Nehal Ahmed Shaikh

Hyderabad, Pakistan | +92 336 2758828 | nehalahmedshaikh@gmail.com | nehalahmedshaikh.github.io

EDUCATION

- *Sep 2020-Jun 2024:* BSc in Economics-Mathematics (Computer Science Minor) from Lahore University of Management Sciences (CGPA: 3.72/4).

RESEARCH INTERESTS

- Developing and improving AI algorithms—using mathematical tools from optimization and game theory—and exploring their applications in both natural and social sciences.

- Advancing the theory of large language models, exploring their applications, and ensuring that they continue to benefit human society.

UNDERGRADUATE RESEARCH

- **Unrolling Optimization Algorithm for Matrix Completion (Jun 2023-Jun 2024)**

  Advisor: Dr. Muhammad Tahir
  Summary: Initially refined a deep-unfolded algorithm for vanilla matrix completion problem (see: convmc-net) and improved upon it by unrolling an existing iterative optimization algorithm into a neural network, eventually proposing a potentially faster and more accurate novel algorithm that does not only work for the standard problem, but also for cases involving GMM noise (see: ConvHuberMC-Net). The results for these two algorithms were compared against existing iterative MC algorithms by replicating their results (see: report).

SELECTED COURSEWORK

Dynamic Programming and Reinforcement Learning (A-), Machine Learning (A), Generative AI for Speech and Language Processing (A), Deep Learning (A), Econometrics II (A-), Convex Optimization (A+), Principles and Techniques of Data Science (A), Introduction to Artificial Intelligence (A), Probability (A).

PROJECTS

- **Speech Recognition and Translation System For Medical Communication (Python)**

  Summary: Combined speech recognition, machine translation, and TTS functionality, selecting relevant vector databases, fine-tuning large language models (LLMs), and employing RAG pipelines to enable users to receive real-time state-of-the-art medical advice and information (see: repository).
  Skills learned: vector database selection, LLM fine-tuning, RAG

- **Panel Data and Tobit Analysis on Health Care Dataset (Stata)**

  Summary: Analyzed healthcare service utilization in Germany employing tobit, fixed effects, and random effects models on a unique panel data set spanning from 1984 to 1995 (see: data), while focusing on socio-economic disparities in healthcare access and hence contributing critical insights to health-care economics. (see: PDF).
  Skills learned: tobit analysis, panel data analysis, policy evaluation

- **Air Pollution & Academic Performance (Python)**

  Summary: Explored the relationship between air pollution and academic performance in Pakistan to know whether there exists any relationship between the two and to predict academic performance scores based on a customized synthesis of ASER (education) and NASA (climate) data (see: blog).
  Skills learned: data synthesis, web-scraping, deep learning for regression analysis

- **Sentiment Analysis on Audio Recordings (Python)**

  Summary: Applied various classifiers (K-NN, logistic regression, naive Bayes, SVM, and neural network) on CREMA-D to predict the emotion signified by an audio and performed comparative metric analysis (see: PDF).
  Skills learned: speech processing, classification algorithms, hyperparameter tuning, sentiment analysis

- **Regression Analysis to Estimate the Effect of Gender on Academic Performance (Stata)**

  Summary: Utilized a multiple linear regression model, featuring a host of explanatory variables, to conclude whether there is any significant difference between the CGPA values of male and female students, using field data collected through questionnaires (see: PDF).
  Skills learned: assumptions validation, data cleaning, econometric modeling, questionnaire design

- **Regression Analysis on the Determinants of New York Housing Prices (R)**

  Summary: Created a multiple linear regression model to determine the significance of factors affecting housing prices in New York and the effect of outliers by performing extensive data analysis (see: data and PDF).
  Skills learned: exploratory data analysis, feature selection, statistical modelling

- **Deep Learning for Computer Vision (Python)**

  Summary: Implemented four deep learning algorithms—CNN, auto-encoder, LSTM, and ViT—for computer vision tasks (see: repository).

WORK EXPERIENCE

- *Jan 2024-Jun 2024:* Teaching assistant for Convex Optimization (graduate-level course), taught by Dr. Hassan Mohy Ud Din at Lahore University of Management Sciences.

- *Jun 2024-Now:* Research assistant for diffusion models in medical imaging, supervised by Dr. Hassan Mohy Ud Din at Lahore University of Management Sciences.

- *Sep 2024-Now:* Data Scientist I at Afiniti.

HONORS AND AWARDS

- Placed four times on Dean's Honor List at Lahore University of Management Sciences for excellent academic performance.

- Graduated with distinction from Lahore University of Management Sciences.

SKILLS

Programming Languages:	C++, Python, R, SQL
Programming Software:	MATLAB, RStudio, Stata, Visual Studio Code
Other Skills:	Canva, Git(Hub), LaTeX, Linux, Microsoft 365 Suite