Kerem Zaman

keremzaman@gmail.com | +1 919 260-7512 | Github: KeremZaman | website: keremzaman.com

Education

The University of North Carolina at Chapel Hill — Chapel Hill, NC, USA
PhD in Computer Science (2022 - Present)
Advisor: *Prof. Shashank Srivastava*

Boğaziçi University — Istanbul, Turkey
BSc in Computer Engineering (2018 - 2022)

Research Interests

Interpretable and explainable machine learning, AI safety, LLMs, multilingual NLP, fairness, synthetic data generation

Publications

Optimization-Free Image Immunization Against Diffusion-Based Editing
Tarik Can Ozden, Ozgur Kara, Oguzhan Akcin, Kerem Zaman, Shashank Srivastava, Sandeep P. Chinchali, James M. Rehg
Under Review

A Causal Lens for Evaluating Faithfulness Metrics
Kerem Zaman, Shashank Srivastava
EMNLP 2025

INTERACT: Enabling Interactive, Question-Driven Learning in Large Language Models
Aum Kendapadi*, Kerem Zaman*, Rakesh R Menon*, Shashank Srivastava
ACL 2025 **[Oral Presentation]**

Fuse to Forget: Bias Reduction and Selective Memorization through Model Fusion
Kerem Zaman, Leshem Choshen, Shashank Srivastava
EMNLP 2024

MaNtLE: Model-agnostic Natural Language Explainer
Rakesh R Menon, Kerem Zaman, Shashank Srivastava
EMNLP 2023

A Multilingual Perspective Towards the Evaluation of Attribution Methods
Kerem Zaman, Yonatan Belinkov
EMNLP 2022 **[Oral Presentation]**

Rank in Style: A Ranking-based Approach to Find Interpretable Directions in StyleGAN
Umut Kocasarı*, Kerem Zaman*, Mert Tiftikci*, Pinar Yanardag
CVPR 2022 5th Workshop on Computer Vision for Fashion, Art, and Design

BURST: Software and Simulation Solutions of an Autonomous Vehicle
Fatih Köse, İbrahim Özcan, Melih Dal, Metehan Yıldırım, , Kadir Türksoy, Kerem Zaman, Tuna Meral, Sinan Öncü
2020 28th Signal Processing and Communications Applications Conference (SIU), Gaziantep, Turkey, pp. 1-4

Experience

UNC-NLP — Chapel Hill, NC, USA
Graduate Research Assistant | Advisor: Shashank Srivastava — *August 2022 - Present*
Working on explainable NLP, faithfulness of natural language explanations, safety, fairness and model merging

IFM MBZUAI Silicon Valley Lab — Sunnyvale, CA, USA
AI Engineering Intern| Advisor: Mikhail Yurochkin — *May 2025 - August 2025*
Worked on improving domain-specific reasoning abilities of LLMs by generating synthetic mid-training data through self-play

CATLAB at Bogazici University — Istanbul, Turkey
Undergraduate Research Assistant | Advisor: Pinar Yanardag — *October 2021 - June 2022*
Worked on improvement of text-based image manipulations for GANs by utilizing automatically extracted masks and improved text generation strategies

Independent Research — Remote
Advisor: Yonatan Belinkov — *April 2021 - June 2022*
Worked on evaluating wide variety of attribution methods on the NLI task along with introducing a novel method for faithfulness evaluation

Independent Research — Remote
Advisor: Tolga Birdal — *September 2021 - September 2022*
Worked on Riemannian replacement of GANs and the concepts like interpolation, random walk and latent manipulation

PragmaCraft — Istanbul, Turkey
Software Engineer / Software Engineer Intern — *July 2018 - June 2022*
Worked on diverse NLP tasks, including retrieval, QA, NER, summarization, and text classification, covering the full lifecycle from training models from scratch to deployment and optimization.

Professional Service

- Reviewer at ICLR 2025, ACL Rolling Review (ARR) 2024-2025, DaSH Workshop @ EMNLP 2022

Skills

- **Programming Languages**: Python, C/C++
- **Technologies**: Git, ROS, Gazebo, MongoDB, Docker
- **Libraries**: PyTorch, numpy, scikit-learn, OpenCV, Transformers, vLLM
- **Familiar with**: HTML, CSS, Javascript, SQL, C#, Verilog, LaTeX, AWS
- **Human Languages**: Turkish (native), English (advanced), Laz (elementary), Ancient Egyptian (elementary), German (elementary)

Other

- **Bogazici University Robotic Systems Team Autonomous EV project (2018 - 2022)** : Led Autonomous Software Team between 2020-2022, developed path planning and control systems, took part in traffic sign detection module, realized and tested them in simulation environment and on the real vehicle.

Additional Experience & Achievements

- Most Original Software Award for Unique Vehicle Category at Teknofest Robotaxi-Full Scale Autonomous Vehicle Competition 2021
- 1st place in Preliminary Design and Simulation Stage at Teknofest Robotaxi-Full Scale Autonomous Vehicle Competition 2019 & 2020
- 242nd place among 2M+ students in University Entrance Exam (2018)
- 3rd place in Istanbul-Asia Region in TUBITAK (Scientific and Technological Research Council of Turkey) Research Projects Challenge for High School Students
- Gave introductory programming courses in high school organizations