platforms, biomedical signal processing applications, computer vision solutions, and mobile AI deployment. Passionate about building scalable AI products that solve real-world problems.

Education

Sakarya University – Software Engineering 2021 – 2025

- Bachelor's Degree in Software Engineering.

Experience

VEXPO – AI Engineer Jan 2026 – Aug 2026

- Architected, developed, and optimized an end-to-end **RAG (Retrieval-Augmented Generation)** platform achieving **94% response accuracy**. Through retrieval and LLM pipeline optimization, reduced inference costs to approximately **$60–70 per 2 million queries**.
- Built document ingestion, embedding, vector database, and retrieval pipelines to deliver scalable and context-aware LLM responses.
- Optimized retrieval quality through prompt engineering, retrieval tuning, and LLM pipeline improvements, significantly increasing response accuracy and reducing operational costs.
- Developed a **ClamAV**-based antivirus scanning service for **VexDrive**, enabling automated malware detection for uploaded files.
- Implemented an AI-powered content moderation pipeline for uploaded images and videos to detect inappropriate content and improve platform safety.
- Enhanced the profile photo upload workflow by integrating biometric photo validation, improving data quality and user experience.

Ford Otosan – Data Annotator 2023 – 2025

- Annotated large-scale autonomous driving datasets using **2D Bounding Boxes**, **Polygon Segmentation**, **Attributes**, and **3D LiDAR Bounding Boxes** for computer vision models.
- Prepared high-quality production datasets for object detection, semantic segmentation, and 3D perception models.
- Performed annotation quality assurance (QA) by reviewing and validating datasets to ensure consistency, accuracy, and compliance with project guidelines.
- Collaborated with AI engineers and annotation teams to maintain high-quality datasets for autonomous driving model development.

Internships

Ford Otosan – Artificial Intelligence Intern 2024

- Contributed to the development of a deep learning-based **Free Space Segmentation** model for autonomous driving perception systems.
- Participated in image preprocessing, dataset preparation, data augmentation, model training, and performance evaluation workflows.
- Developed and optimized semantic segmentation models using **Python**, **PyTorch**, and **OpenCV**.
- Improved model performance through hyperparameter tuning and training pipeline optimization.

MKU Technology – Artificial Intelligence Intern 2024

- Developed a machine learning model for early fetal sex prediction using biomedical signal processing and feature engineering techniques.
- Built a **Random Forest** classifier achieving **96.4%** accuracy in internal evaluations.
- Following my internship performance, I was invited to continue contributing to an AI research project focused on EEG-based intention recognition.
- Developed an EEG-based machine learning model capable of classifying **"Yes"** and **"No"** intentions using data collected with a 14-channel Emotiv headset, achieving approximately **92%** accuracy.

Professional Programs & Certifications

Google AI and Technology Academy 2025

- Selected for Google's 9-month AI and Technology Academy, focusing on Artificial Intelligence, software engineering, and technology entrepreneurship.

Projects

Turkish Coffee Fortune Telling AI Application Graduation Project

- Designed and developed an end-to-end AI-powered Android application for Turkish coffee fortune interpretation using a custom-trained YOLOv11 model deployed with TensorFlow Lite.
- Created and manually annotated a custom dataset of 8,049 images.
- Trained a YOLOv11n classification model achieving 99.5% accuracy.
- Integrated the TensorFlow Lite model into a Kotlin-based Android application for on-device inference.

EEG-Based Yes / No Detection Machine Learning

- Designed and developed an EEG-based intention recognition system using a 14-channel Emotiv headset.
- Performed signal preprocessing and feature engineering to build machine learning-ready datasets.
- Developed a Random Forest classifier achieving approximately 92% accuracy in distinguishing cognitive "Yes" and "No" responses.

Early Gender Detection Machine Learning

- Developed a Random Forest classifier for early fetal sex prediction, achieving 96.4% accuracy in internal evaluations.
- Performed biomedical signal preprocessing, feature extraction, normalization, and dataset preparation from raw fetal heart signal recordings.

Skills

Programming Languages: Python, TypeScript, Java, Kotlin, JavaScript

AI & Machine Learning: PyTorch, Scikit-learn, NumPy, OpenCV, YOLO, Random Forest, SVM, Computer Vision, Deep Learning, Feature Engineering, Signal Processing

LLM & NLP: RAG, LangChain, Prompt Engineering, Vector Databases

Backend & APIs: FastAPI, RESTful APIs , API Testing (Postman, Bruno)

Mobile Development: Android, Kotlin, TensorFlow Lite (TFLite)

DevOps & Tools: Docker, Kubernetes, Git, GitHub Actions , Linux

Data & Computer Vision: LiDAR, 2D/3D Annotation, Dataset Preparation, Data Augmentation

Automation: n8n