EGE KAAN OZALP

Mail: ozalp.egekaan@gmail.com
Location: Istanbul/Turkey
Date of Birth: 27/03/2001
LinkedIn: https://www.linkedin.com/in/egekaanozalp/
GitHub: https://www.github.com/egekaanozalp/

EDUCATION

The University of Edinburgh | Edinburgh, Scotland ***September 2026 – August 2027***
MSc Data Science

Sabanci University | Istanbul, Turkey ***29 August 2020 – 24 June 2025***
BSc Computer Science and Engineering & Industrial Engineering (Double Major)

- **GPA:** 3.52/4.00 (Computer Science and Engineering), 3.59/4.00 (Industrial Engineering)
- **Honors:** 5 terms of High Honor Certificate for finishing the terms with a GPA higher than 3.5 and 3 terms of Honor Certificate for finishing the terms with a GPA higher than 3
- **Awards:** 50% tuition waiver for 5 years for success in the nationwide university placement exam
- **Key Modules:** Data Science, Machine Learning, Deep Learning, Natural Language Processing, Operations Research (LP, IP, MIP), Object Oriented Programming, Data Structures & Algorithms

Huseyin Avni Sozen Anatolian High School | Istanbul, Turkey **16** ***August 2015 – 19 June 2020***
Graduated with 96.78/100 cumulative grade point average.

- **Honors:** 5 High Honor Certificates for each year of education
- Selected to the Physics Olympiad team and participated to the Physics Olympiad exams of TUBITAK and Capa Science High School (2019).

SKILLS

- **Languages:** Turkish (Native), English (C1)
- **Programming Languages:** Python (Advanced), C (Advanced), C++ (Advanced), Assembly (Intermediate), SQL (Intermediate), HTML/CSS (Intermediate), Java (Beginner), C# (Beginner)
- **Frameworks:** Pandas, Scikit-Learn, PyTorch, Darts, LangChain, Langfuse, DeepEval, Optuna
- **Software:** MySQL, SQL Server Management Studio, SQLite, Microsoft Office Applications.

EXPERIENCE

ICRON ***August 2025 – Present***
Full-Time Data Scientist

- Developed and shipped AI-driven solutions and products using machine learning, deep learning, natural language processing, reinforcement learning, and multi-agent orchestration techniques for projects including the ICRON Multi-Agent LLM Platform, RAG-powered RFI Question Answering Agent, and an RL-based bid price management system developed for ACRM operations at THY Cargo.

Turkish Airlines Technology ***November 2024 – August 2025***
Part-Time Data Scientist

- Developed multivariate time-series forecasting models using machine learning and deep learning techniques and utilised natural language processing algorithms across different projects including Customer Lifetime Value (CLV) Prediction, and Prediction of Overseas Operational Expenses.

L'Oréal Turkey ***July 2024 – November 2024***
Information Technology Intern

- Created a program that converts the complaints of the customers from speech to text and automatically classifies them using natural language processing algorithms including speech recognition, sentence-transformers, and cosine similarity function.
- Created new promotions using Magento based on the given digital assets and cart price rules.
- Performed frontend and promotion logic tests of the websites kiehls.com.tr and lancome.com.tr.
- Analysed and solved the purchasing related issues by using Salsify and Magento.

Kariyer.Net ***August 2023 – December 2023***

Machine Learning Intern

- **Internship Project: Skill Extraction and Recommendation System**
    - Contributed to the creation of a system that automatically identifies and extracts soft skills and hard skills from both the job postings and the CVs using natural language processing.

Adel Kalemcilik ***March 2023 – June 2023***

Information Technology Intern

- Managed the data entry of the company using SAP Logon.

PROJECTS

Portfolio Website | Personal ***May 2026 – June 2026***

Programmer

- Built a full-stack portfolio site with Django 6 and a custom glassmorphism-based dark UI, featuring canvas-based neural network and particle animations crafted in vanilla JS.
- Designed a content-managed backend with 10+ Django models (projects, skills, experience, certificates) administered through a fully custom-themed admin panel, with a contact form that verifies every submission using Cloudflare's Turnstile before sending to avoid bot attacks.
- Hardened admin access with django-axes brute-force lockout (5-attempt IP ban) on top of Django's built-in CSRF and authentication layers.
- Deployed on Render with PostgreSQL, WhiteNoise static serving, and Gunicorn — with a one-command build pipeline via render.yaml.

RAG Agent for RFI Question Answering | ICRON ***September 2025 – February 2026***

Programmer

- Designed and implemented a custom document loader pipeline for unstructured documents using LangChain and Unstructured libraries to enable automated extraction and preprocessing of textual content from heterogeneous document formats.
- Generated high-dimensional vector embeddings for processed document chunks and integrated them with the Qdrant vector database for efficient semantic similarity search and retrieval operations.
- Applied semantic, neural, and recursive chunking strategies using LangChain and Chonkie to optimize the Contextual Relevancy metric (DeepEval) and downstream retrieval performance.
- Containerized and deployed the RFI RAG agent using Docker, ensuring portability, reproducibility, and scalable integration of retrieval, embedding, and inference components.

ICRON Multi-Agent LLM Service | ICRON ***September 2025 – February 2026***

Programmer

- Developed an AI-driven forecasting explainer agent that interprets the outputs of the multivariate time-series forecasting models by decomposing the forecast into trend, seasonality, and residuals, identifying changepoints, utilising dynamic time warping for finding the SKUs with similar temporal patterns and applying a wide range of statistical analyses to generate natural language explanations to user questions in Python using LangChain and RabbitMQ.
- Implemented the ToolSelector class to stabilize the tool selection step of the LLM agents using LangChain and ToolRegistry frameworks.
- Implemented the prompt optimization and LLM evaluation modules using DSPy and DeepEval.

Customer Lifetime Value Forecasting | Turkish Technology ***November 2024 – August 2025***

Programmer

- Developed multiple scalable CLV forecasting models for more than 30 million customers over 6-, 12-, 24-, and 36-month horizons using tree-based ensemble learning models for regression and sliding window strategy in Python.
- Performed exploratory data analysis using Matplotlib, Seaborn, and Fitter libraries. Detected outliers using IQR and handled them using Winsorization method.

- Resampled around 70 million customer records to around 30 million with an additional churn model and undersampling methods.
- Executed feature subset selection using Tree Importance, RReliefF and filter methods. Applied feature engineering by creating RFM features using the windows.
- Conducted customer segmentation using K-Means Clustering, and developed segment-based classification and regression models using LightGBM.
- Tuned the model hyperparameters using RandomizedSearchCV and Optuna.
- Delivered multiple CLV models with robust MAEs and MAPEs – e.g., the 6-month model deviated less than 10% of the actual value for 85% of the customers in the test set.

Classification of Customer Complaints | L'Oréal Turkey ***October 2024 – November 2024***
Programmer
- Developed a Python program that transcribes customer complaint voice recordings and classifies them automatically using NLP techniques, including speech recognition, SBERT sentence transformers, and cosine similarity.

Detecting the Severity of Bugs | Sabanci University ***July 2024***
Programmer
- Created a pipeline that identifies the severity of a bug, given its description text, using TF-IDF Vectorizer and a Random Forest Classifier with differing class weights.

Polling Unit Optimization for Local Elections | Sabanci University ***October 2023 – May 2024***
Programmer
- Created a mixed-integer optimization model that accurately finds the user-defined number of neighbourhoods in each city of any country such that analysing solely the results of polls conducted in these selected neighbourhoods would provide a highly accurate overview for the outcomes of the future elections in this city.
- The project consisted of 4 phases: Collecting the data of the past elections from YSK using Selenium, formatting the data using Pandas, creating a mixed-integer program using Gurobi and testing and proving the high accuracy of the model using Python.

Skill Extraction and Recommendation System | Kariyer.Net ***August 2023 – December 2023***
Programmer
- Contributed to the creation of a system that automatically identifies and extracts soft skills and hard skills from both the job postings and the CVs using natural language processing concepts such as creating a custom NER model and a few-shot NER model using a Large Language Model, training an SVM classifier, topic modelling, finding n-grams, creating word embeddings, calculating cosine similarities, and fuzzy matching.
- Developed a Python program that identifies the skills associated with job titles and calculates the frequency of these skills within the respective job title for the recommendation system.

DiSUcord | Sabanci University ***December 2023***
Programmer
- Developed a message networking application by implementing the server and client modules in C# using the socket programming techniques.

SUForum | Sabanci University ***March 2023 – June 2023***
Programmer
- Created a social media application which consists of a back-end application coded in Java Spring Framework using MongoDB as database, and a front-end application coded in Android framework for the course CS310: Mobile Application Development.

University Management System | Sabanci University ***October 2022 – January 2023***
Programmer
- Designed the ER Model of the system and created the website of it using MySQL, PHP, HTML, CSS, and Firebase.

Bitcoin Price Prediction | Sabanci University ***March 2022 – June 2022***
Programmer

- Created a program in Python which performs machine learning algorithms and hypothesis testing on the data taken from Kaggle.com to predict the opening price of Bitcoin for the upcoming day.

EXTRACURRICULAR ACTIVITIES

kAi Sabanci | Sabanci University ***November 2022 – September 2024***
Founding Board Member

- Held meetings with NVIDIA, which sponsors the club.
- Managed organising the workshops and selecting the term projects that are about machine learning, deep learning & artificial intelligence concepts and developed these projects with the club members.

CERTIFICATES

Intermediate Git | DataCamp **September 2025**
Participant & Programmer – 2-Hour program

Introduction to Git | DataCamp **August 2025**
Participant & Programmer – 2-Hour program

Building Conversational AI Applications | NVIDIA Deep Learning Institute **December 2023**
Participant & Programmer – 9-Hour program

Fundamentals of Deep Learning | NVIDIA Deep Learning Institute **December 2023**
Participant & Programmer – 9-Hour program

Introduction to Data Science in Python | University of Michigan **August 2023**
Participant & Programmer – 29-Hour program

Advanced Consultancy Training Certificate | Bogazici University **October 2022**
Participant – 15-Hour program

- Qualified for the case study that was conducted by the company KPMG and selected as the 3rd best group to present the case solution with the team that I was attended to.

VOLUNTEERING

TEMA Foundation | Istanbul, Turkey ***2016 – Present***
Volunteer