Benyamin Tafreshian

benyamint.dev | bentsec@proton.me | linkedin.com/in/benyamint | github.com/btafreshian

Education

Boston University, Metropolitan College **Jan 2024 – May 2025**
Master of Science in Computer Science (Concentration: Security) *Boston, MA*
GPA: 3.97 (out of 4.0)

University of Tabriz **Sep 2016 – Sep 2020**
Bachelor of Science in Computer Engineering *Tabriz, Iran*
GPA: 15.86 (out of 20.00)

Honors and Awards

Excellence in Graduate Studies Award for Computer Science **May 2025**
- Honored by Boston University Metropolitan College as the top computer science graduate student for academic performance and research in security and artificial intelligence (AI)

NCAE Cyber Games **2024 & 2025**
Team Captain (2024) & Mentor (2025)
- Led a nationally-ranked cyber team in advancing from a field of 90+ national teams to the invitational rounds, ranking fifth nationally in 2024 and fourth in 2025

Red Team Village (DEF CON Village) **Apr 2024**
Honorary Member
- Selected as an honorary member, in recognition of fully clearing capture the flag (CTF) boards in multiple competitions hosted by the village

Research Experience

Boston University **May 2025 – Present**
Research Assistant (Advisor: Professor Reza Rawassizadeh) *Boston, MA*
- Developed a framework to decompose large language models (LLMs) into interpretable modular components, enabling targeted visualization and analysis of architectures, parameter groups, and weight distributions
- Investigated task-specific pruning methods for LLMs to create lightweight, resource-efficient variants suitable for deployment on edge and embedded devices under strict latency and memory constraints
- Developed an ESP32 embedded system for a wearable device that performs continuous food image capture and reliable transfer to backend servers to enable real-time calorie estimation
- Designed, 3D printed, and experimentally evaluated custom wearable enclosures, integrating sensors, microcontrollers, and power delivery to ensure user comfort and robustness in daily use

Boston University **Feb 2024 – Jan 2025**
Research Assistant (Advisor: Professor Shengzhi Zhang) *Boston, MA*
- Investigated vulnerabilities in machine learning (ML) models and techniques for manipulating AI systems to induce incorrect or unsafe behavior
- Developed a security framework for network intrusion detection under adversarial conditions, improving threat detection accuracy by 35% and reducing false positives by 12.5%
- Published peer-reviewed research on adversarial machine learning and network intrusion detection in **IEEE TrustCom 2024**

Teaching Experience

Boston University **Sep 2025 – Present**
Teaching Assistant (Course: Big Data Analytics) *Boston, MA*
- Facilitated lab sessions and designed coursework on distributed processing and large-scale data analytics
- Guided 40+ students in using **Spark, Hadoop**, and cloud tools for practical data engineering and analytics projects
- Mentored students in designing scalable data pipelines and deploying ML to production datasets

Boston University **Jan 2025 – May 2025**
Teaching Assistant (Course: Network Security) *Boston, MA*
- Evaluated coursework and provided detailed feedback for graduate students
- Taught labs on firewall configuration, penetration testing, and cloud security
- Developed interactive modules on threat modeling and incident response using real-world scenarios

Publications

- B. Tafreshian (2025). **RoguePrompt: Dual Layer Ciphering for Self-Reconstruction to Circumvent LLM Moderation**, *Preprint*
- B. Tafreshian and S. Zhang (2024). **A Defensive Framework Against Adversarial Attacks on Machine Learning-Based Network Intrusion Detection Systems**, *IEEE 23rd International Conference on Trust, Security and Privacy in Computing and Communications (TrustCom)*

Projects

- **96-Bit Block Cipher** | *C++, CMake, Python*
  Designed a 96-bit block cipher, integrating Rijndael substitution boxes with a novel Rubik's Cube–style permutation box, and evaluated its robustness against linear, differential, key-dependent, and statistical attacks
- **Fraud Detection in Big Data using Apache Ignite** | *Kafka, Ignite, Dask*
  Implemented a streaming fraud detector on Apache Kafka and Ignite, which uses Scikit-learn models to score 10,000 transactions per second, achieving 95% accuracy on PaySim data with latency under 100 milliseconds
- **RedditEngage: ML-Based Comment Score Predictor** | *Spark, MLlib, Scikit-learn*
  Built a four-node Spark pipeline of 37 million+ Reddit comments with 19 engineered features, achieving an AUC of 0.706; used SHAP-based analysis for model interpretability
- **ParamAtlas: LLM and VLM Parameter Inspector** | *Python, PyTorch, Transformers*
  Developed a parameter inspector for large language and vision language models that traverses Hugging Face and vLLM models to construct module trees and export structured summaries in text, CSV, and JSON formats
- **NeuroViz: Neural Network Visualization** | *TypeScript, React, Next.js*
  Created a Next.js interface in TypeScript and React to enable interactive exploration of neural network architectures and experiment results with reusable components and deployment through a Vercel-based workflow

Relevant Coursework

Boston University

- Cryptography
- Network Security
- Generative AI
- Big Data Analytics
- Digital Forensics
- Enterprise Cybersecurity

University of Tabriz

- Data Mining
- Computer Networks
- Microprocessors
- Database Design
- Compiler Design
- Operating Systems

Technical Skills

Certifications: ISC2 Certified in Cybersecurity (Jan 2025)
Programming Languages: Python, C, C++, x86 Assembly, Verilog
Tools and Frameworks: Tamarin, ProVerif, Z3, TransformerLens, DeepSpeed, Scikit-learn, Dask, Docker, Git, Bash, PowerShell, LaTeX

Additional Work Experience

Poyan Sannat Farnad **Nov 2020 – Sep 2023**
IT Manager *Tabriz, Iran*

- Established security policies and lightweight audits, reducing cyber risk by 25%
- Conducted company-wide risk assessments and improved compliance, decreasing security incidents by 15%
- Implemented and monitored secure network designs, ensuring system performance across teams

Poyan Sannat Farnad **May 2019 – Sep 2019**
IT Intern *Tabriz, Iran*

- Provided technical support and diagnostics for networked systems, reducing system downtime by 20%
- Assisted senior engineers with security audits and initiatives to improve system uptime and reliability