Mariana Khachatryan, PhD

marishnem@gmail.com, +1-757-337-9769 https://www.linkedin.com/in/marianakhachatryan
https://mariana-khach.github.io https://github.com/mariana-khach
https://scholar.google.com/citations?user=GFP-ZS8AAAAJ

Data Scientist with experience in developing and deploying end-to-end machine learning pipelines. Skilled in predictive modeling, quantitative analysis, data visualizations of big data using Python, SQL and C++. I am a trained Data Scientist from The Erdös Institute and a nanodegree holder of Udemy for Deep Learning and Artificial Intelligence.

Skills

- **Languages** & **Tools**: SQL, Python(Libraries: Numpy, SciPy, Pandas, Matplotlib, Tensorflow, Scikit-learn), Microsoft Power BI (view reports), C/C++, Git, Shell, Latex, SLURM (distributed computing).
- **ML** & **AI**: Deep Learning (CNNs, LSTMs, Autoencoders, Transformers), NLP, Supervised & Unsupervised Learning, Statistical Modeling, Time Series Forecasting, GenAI (LLM Fine Tuning, OpenAI API, Langchain, Building Chatbot, Prompt Engineering).

Experience

- **Erdös Institute Data Science Bootcamp** Columbus, OH
  Data Science Fellow *Sep. 2024 - Dec. 2024*
  - Led a team of PhD researchers in building machine learning models for predicting car price and identifying features driving model prediction using SHAP values. Achieved R^2 of 0.88 from XGBoost, reducing prediction error by 2×.
  - Communicated findings to stakeholders (CarMax, Upstart); built reproducible codebase in Python with Git.

- **Florida International University (FIU)** Miami, FL
  Postdoctoral Research Associate *Jan. 2020 - May 2023*
  - Used Python, probability and statistics for exploratory data analysis, data engineering and model development for $\sim$ 3.9 PBs particle physics data from GlueX collaboration at Jefferson Lab.
  - Used Least Squares and Maximum Likelihood statistical methods to clean data via Probabilistic Event Weightings. Implemented multivariate classification algorithms in C++ to classify particles. Applied bootstrapping techniques for error estimation. Used distributed computing on computing cluster with 5k nodes. Collaborated with a global team of 300+ scientists to exchange findings, refine data analysis strategies, and improve model accuracy.
  - Participated in collection, monitoring and quality check of $\sim$ 1.5 PBs data.
  - Mentored junior group members. Delivered presentations for technical and non-technical audiences. 40+ publications in peer reviewed journals. Recognized with a Certificate of Appreciation from Executive Dean.

- **Old Dominion University** Norfolk, VA
  Research Assistant *Jan. 2014 - 2019*
  - Utilized MS SQL to process the Jefferson Lab CLAS collaboration (over 200 physicists) large scale data using SLURM distributed computing on Lustre file systems. Used C++ for Monte Carlo simulations, error estimation and prediction of Electron-beam energy reconstruction. Mentored junior researchers. Received recognition and award. Data analysis results published in Nature journal. 3 media appearances (ODU1, ODU2, JLab).

Selected Machine learning and Deep learning Projects

- **Car sales price prediction 2024 GitHub URL**: Predicted car sales price based on various features.
- **Clothing Sales Forecasting 2025 GitHub URL**: Developed time series forecasting models using LSTM networks and Linear Regression to predict clothing sales trends. Achieved 2× improvement in model performance using Linear Regression.
- **Large Language Model (LLM) fine tuning, 2025 GitHub URL**: Fine tuning model from OpenAI API for email classification and serving it with FastAPI.

Recognition & Awards

- **Certificate of Appreciation**: Florida International University Executive Dean (2021)
- **Research Excellence**: Jefferson Science Associates Graduate Fellowship (2018-2019)
- **1st Place Prize**: Jefferson Lab Users Organization Poster Competition (2018)

Education

Ph.D. in Experimental Nuclear Physics, Old Dominion University 2019
M.Sc. in Physics, Old Dominion University 2015
M.Sc. and B.Sc. in Physics, Yerevan State University 2007-2013

[0]Languages: **English** (Fluent), **Russian** (Fluent), **Armenian** (Native)