Nehal Ahmed Shaikh

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

EDUCATION

University:	Lahore University of Management Sciences	Major:	Economics-Mathematics
Duration:	July 2020 - July 2024	Minor:	Computer Science
Degree Type:	BSc	CGPA:	3.72/4.00

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.

RESEARCH EXPERIENCE

Topic:	Unrolling Optimization Algorithm for Matrix Completion
Advisor:	Dr. Muhammad Tahir
Expected Duration	12 months
Status:	Ongoing

Summary: Initially refined a deep-unfolded algorithm for vanilla matrix completion (MC) 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).

RELEVANT COURSEWORK

Dynamic Programming and Reinforcement Learning, Machine Learning, Generative AI for Speech and Language Processing, Deep Learning, Econometrics II, Convex Optimization, Principles and Techniques of Data Science, Introduction to Artificial Intelligence, Numerical Analysis, Statistics, Probability, Introduction to Differential Equations.

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 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 on CREMA-D data—K-NN, logistic regression, naive Bayes, SVM, and neural network—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 modelling, 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).

EXPERIENCE

Teaching assistant for EE 563/MATH 325: Convex Optimization (graduate-level course).

HONORS AND AWARDS

Placed four times on Dean's Honor List at LUMS for excellent academic performance.

SKILLS

- Programming: C++, MATLAB, Python, R
- Software: Canva, Microsoft 365 Suite, RStudio, Stata, Visual Studio Code
- Other: Git, LaTeX