

Halil Erdogan

Senior Machine Learning Engineer | Software Engineer

✉ herdoganturkey@gmail.com | 💻 herdogan.dev | 🌐 github.com/halil-erdogan | 🔗 halil-erdogan | 📞 +352691131586

Technical Skills

- **Programming Languages:** Python, C++, SQL, Zig, Shell Scripting
- **GenAI Applications:** Conversational AI systems, LLM-powered content generation, RAG Systems, Voice Agents, Text Agents, Video Agents, Automated Copywriting, Semantic Search, AI-powered Analytics, Multi Agents
- **ML & GenAI Frameworks:** PyTorch, FastAPI, MLflow, scikit-learn, VAPI, Tavus, LangChain, OpenAI API, Langfuse, LlamaIndex, Transformers, Anthropic Claude API, Pinecone, ChromaDB, FAISS
- **Big Data:** Apache Airflow, Apache Spark, PostgreSQL, MongoDB, DuckDB, Redis, Pandas
- **DevOps & Cloud:** Docker, Kubernetes, AWS, GCP, Azure, Cloudflare, Terraform, Git, CI/CD
- **Monitoring & Analytics:** Grafana, Prometheus, ELK Stack, PowerBI, Looker, Tableau

Professional Experience

| | |
|---|-------------------------------|
| INFINIT | February 2025 - Present |
| Senior Machine Learning Engineer | Luxembourg |
| <ul style="list-style-type: none">• Developed conversational AI voice agent using VAPI and OpenAI that reduced human dependency by 70% for routine inquiries while maintaining 95% customer satisfaction• Architected ML-based vehicle evaluation system reducing inventory review time by 50%, enabling efficient prioritization of high-value listings• Engineered LLM-powered advertisement generation platform using OpenAI GPT models, reducing content creation time by 85% and increasing engagement by 40%• Implemented Azure-based MLOps pipeline with Langfuse observability, achieving 99.8% model availability and enabling rapid iteration cycles | |
| Aiva Technologies | January 2019 - September 2024 |
| Senior Machine Learning Engineer | Luxembourg & Remote |
| <ul style="list-style-type: none">• Progressed from intern to technical team lead, driving growth to \$1M ARR and securing investment at \$30M+ valuation• Architected high-performance ML pipeline processing 1M+ monthly compositions with 99.9% reliability and 90% reduced latency• Led migration from traditional ML to transformer-based ML resulting in 300% YoY user growth and scaling to 50,000+ MAU• Built real-time music generation APIs achieving 50ms response time and designed fault-tolerant microservices architecture using Docker/Kubernetes | |
| Fraunhofer IAIS | June 2018 - September 2018 |
| Research Intern | Sankt Augustin, Germany |
| <ul style="list-style-type: none">• Engineered first-of-its-kind speech segmentation system reducing manual annotation time from hours to minutes, leading to widespread adoption across organizational projects | |
| Telecom ParisTech | June 2017 - September 2017 |
| Research Intern | Paris, France |
| <ul style="list-style-type: none">• Developed novel probabilistic model for speech enhancement, published in ICASSP 2018 | |

Education

| | |
|--|--------------------------------|
| Friedrich-Alexander-Universitat Erlangen-Nurnberg | September 2019 - December 2021 |
| MSc. in Communication and Multimedia Engineering | Erlangen, Germany |
| <ul style="list-style-type: none">• GPA: 1.9/1.0 (German grading system, equivalent to distinction)• Thesis: 📄 <u>Data-Driven Singing Activity Detection for Music Recordings</u> | |
| Sabanci University | September 2014 - January 2019 |
| BSc. in Electronics Engineering | Istanbul, Türkiye |
| <ul style="list-style-type: none">• GPA: 3.78/4.00 | |