Konstantinos Zafeirakis

Curriculum Vitae

Faculty of Science
University of Amsterdam
Science Park 904, 1098 XH Amsterdam,
The Netherlands

Tel: +30 699 608 8489
Email: konstantinos.zafeirakis@student.uva.nl
Website: kzafeir.github.io
Google Scholar: Link
LinkedIn: Link

Research Interests

Reliable and Trustworthy Machine Learning, Adversarial Robustness and Verification, Privacy & Security in Neural Networks, Information Retrieval, Natural Language Processing, Computer Vision & Signal Processing

Education

University of Amsterdam, Amsterdam, The Netherlands **2024 – 2026**
MSc in Artificial Intelligence
GPA: 8.15/10

Technical University of Crete, Chania, Greece **2019 – 2024**
Diploma in Electrical and Computer Engineering
GPA: 8.81/10 | **Rank:** 1st in cohort upon admission
Thesis: *"Hallucination Detection in Image Inpainting"*
Supervisor: Dr. Grigorios Tsagkatakis

Research Experience

Elsevier B.V. & IRlab UvA, Amsterdam, The Netherlands **June 2025 – Present**
Machine Learning Researcher - Supervisor: Evangelos Kanoulas

- Conducted an in-depth literature review on advanced machine unlearning techniques, focusing on post-training data erasure, model compliance, and privacy-preserving ML systems.
- Analyzed methods including certified removal, influence functions, and distillation-based unlearning for dynamic model adaptation without full retraining.
- Ran experiments across multiple models and data partitions; investigated prompting strategies to optimize unlearning performance.

Jobly.ai.(pre launch), Los Angeles, CA, USA **Apr 2025 – Present**
Co-founder, link

- Co-founded an AI-powered platform for matching candidates with opportunities using NLP-based resume parsing and transformer models.
- Fine-tuned a cross-encoder (MS-MARCO MiniLM) on user feedback to improve future match quality for candidates.
- Built a robust end-to-end matching pipeline including preprocessing, model training, scoring and ranking.

Foundation for Research and Technology - Hellas, Heraklion, Greece **June 2023 – Aug. 2024**
Undergraduate Research Fellow - Supervisor: Grigorios Tsagkatakis

- Applied TensorFlow and PyTorch for computer vision challenges: ViT training, fine-tuned image classifiers, CAMs, and data preprocessing.
- Conducted advanced image processing research for hallucination detection in image inpainting.
- Designed and implemented an innovative architecture to identify and quantify hallucinations in remote sensing image inpainting, providing a novel means to precisely evaluate model performance and reliability.
- Presented findings at the AstroML Journal Club, contributing to actionable insights for journal submissions.

Teaching

University of Amsterdam, Amsterdam, The Netherlands **Fall 2025**
Teaching Assistant

- 52041COV6Y, Computer Vision 1. Graduate Teaching Assistant.
  Course Coordinators: Martin R. Oswald, Dimitrios Tzionas.

Publications

Submitted

Zafeirakis, K., & Tsagkatakis, G. (2025).
A Framework for Detecting and Quantifying Hallucinations in Remote Sensing Image Inpainting.
IEEE Geoscience & Remote Sensing letters. Link

Journal Articles (peer-reviewed)

van Erven, O., Zafeirakis, K., Smit, J., Smidi, J., & Buijs, L. (2025).
[Re] Cooperate or Collapse: Emergence of Sustainable Cooperation in a Society of LLM Agents.
Transactions on Machine Learning Research (TMLR). Reproducibility Certification. Link

Leadership & Student Activities

VIA - Amsterdam Information Sciences Association, Amsterdam The Netherlands **Sep 2025 – Present**
Member

IEEE-TUC Student Branch, Chania, Greece **Nov. 2019 – Aug. 2024**
Member

- Developed a machine learning-based anomaly detection system for post-flight rocket telemetry data, identifying 2 critical anomalies.
- Mentored 10+ new members on multiple projects, improving team performance.
- Led a workshop on "Introduction to Deep Learning with TensorFlow" for 30+ members, boosting student proposals in deep learning by 20

Technical Skills

Programming: Python (NumPy, Pandas), Java, C++, C, MATLAB, SQL
Machine Learning / AI: Deep Learning: PyTorch, TensorFlow, Keras; Classical ML: Scikit-learn; NLP: Transformers, HuggingFace; Computer Vision: CNNs, ViTs, Image Processing;
Data Analysis & Visualization: Matplotlib, Seaborn, Pandas, NumPy
Tools & Platforms: Git, MLflow, W&B, Docker, AWS, Azure, Linux
Signal / Image Processing: Fourier / Wavelet Analysis, Filtering, Computer Vision Pipelines