

Fatma Betül ARSLAN Data Scientist

 betularsln01@gmail.com

 +90 546 911 1024

 İstanbul, Turkey

 linkedin.com/in/fatma-betül-arslan

 <https://github.com/fatmabetularslan>

About Me

Data Scientist with a BSc in Mathematics and hands-on experience in Python and SQL. Worked on ML model development and data pipelines using tools such as scikit-learn and Airflow, with applied experience in image recognition, churn prediction and time-series forecasting.

Skills

Python: Pandas, NumPy, scikit-learn, XGBoost

Machine Learning: Classification, churn, time-series, imbalanced data

Tools: Apache Airflow, Docker, Linux, REST API integration

Data Processing: SQL (PostgreSQL, Oracle), preprocessing, feature engineering, quality checks

Visualization: Matplotlib, Seaborn, Power BI, Plotly

Professional Experience

Data Science Intern, Ordulu Technology

- Built image-based product recognition models in Python. (classification & detection) Achieved ~80%+ accuracy.
- Designed ML pipelines with Airflow, improving automation & runtime.
- Collaborated on integrating models into production APIs.

Oct 2023 – May 2024

Database Administrator- Intern, DMC Information Technologies

- Wrote SQL queries & automated reporting dashboards.
- Optimized database performance (indexes, joins).
- Ensured data accuracy across large relational datasets.

Jul 2023 – Sep 2023

Education

Bachelor's Degree - Mathematics, Ankara Yıldırım Beyazıt University

Sep 2018 – Jul 2024

Projects

FinTurk Finansal Asistan, Streamlit, Gemini AI, XGBoost, RAG

Price forecasting, news sentiment & portfolio analysis; deployed an interactive finance assistant with real-time BIST data.

Aug 2025

AI Portfolio Chatbot, TR/EN, Gemini API, RAG

Developed an interactive AI portfolio assistant that answers questions about my CV, projects and experience using a RAG-enhanced LLM pipeline. Built with Streamlit UI, FastAPI backend and FAISS-based vector search for real-time context-aware responses.

Jul 2025

Customer Churn Prediction, Apache Spark, SMOTE, Imbalanced Data Handling

Built a customer churn prediction pipeline on large-scale datasets using Apache Spark and improved F1 score through feature engineering and imbalanced-data techniques (e.g., SMOTE).

Dec 2024

Smart Home Energy Management Application, REST API, Pytest, Docker

Time-series ML model deployed via REST API; achieved ~8% MAPE and built unit & endpoint tests with Pytest.

Dec 2024

Competitions

BTK Datathon 2024

- 2nd place out of 364 teams.

Gdz Electric Datathon 2024

- Ranked in the top 7% with a time-series forecasting model.

Certifications & Courses

Machine Learning Specialization, Stanford Online

Big Data Camp, Türk Telekom

AWS Cloud Solutions Architect Professional Certificate, AWS

Languages

Turkish — Native • **English** — Upper-intermediate