

ENES KUZUCU

+905064024980 | enesesvetkuzucu@gmail.com | [Github](#) | [LinkedIn](#)

SKILLS

- **Programming:** Python | C | JAVA | MATLAB | SQL | AMPL
- **Software & Tools:**
 - LLM:** Langchain, Agno, LangGraph, LlamaIndex, HuggingFace, OpenAI
 - ML:** PyTorch, Tensorflow 2, Scikit-learn, BigQuery ML, Prophet
 - Computer Vision & Image Processing:** OpenCV, StableDiffusion, ComfyUI, torchvision, PyTorch Model Zoo,
 - Data Processing & Databases:** Pandas, PySpark, Chroma, Airflow & BigQuery, MongoDB, PostgreSQL, MySQL
 - Other:** Git, Flask, FastAPI, Django, Docker, Kubernetes, AWS, GCP
- **Languages:** English (TOEFL IBT 90) | Turkish | Russian
- PyPI packages: [LLMService](#) [proteas](#) [facexformer-pipeline](#) [categorizer](#) [serpengine](#) [brightdata](#)

EXPERIENCE

AI Solutions Engineer *Affirm, USA (Remote)*

JUN 2025 – (ONGOING)

- **Responsible for the analysis of internal workflows and designed custom integrations and automations to significantly improve operational efficiency across teams.**
- Designed and deployed a custom Slack bot that answers IT-related questions using the company's Notion knowledge base and agentic reasoning, resulting in a **45% reduction in human-managed IT tickets within two months.**
- Implemented a production-grade Slack bot architecture with comprehensive guardrails, deployment tracking, and log-driven self-improvement mechanisms.
- Built an internal LLM-powered analytics tool that translates natural language into SQL, executes queries against a **100-table test database**, and generates visualizations for rapid data exploration (internal use only).
- Developed **custom agentic workflow** to automatically **generate AI-powered sales slide decks**, currently in late prototype phase, enabling the **sales team** to rapidly create **high-quality presentations** with **minimal manual effort**.

Open-Source Maintainer BrightData

JAN 2025 – DEC 2025

- Invited by Bright Data to officially maintain the Python SDK after developing an independent implementation; contributed to and advised on a major SDK rewrite, with ongoing open-source maintenance and reviews.

Data Scientist *EETech, USA (Remote)*

FEB 2021 – JUN 2025

- **LLM Pipeline for Data Imputation and Normalization:** Built a production FastAPI service combining web scraping, SERP data, and LLM-based extraction to recover missing product attributes; implemented custom async orchestration and context-aware extraction. Adopted by **two electronic component distributors**.
- **LLM-based Document Processor (RAG):** Designed and evaluated a RAG pipeline for technical PDFs using Chroma, custom chunking strategies, and page-aware context injection; benchmarked multiple LLMs (cloud, local, vLLMs) with labeled datasets, improving similarity scores from **55% to 68%** and significantly reducing manual extraction time..
- **Demand Forecasting and Custom Data Quality Metrics for Time-Signals:** Developed a multi-model demand forecasting application, achieving 93% accuracy for increase/decrease predictions based on 3 months of cross-validation results. Created custom data quality metrics, such as data integrity score, time series cohesion score, and predictability score, to segment datasets based on their forecastability.
- **Data Cleaning Strategy for Electronic Component Dataset:** Implemented a robust data cleaning strategy for a crawled electronic component dataset consisting of 120 million rows. Developed multiple custom tailored filters and created a pipeline mechanism using the Prefect package to process them. Successfully eliminating 98% of noise while retaining 96% data coverage. Used Pandas and Spark extensively
- **Team Growth in Turkey:** Responsible for expanding our developer team in Turkey, I collaborated with three recruiter agencies and conducted interviews for various positions, including technical data science interviews. Successfully grew the team size from 2 to 9 members.

Founder / Lead Backend Engineer *budgety.ai*

JULY 2024 - MAY 2025

- Founded and built an LLM-driven personal finance platform that extracts and categorizes transactions from bank statement PDFs.
- Designed a multi-stage document processing pipeline combining PDF parsing, OCR, and AWS services to handle highly heterogeneous bank statement formats.

- Implemented robust table normalization logic to standardize dates, currencies, and transaction schemas across multiple banks.
- Built a configurable LLM-based categorization engine using YAML-defined taxonomies; optimized throughput via async batching and keyword/pattern-based fast paths.
- Developed and deployed a production-grade FastAPI backend with SQLAlchemy, JWT authentication, background workers, and CI/CD on AWS ECS.

EDUCATION & PUBLICATIONS

Master's of Science in Big Data & AI 08/2020 - 07/2022
Novosibirsk State University, Russia

Bachelor of Biomedical Engineering 09/2014 - 06/2018
Kocaeli University, Turkey

“**Button Press Detection from EEG Signals Using Deep Learning**”, 2022 IEEE 23rd International Conference of Young Professionals in Electron Devices and Materials (EDM), <https://doi.org/10.1109/EDM55285.2022.985519>