Halil İbrahim Ergül

Türkiye | +905522390268 | halilibrahimergul55@gmail.com | LinkedIn | Github

WORK EXPERIENCE

Hepsiburada (NASDAQ: HEPS)   **Turkey, Istanbul**
Data Scientist II, Search Algorithm Team   *May 2024 - Present*
- Engineered robust data pipelines for high-quality collection, processing, and cleaning.
- Developed N-gram and topic extraction-based search suggestions, which drove a 10% increase in click-through and 7% higher order conversions.
- Improved search relevance with point-wise ranking experiments, driving an 11% and 2% gain in offline and online metrics
- Debugged and fine-tuned in-house NER model training, leading to 6% greater entity recognition accuracy.
- Established systematic evaluation for typo correction models.
- Designed and launched "Related Search Term" feature using BERT and recommendation algorithms

Upily   **Turkey, Istanbul**
Data Scientist   *2022 - 2024*
- Built automated and scheduled ETL pipelines for processing millions of daily events via BigQuery and creating dashboards
- Ad-hoc data analysis, applying end-to-end ML models for churn prediction to improve conversion rates
- Clustering users and increasing user segmentation accuracy by 22%
- Collaborated with Product, Marketing, and Growth teams on ad-hoc A/B test results analysis and custom KPI reporting for 5+ monthly experiments

ML Engineer   *2022 - 2023*
- Web Scraping for preparing training data (via Scrapy/Selenium/Bs4)
- Prepared and processed 500+ hours of audio data for TTS model training, achieving 80%+ Mean Opinion Scores from domain experts
- Integrating an LLM assistant into backend for user interaction and optimizing API response times by 30%

Behavioral Analytics and Visualization Lab   **Turkey, Istanbul**
Research Assistant   *2023 - 2024*
- LLM fine-tuning for various sub-domain tasks, from summarizing a custom dataset to behavior classification
- Worked with spatial-behavioural data to implement a recommendation algorithm for merchant POIs
- Worked on an NLP deep learning project for Turkish sentiment tone detection with BERT models, supported by TUBITAK

EDUCATION

Sabanci University   **Turkey, Istanbul**
M.Sc with Thesis, Data Science, GPA: 3.60/4.00   *Graduation Date: 2024*
Thesis Title: Instruction-based fine-tuning of open-source LLMs for predicting customer purchase behaviors

Boğaziçi University   **Turkey, Istanbul**
BA, *International Relations & Sociology,* **GPA: High Honors, 3.94/4.00 (1st among cohorts)** *Graduation Date: 2021*

SKILLS

Tech: Python, SQL, R Studio, PySpark, Google Cloud Platform, Google BigQuery, Elastic Search, Redis, PyTorch, MLflow, Airflow, Tableau, Looker, SPSS, Stata, GA4, Docker

PROJECTS

- Question-Answer Application: End-to-end Question-Answer App with LLamaindex and ElasticSearch
- NLP-Powered Recommendation System For Hotels: Retrieving relevant hotels and ranking

PUBLICATIONS

- Instruction-Based Fine-tuning of Open-Source LLMs for Predicting Customer Purchase Behaviors, arXiv
- LDA Topic Modeling on Twitter Data Concerning Immigrants and Refugees, IEEE

CERTIFICATES

- Advanced Learning Algorithms, AWS Cloud, ML in Production, Big Data & Business Analytics, NN and Deep Learning