# Amirabbas Afzali

*Curriculum Vitae*

+98 911 317 2840
amir8afzali@gmail.com
socials: - - -

## Research Interests

- Theory of Optimization machine learning and mathematics of data science
- Statistical Learning: Bandits & Probabilistic Graphical Models
- Natural Language Processing
- Reinforcement Learning and Preference Optimization in deep networks
- Graph Signal Processing and Graph Neural Networks
- Signal Processing

## Education

2021–2025 **Sharif University of Technology** | **SUT**, *Bachelor's degree*, Tehran, Iran,
Major in Electrical Engineering, Communication Systems sub-major, Minor in Mathematics.
*GPA – 19.39/20; Ranked 9th out of 184 students.*

2018–2021 **Shahid Beheshti High School (NODET)**, *High School Diploma*, Sari, Iran,
*GPA – 19.40/20.*

## Publications

2024 **Bias-Resilient Preference Optimization: Addressing Content-Aware, Multi-Source Biases in Preference Learning**, *Under review for NAACL 2025*, Stanford University, Advised by Prof. Sanjay Lall ,
This paper presents Bias-Resilient Preference Optimization (BRPO), a framework designed to improve preference alignment in LLMs by addressing multiple sources of bias. By using a multi-objective optimization approach and backdoor attack mechanisms, BRPO effectively separates genuine human preferences from biases.

2024 **Aligning Visual Contrastive learning models via Preference Optimization**, *Under review for ICLR 2025*, L3S Research center , OpenReview, Arxiv
This paper introduces a novel paradigm for enhancing contrastive learning models using Preference Optimization (PO) to align model behavior with specific preferences, improving robustness and fairness. Our approach addresses typographic attacks and mitigates gender bias, outperforming traditional contrastive methods in adversarial settings across vision-language tasks.

2024 **ULTra: Unveiling Latent Token Interpretability in Transformer-Based Understanding**, *Under review for CVPR 2025*, SUT, Advised by Dr. Sajjad Amini , Arxiv
This paper introduces a framework for interpreting Vision Transformers by examining how their internal token representations capture image semantics. We use gradient-based saliency maps to show that ViTs can perform zero-shot unsupervised semantic segmentation, identifying image regions without extra training.

2024 **Clustering Time Series Data with Gaussian Mixture Embeddings in a Graph Autoencoder Framework**, *Published at AAAI 2025 AI4Research Workshop*, Multimedia lab | SUT, Advised by Dr. Arash Amini , OpenReview, Arxiv
I have conducted research in time series data analysis, focusing on novel graph-based clustering methods. My work introduces a graph autoencoder designed for clustering tasks, leveraging a Gaussian mixture embedding for improved data separation. This approach has outperformed state-of-the-art methods and has been validated on time-series datasets.

### In-Process

2024 **Towards Fairer Retrieval: Sense Disentanglement and Bias Control Using Backpack Language**, *Expected Submission: ARR Conference, winter 2024*, IRLab | University of Amsterdam, Advised by Dr. Aliannejadi
In this project, we transformed a decoder-only language model named BackPack into an encoder model using an unsupervised approach. The model was trained specifically for ranking-based tasks. By breaking the monolithic structure of traditional transformer-based models, we leveraged this architecture to control and mitigate various biases, such as gender bias, in a ranking framework.

2024 **Efficient FMCW Index Modulation Receivers**, *Expected Submission: IEEE Journal on Selected Areas in Communications, winter 2024*, SUT , Advised by Dr. Mojahedian
This work developed a memory-efficient receiver for chirp signal index modulation, using the Wigner Distribution Function to reduce computational overhead. A transformer-based receiver with frequency attention was also designed to enhance chirp signal classification accuracy.

## Course projects

2024 **GAN-BERT**, *Deep Learning project*, SUT, **source**
Implemented the GAN-BERT architecture, which adversarially trains a BERT-based generator against a discriminator to detect and classify LLM-generated texts to the specific model used for generation.

2024 **Hidden Markov Map Matching Through Noise**, *AI project*, SUT, **source**
Implement a map-matching algorithm that uses a Hidden Markov Model (HMM) to find the most likely road route represented by a time-stamped sequence of latitude/longitude pairs.

2023 **Regression problem on CarPrice-assignment dataset**, *Applied Statistics project* , SUT, **source**
A project with the purpose of getting familiar with Regression and some classical ML methods.

2022 **Social Media app**, *OOP project* , SUT, **source**
Developed an application in Java with a GUI using JavaFX and CSS, integrated with a MySQL database for efficient data management.

## Experience

### Research Experience

Present **Bachelor's project**, Advised by Dr. Sajjad Amini, in collaboration with Dr. Moosavi-Dezfooli, Prof. Sanjay Lall, and Amirhossein Afsharrad, SUT, (Initial Presentation)
Finetuning contrastive learning models to have adversarially robust embeddings via constrained optimization.

Present **Volunteer researcher**, Conducting research and contributing to the ongoing projects at Aktus AI, focusing on leveraging artificial intelligence for innovative applications.

2023–2024 **Research project**, Conducting research on early Alzheimer's disease risk assessment using pretrained language and audio models. Analyzing speech patterns of healthy individuals and Alzheimer's patients, and investigating neural mechanisms underlying decision-making and cognitive dementia recognition, Ambient Intelligence Research Lab | SUT.

### Voluntary Teaching Assistance

- Deep Generative Models
- Artificial Intelligence
- Engineering Probability and Statistics
- Object Oriented Programming
- Deep Learning
- Signal processing
- Computer Structure
- Electronics 1 & 2

### Relevant Coursework

The symbol "+" denotes graduate coursework.

- High Dimensional Probability +
- Stochastic Process
- Signals and Systems
- Artificial Intelligence
- Machine Learning with Graphs + (Auditing)
- Reinforcement Learning +
- Deep Learning +
- Convex Optimization 1
- Machine Learning
- Online Learning & Bandits + (Auditing)

## Skills

### Technical Skills

**Programming:** C, C++, Java, Python, MATLAB, Julia, R, Verilog, Assembly, SQL

**Machine Learning Libraries:** PyTorch, Tensorflow, Pytorch-Geometric, OpenCV, Scikit-learn, NumPy, Pandas, Matplotlib

### Soft Skills

**Languages:** Persian (native), English (advanced)

**Misc:** Problem-Solving, Collaboration, Communication, Teaching