Ali Firooz

(Seyed Ali Firooz Abadi)
Computer Science, PhD · Chemical Engineer, MS
State College, PA 16801

☎ +1 (803) 662-6908 | ✉ Ali.emad649@gmail.com | ⬡ AFirooz | in Ali-Firooz | ☑ **Permanent Resident** | ⌂ Open to Relocate

Summary

- Highly motivated interdisciplinary scientist with expertise in **data science, machine learning, artificial intelligence,** and **chemical engineering,** driven by a passion for leveraging cutting-edge technologies to solve complex real-world problems.
- **3+ years** of hands-on experience in developing ML,DL, and GenAI techniques, data management pipelines, SQL/No-SQL databases, to process, analyze, and extract insights from complex, high-throughput datasets.
- Proficient in applying advanced data processing techniques, including dimension reduction, machine learning algorithms, and deep learning models, to build accurate predictive models and enable data-driven decision-making.
- Experienced in leveraging deep learning techniques to solve real-world problems in areas such as signal classification, image classification, sentiment analysis, and natural language processing.
- Collaborating with prestigious institutions, including **Prisma Health Cancer Institute** and **School of Medicine Greenville,** to drive innovative research and develop novel approaches to cancer profiling and patient outcome prediction.

Skills

Programming	**Fluent:** Python, SQL, Shell/Bash **| Prior Experience:** Neo4j, R, C++, C#
LLMs	
Frameworks	PyTorch, Scikit-learn, Multiprocessing & AsyncIO, Node, .Net
Web Development	Flask, HTML, CSS, JavaScript, TypeScript, PHP, Bootstrap
Technologies	**Fluent:** Azure, Git, GitHub, Azure DevOps, Docker, LaTeX, Linux
	Prior Experience: AWS, GCP, SAS, Tableau
Tools	Experimental Design and Statistical Analysis (Minitab), Design Expert, Emacs & Vim, VSCode & JetBrains

Experience

University of South Carolina
Graduate Research Assistant

Jan 22 - Present
Columbia, SC, US

- Managed high-throughput liquid chromatography-mass spectrometry (LCMS) data (**over 6.9 billion data points** across 78 files), storing it into a MySQL database by leveraging the advantages of SQLAlchemy's asyncio methods to expedite the process. Explored neural networks like ResNet and CNN to differentiate between cancerous and surrounding tissues based on the processed data
- Applied dimension reduction techniques (PCA, t-SNE, ISOMAP) to perform cluster analysis and reveal natural groupings in data comprising over **77,000** small molecules and their chemical descriptors. Performed retention time prediction for these small molecules based on their chemical descriptors. Trained baseline models for initial feature selection and data understanding.
- Designed a Cancer-Profiling Database adhering to normal form standards, and utilized SQLAlchemy's Object Relational Mapping (ORM) for MySQL connectivity. The database currently holds over 10,000 records and continues to grow
- Developed a Python pipeline for parsing, cleaning, verifying, and loading data onto the database, leveraging multiple frameworks including Numpy, Pandas, etc. The data are processed from Prisma Health, enhancing the efficiency and accuracy of data management processes
- Conducted research on machine learning (ML) algorithms (using the Scikit-Learn library), including tree-based methods (Decision Trees, Random Forest), and gradient boosting techniques (XGBoost, AdaBoost, CatBoost), to classify patient responses to different combinations of 65 medications. Employed evaluation metrics (CM, F1, recall, precision, etc.) to assess model performance. Visualized the result (using SHAP) for enhanced interpretability and decision-making, while addressing class imbalance issues (using down sampling, SMOTE) and ensuring robust and reliable predictions.
- Contributed to the development of TransONet, a deep learning model with ResNet encoders and Transformer bridges integrated into a UNet structure, achieving **93.5% Dice accuracy** for aortic segmentation in CT scans.
- Led a **5-member team** in developing an End-to-End Web App (Certificate Management System) in an Azure Cloud Native environment. Leveraging various Azure services including SQL Database, Service Bus, Function Apps (with Blob Storage and Service Bus triggers), Logic Apps (with Microsoft 365 Outlook connector), and Data Factory. Applied agile methodologies and CICD practices throughout the development process. Implementing features such as data integration from external feeds, CRUD operations with notifications, dashboard, user profile, and certificate catalog. Used Dotnet (.NET) Entity Framework Core, .NET Core 7.0 (C#), REST API, and React, in addition to following a microservice architecture with serverless components.
- Classified images using Vision Transformer (ViT) architectures using PyTorch and Lightning following object orientation programming (OOP).
- Conducted sentiment analysis on TripAdvisor reviews dataset using Hugging Face ALBERT model, including text tokenization, and fine-tuning.

Chemistry and Chemical Engineering Research Center of Iran *Jan 16 - May 20*
Graduate Research Assistant (Chemical Eng) *Tehran, Iran*

- Skills: Computational Fluid Dynamics (CFD) using ANSYS Fluent, SpaceClaim, and ICEM.
- CFD Simulation of a pilot-scale Bubble column reactor: Hydrodynamics, Temperature, and Concentration
- Study of Process Parameters on Crystal Size Distribution and Crystallization Hydrodynamics
- Polymeric composite rheology analysis and characterization

Projects

2024	**Classify steel scraps using ResNet and Vision Transformer (ViT) architectures in PyTorch**,
2024	**Performed sentiment analysis on TripAdvisor hotel reviews dataset using ALBERT (A Lite BERT) model from Hugging Face library**, tasks: text preprocessing, tokenization, and fine-tuning
2024	**Retention time prediction of over 77,000 small molecules based on their chemical descriptors**,
2024	**Creation of an End2End Web App (Certificate Management System) in Azure Cloud Native environment**,
2024	**Led a team of four undergrads at the USC hackathon, developing a chatbot using RAG (Retrieval-Augmented Generation) and local LLMs (Large Language Models)**,

Leadership & Volunteer Experiences

- Led a team of 4 undergraduates at the USC Hackathon to develop a chatbot leveraging Retrieval-Augmented Generation (RAG) technique. The project utilized MongoDB Atlas Vector Search, LangChain for orchestration, Ollama embeddings for enhanced contextual understanding, and LangChain text splitters for efficient document parsing and management. *2024*
- Supervised 2 undergrads in the development of a Flask web interface for database, implementing security measures, such as input validation, and parameterized queries to prevent vulnerabilities like SQL injection and cross-site scripting (XSS) attacks. *2023*
- Voluntarily graded and provided feedback for Advanced C++ Programming Techniques course. *2023*
- Taught Linux workshop with 20 participants, at SC INBRE Bioinformatics Core Summer Workshops. *2022*
- Assisted 20+ first-year CS students with Web Development in HTML, CSS, and JavaScript. *2022*

Education

University of South Carolina (USC) *Jan. 2022 - Present*
PhD & MS in Computer Science and Engineering *SC, US*

- GPA: 3.72 /4 (updated Jan 2025)
- Enrolled in a dual-degree program and **awarded my MS in CS in December 2024**
- Relevant coursework: Advanced Algorithms, Database System Design, Artificial Intelligence, Deep Learning

Chemistry and Chemical Engineering Research Center of Iran (CCERCI) *Sep 2016 - Mar 2019*
MS in Chemical Engineering *Tehran, Iran*

Arak University *Sep 2012 - Aug 2016*
BS in Chemical Engineering *Arak, Iran*

Honors & Awards

2023	**AACR Associate**,	The American Association for Cancer Research	*US*
2023	**First Place in the Cancer Category**,	INBRE SE IDeA Conference	*US*
2023	**Highest Scoring Graduate Student**,	INBRE SE IDeA Conference	*US*

Certificates

2023	**Machine Learning Specialization**,	tinyurl.com/ML-specialization	*Stanford University*
2023	**Deep Learning with PyTorch**,	tinyurl.com/DL-PyTorch2	*DataCamp*
2023	**Advanced Learning Algorithms**,	tinyurl.com/adv-algo	*Stanford University*
2023	**Biomedical Investigator**,	tinyurl.com/citibiomed	*CITI Program*

Publication

- A. B. Rajeoni, B. Pederson, **A. Firooz**, et al. "TransONet: Automatic Segmentation of Vasculature in Computed Tomographic Angiograms Using Deep Learning." in 2023 International Conference on Computational Science and Computational Intelligence (CSCI), 2023, doi.org/10.48550/arXiv.2311.10328.

- B. K. Austin, **A. Firooz,** H. Valafar, and A. V. Blenda, "An Updated Overview of Existing Cancer Databases and Identified Needs", Biology, vol. 12, no. 8, 2023.

- **A. Firooz,** A. T. Funkhouser, J. C. Martin, W. J. Edenfield, H. Valafar, and A. V. Blenda, "Comprehensive and User-Analytics-Friendly Cancer Patient Database for Physicians and Researchers", in 2022 International Conference on Computational Science and Computational Intelligence (CSCI), 2022, pp. 1615–1621.

- **A. Firooz,** A. T. Funkhouser, J. C. Martin, W. J. Edenfield, H. Valafar, and A. V. Blenda, "Analysis of cancer patients' molecular and clinical data using artificial intelligence and machine learning approaches", Cancer Research, vol. 83, no. 7 Supplement, pp. 5425–5425, 04 2023.

- **A. Firooz,** S. M. Noblitt, J. Martin, W. J. Edenfield, A. Blenda, and H. Valafar, "Application of machine learning in predicting breast cancer patient outcome," BMC Proceedings, vol. 17, no. 19, pp. O11, 2023. doi.org/10.1186/s12919-023-00281-y.