

# Taha ValizadehAslani

Taha.Valizadeh@Gmail.com  
tahaaslani.github.io

Scholar.Google.com/citations?user=WxaHx0sAAAAJ  
GitHub.com/TahaAslani

## Experience

### Comcast

Software Engineer

Sept 2023 – Present

Philadelphia, PA

- Building a metaheuristic pipeline for designing the topology of the fiber cable network at layer 3 including choosing core-site and core-circuits
- Using the Monte Carlo simulation for detecting network capacity overloads in different stochastic traffic distribution scenarios
- Forecasting future network traffic utilization based on the previous time-series data, at national, regional, and site level with top-down scaling
- Developing data engineering and analysis pipelines to monitor network traffic utilization while ensuring the quality of the underlying data.
- Estimating network link latency based on delays in probe readings as a linear programming problem

## Education

### Drexel University

Ph.D. in Electrical Engineering

Sept 2018 – Mar 2024

Philadelphia, PA

- Predicting individual treatment effect of Product-Specific Guidances on submission of generic drugs using cause-effect inference models
- Model blind enhancement of embedding representation of a large language model (LLM) (e.g. GPT-3 and GPT-4) for down-stream tuning
- Pre-training PharmBERT: a domain-specific LLM for pharmacological text mining that outperforms the current state-of-the-art models
- Enhancing the product-specific guidance assessment process of generic drug products for FDA using modern NLP solutions
- fine-tuning LayerNorm as a novel method for decreasing the computational complexity of training transformer-based LLMs by 99%.
- Proposing two-stage fine-tuning as a new deep learning training strategy for learning imbalanced data
- Using LLMs for relationship extraction to detect adverse reactions and drug-drug interactions from drug labels
- Developing NLP solutions based on LLMs for sentiment analysis, question answering, relationship extraction, and regression
- Deep learning data augmentation using back translation, text editing, and hidden layer embeddings representation mixing
- Predicting antibiotic resistance from the genome of bacteria with a novel feature: counting amino-acid k-mers.
- Genomic data analysis and statistical hypothesis testing of cancer mutation signatures
- Statistical analysis of temporal habits for eating, sleeping, and light exposure data of healthcare shift workers

### Iran University of Science and Technology

M. Sc. in Electrical Engineering - Communication - System

Sept 2014 – Jan 2017

Tehran, Iran

- Mathematical analysis and improvement of error control performance of IS-LDPC codes, by a heuristic search for a high girth graph
- Analysis and simulation of peak-to-average- power ratio (PAPR) in orthogonal frequency division multiplexing (OFDM)
- Designing a CCTV image denoising system based on frequency-selective filters

### LoRESTAN University

B. Sc. in Electrical Engineering - Power

Feb 2009 – May 2013

Khorramabad, Iran

- Real-time automatic reactive power control in a power system using a feed-forward neural network

## Skills

**Programming languages:** Python, SQL, R, MATLAB, Bash, C, C++, Scala

**Machine learning libraries:** PyTorch, TensorFlow, Keras, XGBoost, Random forest, SVM, scikit-learn, statsmodels

**Natural language processing:** Hugging Face Transformers, BERT, GPT, LLaMA, Mistral, Instructor, fast.ai, NLTK

**Metaheuristic search:** Particle Swarm Optimization, Genetic Algorithm, Ant Colony, Differential Evolution, Multi-objective optimization

**Data science tools:** Apache Spark, Apache Hadoop, Apache Hive, Databricks, KNIME, t-SNE, PCA, Tableau, Matplotlib, seaborn, ggplot2

**High-performance computing:** Slurm Workload Manager, SGE, AWS, Microsoft Azure, Google Cloud Platform, Cerebras (CS-2), Google Colab

**Bioinformatic packages:** Biopython, BLAST, SAMtools, MAFFT, Ray assembly, MetaBAT 2, Bowtie

**Miscellaneous:** pandas, NumPy, SciPy, Git, Docker, Conda, Distributed systems, Big data, HTML, Linux, TeX, Microsoft Office

## Research

### Publications

**Taha ValizadehAslani**, Yiwen Shi, Jing Wang, Ping Ren, Yi Zhang, Meng Hu, Liang Zhao, Hualou Liang. “Two-Stage Fine-Tuning with ChatGPT Data Augmentation for Learning Class-Imbalanced Data”. *Neurocomputing*, 127801. 2024.

**Taha ValizadehAslani**, Hualou Liang. “LayerNorm: A key component in parameter-efficient fine-tuning”. arXiv:2403.20284.

**Taha ValizadehAslani**, Yiwen Shi, Ping Ren, Jing Wang, Yi Zhang, Meng Hu, Liang Zhao, Hualou Liang. “PharmBERT: a domain-specific BERT model for drug labels”. *Briefings in Bioinformatics* 24 (4), bbad226. 2023.

Yiwen Shi, Ping Ren, Jing Wang, Biao Han, **Taha ValizadehAslani**, Felix Agbavor, Yi Zhang, Meng Hu, Liang Zhao, Hualou Liang. “Leveraging GPT-4 for Food Effect Summarization to Enhance Product-Specific Guidance Development via Iterative Prompting”. *Journal of Biomedical Informatics*, 148, 104533. 2023.

Yiwen Shi, Jing Wang, Ping Ren, **Taha ValizadehAslani**, Yi Zhang, Meng Hu, Hualou Liang. “Fine-Tuning BERT for Automatic ADME Semantic Labeling in FDA Drug Labeling to Enhance Product-Specific Guidance Assessment”. *Journal of Biomedical Informatics* 138, 104285. 2023.

Yiwen Shi, **Taha ValizadehAslani**, Jing Wang, Ping Ren, Yi Zhang, Meng Hu, Liang Zhao, Hualou Liang. “Improving Imbalanced Learning by Pre-finetuning with Data Augmentation” *Fourth International Workshop on Learning with Imbalanced Domains: Theory and Applications*. 2022.

Waleed Iqbal, Elena V. Demidova, Samantha Serrao, **Taha ValizadehAslani**, Gail Rosen, Sanjeevani Arora. “RRM2B Is Frequently Amplified Across Multiple Tumor Types: Implications for DNA Repair, Cellular Survival, and Cancer Therapy”. *Frontiers in genetics*, 12, 628758. 2021.

**Taha ValizadehAslani**, Zhengqiao Zhao, Bahrad A Sokhansanj, Gail L Rosen. “Amino Acid k-mer Feature Extraction for Quantitative Antimicrobial Resistance (AMR) Prediction by Machine Learning and Model Interpretation for Biological Insights”. *Biology* 9 (11), 365. 2020.

Chiahui Chen, **Taha ValizadehAslani**, Gail Rosen, Carla Jungquist, Laura Anderson. “Healthcare Shift Workers’ Temporal Habits for Eating, Sleeping, and Light Exposure: A Multi-Instrument Pilot Study”. *Journal of Circadian Rhythms* 18. 2020.

**Taha ValizadehAslani**, Abolfazl Falahati. “An analysis and improvement of error control performance of IS-LDPC codes with a large number of subsets”. *Physical Communication* 31, 79-86. 2018.

## Academic Peer Review

Reviewed for *PLOS Digital Health*

## Teaching

### Drexel University

Teaching Assistant

- ECE T480 Applied Robotics
- ECE 391 Introduction to senior design
- ECE T380 Introduction to digital signal processing
- ECEP 371 Introduction to Nuclear Engineering
- ECES 352 Introduction to digital signal processing
- ECES 306 Analog and digital communications
- ECE 304 Remote Sensing and Control
- ECES 304 Dynamic Systems and Stability
- ECE 303 Micro controllers
- ECES 303 Transform Methods II
- ECE 301 Foundations of Electric Circuits II
- ECES 301 Signals and systems
- ENGR 231 Linear engineering systems
- ECE 201 Foundations of electric circuits
- ENGR 131 Introductory programming for engineers

Sept 2018 – Jun 2023

Philadelphia, PA

### Iran University of Science and Technology

Teaching Assistant

- Fundamentals of communication systems

Feb 2016 – Apr 2018

Tehran, Iran

### Lorestan University

Instructor

- MATLAB programming

Sept 2010 - Jan 2011

Khorramabad, Iran

## Extracurricular

- Tango dancer
- Translator of “FlyPaper” by Robert Musil into Farsi, Ruzegare Now publication, Tehran, 2011
- Founding member of Kanun Nevisandegan Lorestan (Association of Writers of Lorestan)