

CAVIT ÇAKIR

Munich, Germany | cavitcakir3@gmail.com | linkedin.com/in/cavitcakir | github.com/cavitcakir

SUMMARY

Machine Learning Engineer with 3+ years in AI, focused on Computer Vision and NLP. Holds a Master's in Informatics from TUM. Skilled in building scalable AI applications, optimizing data pipelines, and accelerating image processing by 5x in current role. Experienced in deploying on AWS and GCP, with strong Python, C++, and PyTorch expertise. Passionate about driving innovation through technology.

EDUCATION

Technical University of Munich (TUM)

Master of Science in Informatics

Munich, Germany

Oct 2021 – Mar 2024

- Specializations: Computer Vision, Natural Language Processing

Sabancı University

Bachelor of Science in Computer Science and Engineering

Istanbul, Turkey

Sep 2016 – Jun 2021

- Specializations: Software Engineering, Natural Language Processing

TECHNICAL SKILLS

Languages: Python, C++

Machine Learning: PyTorch, NumPy, pandas, TensorFlow, Scikit-Learn, OpenCV, BentoML, Ollama, Langchain, RAG, LLM

Databases: Qdrant, MongoDB, MySQL

Developer Tools: Git, Docker, Bash, AWS, Google Cloud Platform

Web Development: React, Node.js, FastAPI

Natural Languages: Turkish (native), English (advanced), German (basic - learning)

EXPERIENCE

Machine Learning Engineer

Quasara GmbH

May 2024 – Present

Munich, Germany

- Designed and implemented a scalable vectorization pipeline that leverages state-of-the-art zero-shot multimodal models to process large-scale image datasets from S3 and store them in a Qdrant database, achieving a **5x increase in vectorization speed**, thereby reducing computational costs and generating substantial savings for the company.
- Enhanced small object recognition with multimodal LLMs and computer vision techniques, achieving **2x faster processing with 3% accuracy trade-off** and **3x speed with 5% trade-off**.
- Built and maintained a FastAPI backend, collaborating with client engineering teams for integration.
- Deployed scalable solutions on Google Compute Engine and AWS EC2.
- Adapted to rapidly changing customer requirements to ensure timely project delivery.

Master Thesis Student

CARIAD - a subsidiary of Volkswagen |  [View Thesis](#)

May 2023 – Nov 2023

Ingolstadt, Germany

- Developed a novel self-supervised learning approach for 3D LiDAR semantic segmentation, utilizing Neural Radiance Fields (NeRFs) to extract volumetric features from image data.
- Distilled 2D image features into 3D domains, allowing LiDAR segmentation models to be trained with geometrically aware features without relying on labeled data.
- Demonstrated the potential of using self-supervised techniques to overcome the limitations of densely labeled 3D datasets.

Computer Vision Engineer Working Student

Quasara GmbH

Dec 2022 – May 2023

Munich, Germany

- Led the end-to-end development of a Damage Classification Project using state-of-the-art transformer models, achieving over 93% accuracy in image classification.
- Implemented advanced cleaning and augmentation techniques to overcome imperfect labeled data.
- Enhanced client communication through regular updates and detailed presentations of project outcomes.

Natural Language Processing (NLP) Intern

FineSci Technology

Jul 2020 – Oct 2020

Istanbul, Turkey

- Contributed to the News Classification and Clustering Project by applying state-of-the-art transformer language models to classify Turkish news.
- Achieved over 96% accuracy in news classification, significantly improving efficiency and accuracy of categorization.

Undergraduate Teaching Assistant

Sabanci University

Feb 2019 – Feb 2020


Istanbul, Turkey

- Independently managed lab sessions for 20-30 students, providing hands-on support and guidance, and offered mentoring during office hours to enhance students' understanding and engagement with course materials.
- Utilized C++ as the primary programming language to facilitate practical learning in the Introduction to Computing course, demonstrating strong technical expertise and application skills.

SELECTED PROJECTS IN COMPUTER VISION

Implementation of Panoptic Neural Field


Oct 2022 - Mar 2023

Technical University of Munich | Python, Pytorch, NeRF |  [View Project](#)

- Implemented the *Panoptic Neural Field (PNF)* paper from scratch using *Kaolin Wisp* and the *KITTI 360 dataset*, demonstrating initiative and expertise in translating theoretical concepts into practical, executable code without existing implementations.
- Optimized the *PNF* architecture for better performance, employing advanced techniques to significantly improve model efficiency and accuracy, highlighting my ability to innovate and enhance existing methodologies.

Improving Point Cloud Transformer using Curve Aggregation

Apr 2022 - Sep 2022

Technical University of Munich | Python, Pytorch, Transformers |  [View Project](#)

- Enhanced the shape analysis capabilities of a *Point Cloud Transformer* by implementing a *Curve Aggregation* method, improving model performance for detailed 3D point cloud processing by 0.35%. This innovation led to more accurate shape interpretations on the *ShapeNet Parts* dataset.
- Converted the *Point Cloud Transformer* implementation from *Jittor* to *PyTorch*, demonstrating technical proficiency and versatility in adapting complex models to widely used frameworks.

SELECTED PROJECTS IN NLP

NLP and Knowledge Graphs for Research Cluster Prediction and Analysis


Oct 2022 - Mar 2023

Technical University of Munich | Python, Pytorch, Transformers, LLM |  [View Project](#)

- Contributed to the *Unsupervised Classification of Research Papers Project* by proposing a novel *Hierarchical Classification Method* and successfully applying existing techniques.
- Used the cutting-edge *SPECTER* embedding model to enhance the quality of embeddings, enabling a more precise understanding of research paper content and significantly contributing to the project's success.

Emotional Clustering of Social Media Users


Apr 2022 - Sep 2022

Technical University of Munich | Python, Pytorch, BERT |  [View Project](#)

- Utilized a *BERT model* to extract embeddings and cluster users based on their textual data, demonstrating proficiency in advanced NLP techniques for user segmentation.
- Preprocessed Reddit user posts to meet the requirements of the *BERT model*, showcasing the ability to prepare and adapt large-scale datasets for complex NLP tasks.
- Applied various *Dimensionality Reduction Methods*, including *HDBSCAN* and *KMeans*, to efficiently analyze and cluster high-dimensional data.

Meeting Scheduler Chatbot — Graduation Project

Sep 2020 - Jun 2021

Sabanci University | Python, JavaScript, React, Node.js, Docker |  [View Project](#)

- Developed an advanced chatbot using the *RASA Bot Framework*, integrating pre-trained natural language understanding methods to accurately process and interpret user inputs. This demonstrated expertise in utilizing state-of-the-art AI technologies to create intelligent conversational agents.
- Implemented a graphical user interface for the chatbot with *React* and *Node.js*, encapsulated in *Docker* containers to ensure ease of deployment and scalability. This end-to-end development highlighted my full-stack development skills and commitment to delivering robust, scalable applications.