Ömer Veysel Çağatan

https://asparius.github.io
omerveyselacademic@gmail.com

Education

Koç University, Istanbul — Sep 2020–Jun 2024
BSc, Computer Engineering
Tracks: Artificial Intelligence, Data Analytics

Professional Experience

Research Engineer, KUIS AI — Apr 2025–Present
Working on large-scale preference learning, RLHF, and LLM fine-tuning research.

NLP Intern, FineSci Technology supervised by Assoc. Prof. Alptekin Kupcu Jul 2022–Oct 2022
Built Turkish sentiment analysis models; created large datasets and conducted cross-lingual benchmarking.

Research Experience

Undergraduate Research, Koç University — Nov 2022–Mar 2025

- Data-Efficient Reinforcement Learning – supervised by Asst. Prof. Barış Akgün
- Robustness of Self-Supervised Models – supervised by M. Emre Gürsoy
- Novel Vision Self-Supervised Learning Objective
- Non-Contrastive Sentence Embeddings – with Prof. Deniz Yuret and Prof. Alper Erdoğan

Teaching

Guest Lecture – COMP 442 (NLP) — May 2024
TA – ENGR200, Probability — Oct 2022–Jan 2023
Tutor – MATH204, Differential Equations — Feb 2022–Jun 2022

Awards

Vehbi Koç Scholar
Anatolian Scholarship Program

Skills

Programming: Python, C, C++, Java, LaTeX
Deep Learning: PyTorch, Flax/JAX

Academic Service

Reviewer: ICLR 2025, ACML 2023

Publications

O. V. Cagatan, Ö. F. Tal, and M. E. Gürsoy. *Adversarial Robustness of Discriminative Self-Supervised Learning in Vision.* arXiv preprint arXiv:2503.06361, submitted to the IEEE/CVF International Conference on Computer Vision, 2025 (ICCV '25).

O. V. Cagatan and B. Akgun. *Uncovering RL Integration in SSL Loss: Objective-Specific Implications for Data-Efficient RL.* Submitted to the Reinforcement Learning Conference, 2025 (RLC '25).

Accepted to the NeurIPS 2024 Workshop: Self-Supervised Learning - Theory and Practice (NeurIPS SSL Workshop '24).

MMTEB Team, O. V. Cagatan. *MMTEB: Massive Multilingual Text Embedding Benchmark*. To appear in the Thirteenth International Conference on Learning Representations, 2025 (ICLR '25).

O. V. Cagatan. *SigCLR: Sigmoid Contrastive Learning of Visual Representations*. To appear in the NeurIPS 2024 Workshop: Self-Supervised Learning - Theory and Practice (NeurIPS SSL Workshop '24).

O. V. Cagatan. *UNSEE: Unsupervised Non-contrastive Sentence Embeddings*. To appear in the 18th Conference of the European Chapter of the Association for Computational Linguistics (EACL '24).

O. V. Cagatan. *ToddlerBERTa: Exploiting BabyBERTa for Grammar Learning and Language Understanding*. To appear in the CoNLL–CMCL 2023 Shared Task: The BabyLM Challenge (CONLL'23).

O. V. Cagatan and B. Akgun. *BarlowRL: Barlow Twins for Data-Efficient Reinforcement Learning*. To appear in the Asian Conference on Machine Learning, 2023 (ACML '23).