Cemre Dag

cemred.istanbul@gmail.com | linkedin.com/in/cemred | github.com/gumaruw

Education

B.Sc. in Software Engineering, Dogus University, TR — Oct 2021 – Jul 2025

Experience

Volunteer Data Analyst, Superint Software – Istanbul, TR — Jan 2025 – Apr 2025

- Led AI component development for a start-up project. Built personalized engine using Gemini API achieving 88% food recognition accuracy and 99.7% text-based accuracy
- Optimized query time with TF-IDF, reduced storage by 60%, and improved user retention
- Integrated USDA food database for accurate data with offline support using SQLite and cloud-based infrastructure

AI & NLP Intern, UNLU & Co – Istanbul, TR — Jan 2025 – Feb 2025

- Developed a RAG-based system, integrated local LLM inference (Ollama / Phi-4), preserving on-prem data privacy
- Applied NLP, IR, and transformer models to deliver a secure system with full regulatory compliance

SQL Development Intern, FLO Group – Istanbul, TR — Jul 2024 – Aug 2024

- Accelerated data analysis, reduced query times by ~30% across databases

Full Stack Developer Intern, Satel Informatics – Antalya, TR — Jul 2023 – Aug 2023

- Built real-time tourism news platform aggregating X (Twitter) for Antalya, Istanbul, Izmir, enabling timely coverage
- Implemented RESTful APIs and SQL pipelines for dynamic data ingestion, owning end-to-end integration

Projects

Portfolio

Vitalight – **rPPG Heart Rate Detection** — github.com/VitaLight

- Developed rPPG pipeline to estimate heart rate from facial video using OpenCV/MediaPipe, multi-ROI extraction and spectral methods, reducing estimation error to ~6.5% after iterative improvements
- Integrated CHROM, POS and ICA-based fusion with adaptive filtering and signal-quality scoring to stabilize estimates
- Fixed UBFC ground-truth parsing and benchmarked the pipeline on UBFC dataset to enable reliable evaluation

SifAI – **Skin Cancer Analysis** — github.com/Skin-Cancer

- Built U-Net segmentation model with 95.83% accuracy and 84.95% IoU on 10,000+ (HAM10000) images
- Trained ResNet50-based binary classifier with 89.55% accuracy and 93.88% AUC
- Conducted extensive model training with multiple architecture iterations, achieving 91.83% Dice coefficient

DocuMind – **Enterprise RAG System** — github.com/DocuMind

- Delivered RAG platform for financial reporting, cut manual research time by 65% with scalable architecture supporting 500–1,000 employees
- Developed retrieval layer with FAISS and ChromaDB using optimized embeddings, achieving sub-second query latency

SignFlow – **Real-Time ASL Recognition** — github.com/SignFlow

- Led development of the system, coordinating end-to-end project lifecycle
- Designed and trained a CNN for letter recognition that achieved 92%+ validation accuracy
- Preprocessed and normalized 1000+ images, corrected missing pixels via interpolation using OpenCV and scikit-image

Skills

Programming & Tools:	Python, SQL, Git, CI/CD, GCP, Power BI, Roboflow
Libraries & Models:	TensorFlow, PyTorch, Keras, Scikit-learn, NumPy, SciPy, MLflow, CNNs, Transformers, RAG
Computer Vision & NLP:	OpenCV, YOLO, MediaPipe, Signal Processing, NLP Pipelines
Languages:	Turkish (Native), English (C1), German (B1)