Gökmen Öz

Contact: (+49) 160 924 08 473 ⋄ gokmenoz89@gmail.com

WORK EXPERIENCE

Applied Scientist II   Feb 2020 - *present*
Amazon, Artificial General Intelligence - AGI, (Alexa AI)   *Berlin, Germany*

· Role: Science Lead of large scale projects that include data infrastructure, LLM inference, optimization, and deployment. Leading and implementing workflow and system designs. Building communication with external teams and bridges with their services.
· Topics: Large Language Models (LLM Training and Evaluation), Generative Artificial Intelligence (AI), Transformer-based Natural Language Understanding (NLU) models, End-to-end (E2E) pipeline setup, Web UI development (Front-end, Back-end), Synthetic Data Generation, Dialog Evaluation and Online & Offline Prediction of Dialog Defects, A/B testing, Amazon Web Services (AWS), e.g. Bedrock, SageMaker, EC2, S3, Athena, StepFunctions, EventBridge.
· (Some) Achievements: 1) Migrated several locales (arSA, itIT) to first DNN and then Transformer-based models and achieved more than 10% accuracy improvement. 2) Built E2E NLU infrastructure for koKR, trTR, plPL and launched to Beta. 3) Built a key algorithm for multilingual data generation that was adopted globally for Alexa's NLU system.
· Coding: Production code for several internal tools and repositories (development and unit testing).
· Started as Research Scientist I in 2020, promoted to Research Scientist II in 2022, and then to Applied Scientist II in 2024.

Ph.D. Research & Teaching Assistant   Sep 2014 - Jan 2020
The Pennsylvania State University Economics Department   *Pennsylvania, USA*

· Dissertation: Studied Applied Game Theory, Bayesian Inference, Dynamic Reputation Games.
· Research Assistant: The lead of an applied mechanism design project. Built a tool for web scraping, data cleaning and engineering, statistical inference with a structural model.
· Instructor: Taught my own course in Introductory Microeconomics Analysis and Policy
· Teaching Assistant: Intermediate Macroeconomics Analysis, Economics of Public Expenditure, International Finance and Open Economy Microeconomics.

EDUCATION

The Pennsylvania State University   Sep 2014 - Jan 2020
Ph.D. in Economics   *Pennsylvania, USA*

· Applied Game Theory, Bayesian Inference, Structural Models, Statistical Inference, Econometrics, Auction Theory, Dynamic Reputation Games

The Data Incubator   March 2019 - May 2019
3 months Fellowship   *New York City, USA*

· Won the Fellowship award with less than 2% selection rate.
· Selected for the unique Capstone project for a company.
· Intense Data Science training with hands-on training for all Machine Learning Algorithms from Regression Analysis, Conditional Random Fields, K-means Clustering, Regression Analysis, Daily coding exercises in Python, SQL, Bash.

Sabanci University	August 2012 - June 2014
M.A. in Economics	*Istanbul, Turkey*

· M.A. Thesis: "Parking Fees and Retail Prices" (joint with C. Robin Lindsey and Eren Inci) published by **Journal of Transport Economics and Policy**, July 2018.

Bogazici University	August 2006 - June 2012
B.S. in Industrial Engineering	*Istanbul, Turkey*

· Time Series Analysis, Forecasting, Auto Correlation models (AR,ARMA, ARIMA, SARIMA), Operations Research, Statistics.

REFERENCES

Bei Chen
Research Science Manager, Amazon Alexa AI
Contact: chenbe@amazon.de
Karolina Owczarzak
Research Science Manager, Amazon Alexa AI
Contact: karowc@amazon.com
Felipe Blanes
Technical Program Manager, Amazon Alexa AI
Contact: blanesf@amazon.com
Charith Peris
Research Scientist, Amazon Alexa AI
Contact: perisc@amazon.com
Kalyan Chatterjee
Distinguished Professor of Economics and Management Science, Penn State University
Contact: kchatterjee@psu.edu
Vijay Krishna
Distinguished Professor of Economics, Penn State University
Contact: vkrishna@psu.edu
Henrique Roscoe de Oliveira
Professor of Economics, Penn State University
Contact: hzd1@psu.edu

PUBLICATION AND RESEARCH

"Towards Equitable Natural Language Understanding Systems for Dialectal Cohorts: Debiasing Training Data" published by **LREC-COLING 2024**, `https://aclanthology.org/2024.lrec-main.1433.pdf`

"CALICO: Conversational agent localization via synthetic data generation" accepted to **NeurIPS 2023 Workshop on SyntheticData4ML**, `https://openreview.net/forum?id=bpO14yuQQg`

"Alexa Teacher Model: Pretraining and Distilling Multi-Billion-Parameter Encoders for Natural Language Understanding Systems" published by **KDD 2022**, `https://arxiv.org/abs/2206.07808`

"Knowledge distillation transfer sets and their impact on downstream NLU tasks", published by **EMNLP 2022**, `https://www.amazon.science/publications/knowledge-distillation-transfer-sets-and-their-impa`

"Using multiple ASR hypotheses to boost i18n NLU performance" published by **ICON 2020**, `https://arxiv.org/abs/2012.04099`

Ph.D. Thesis and Job Market Paper: "Sustainable Reputations with Biased Review Platforms" (Microeconomics Theory, Game Theory, Dynamic Reputation, Industrial Organization, Automated Web Scraping, Regression Analysis)

M.A. Thesis: "Parking Fees and Retail Prices" (joint with C. Robin Lindsey and Eren Inci) published by **Journal of Transport Economics and Policy**, July 2018.

SKILLS

Machine Learning	LLM, Transformers, Neural Networks, Natural Language Processing, Clustering, K-Nearest, Random Forest, Linear Regression, Time Series Analysis
Economics	Microeconomics Theory, Game Theory, Auction Theory, Regression Analysis, Linear Algebra, Hypothesis testing
Software/Tools	Python (Pandas, Numpy, Tensorflow, Keras), Matlab, SQL, Unix/Bash, Git Data Visualization (Bokeh, Geopandas, Matplotlib), Spark, HTML, CSS, AWS

ACADEMIC ACHIEVEMENTS

Ranked in Top 0.015% (303/2,000,000 candidates) in Turkey National University Entrance Exam

Top 0.004% (13/350,000 candidates) in Turkey National Grad School Entrance Exam

Being selected to compete at the high school level National Math Olympics.