

# M. Erdi ARI

+90 (507) 889 46 81 | Şişli, İstanbul | [me@erdiari.dev](mailto:me@erdiari.dev) | [github.com/erdiari](https://github.com/erdiari) | [linkedin.com/in/erdiari](https://linkedin.com/in/erdiari) | [hf.co/erdiari](https://hf.co/erdiari)

## EDUCATION

**İstanbul Technical University**

Maslak, İstanbul

*Physics Engineering*

- Relevant Coursework: Python, Fortran, Statistical Learning From Data and General Mathematics

## WORK EXPERIENCE

**AI engineer**

Oct 2022 — Present

VNGRS

*Remote*

### Notable Projects:

- LLM Training: Trained [Kumru AI](#) and VBART, from scratch. (Data collection, filtering, pre-training, post-training)
- Real-Time Transcription: Developed systems for converting live speech to text instantly, enabling immediate text output during conversations or presentations
- Text Categorization: Created automated classification systems to organize textual content into relevant categories, streamlining information management
- Graph-based RAG: Implemented a Retrieval-Augmented Generation system using graph structures to enhance context understanding and information retrieval
- Video Generation: Engineered AI-driven solutions for automated video content production
- Specialized Data Curation: Assembled and refined datasets specifically tailored for model fine-tuning, improving AI model performance for targeted use cases

## SKILLS

- **Python:** PyTorch, TensorFlow, Keras, scikit-learn, NumPy, Pandas, Matplotlib, Seaborn, Plotly, JAX, Langchain, Hugging Face
- **Machine Learning, AI:** Machine Learning, Deep Learning, Neural Networks, Natural Language Processing (NLP), Transformers, Attention Mechanisms, Transfer Learning, Reinforcement Learning, Model Lifecycle, CNN, RNN, LSTM, Generative AI, Data Science, GAN, Hyperparameter Tuning, Algorithm Design, Big Data
- **Proficiency in UNIX/LINUX Environment:** Ubuntu, NixOS, CentOS, Bash Scripting, Shell Commands, Package Management, Anaconda, Virtual Environments, System Administration, Performance Monitoring
- **Cloud Services (AWS):** AWS EC2, S3, SageMaker, IAM, Docker, Containerization, Kubernetes, APIs, Cloud Computing
- **Development Cycles:** Git, GitHub/GitLab, CI/CD Pipelines, Unit Testing, DevOps, Data Engineering, Documentation

## EXTRA

### Publications

**VBART: The Turkish LLM,** (<https://arxiv.org/abs/2403.01308>)

**VNLP: Turkish NLP Package,** (<https://arxiv.org/abs/2403.01309>)

**Ahşapta İç Kusurların Tespiti...,** ([www.researchgate.net/...](http://www.researchgate.net/...))

### Certificates

**Toelf C1 (110/120), ETS**

11 February 2026

**Getting Started with Deep Learning, NVIDIA**

29 June 2022 [Verify](#)

**Data Science Certificate Program, Ginova**

June 2020