

# SHEHRYAR HANIF

shehryarhanif123@gmail.com | www.shehryarhanif.com & www.github.com/ShehryarHanif | +971 501 450 552

---

## EDUCATION

### **New York University**

Abu Dhabi, UAE & New York, USA

*BS in Computer Science, BS in Mathematics*

Aug. 2019 – May 2023

Class Rank: Magna Cum Laude CGPA: 3.987/4.000

Relevant Coursework: Applied Internet Technology, Data Management and Analysis, Natural Language Processing, Parallel Computing, Operating Systems, Computer Networks, Algorithms, Data Structures, Linear and Nonlinear Optimization

## PROGRAMMING SKILLS

Skilled In: **Python** (Including **Scikit-learn**, **NLTK**, **Pandas**), **JavaScript**\**TypeScript** (Including **Express.js**, **Node.js**), **React** (Including **Hooks**, **Redux**, **Routers**), **PostgreSQL**\**MySQL**, **GCP**, **HTML**, **CSS** (Including **Bootstrap**)

Familiar With: **C++**\**C** (Including **MPI**, **OpenMP**, **CUDA**), **MATLAB**, **R**, **MongoDB**, **AWS**, **Heroku**, **Firebase**

## PROFESSIONAL EXPERIENCE

### **Software Development Engineer**

Jun. 2023 – Present

*noon*

Dubai, UAE

- Develop the backend for noon Food's marketplace system, specializing in catalog, search, and content management system(s)
- Led engineering efforts to revamp the order review system, specifying product requirements and coding the feature in **Python** and **MySQL**
- Optimize query performance for **Google Cloud Platform (GCP)** tools, including **BigQuery** and **Spanner**

### **Data Science Intern**

Jun. 2022 – Aug. 2022

*Selesnick Lab @ NYU Tandon School of Engineering*

New York City, USA

- Optimized and parallelized **MATLAB** program for pair-sparse signal recovery, while simultaneously prototyping potential improvements for the base algorithm in **Python** (particularly **Scikit-learn** and **NumPy**)
- Visualized stepwise recovery process by using **Matplotlib** for **Python** and in-built plotting library for **MATLAB**
- Improved recovery accuracy by ~50 % for common use cases involving a signal-to-noise ratio of about 100

### **Full-Stack Development Intern**

Jun. 2021 – Nov. 2021

*Frich*

New York City, USA (Virtual)

- Devised, designed, and developed web application for student recruits to register, log users, and chart progress through college ambassador program
- Built back-end **REST API** in **Express.js**, and served **MySQL** data to custom-built **React** application through **React Routers** with **Axios**
- Implemented **JWT** authentication, designed UX interface, and maintained hosting on **Heroku** with **AWS (RDS)**

## RESEARCH EXPERIENCE

### **Machine Learning Research Assistant**

Apr. 2021 – May 2024

*New York University Abu Dhabi*

Abu Dhabi, UAE

- Developed multilingual text-to-code translation models with Assistant Professor Riyadh Baghdadi (Computer Science)
- Implemented custom **Python** adaptation of **GPT-J** language model by using **PyTorch**, **Scikit-Learn** and **TensorFlow**, with testing indicating accuracy improvement of ~15 %
- Web-scraped, cleaned, and analyzed code snippets and explanations from online tutorial websites (e.g., geeksforgeeks.com) to build novel consolidated dataset for **Python** and **Java** text-code pairs

### **Data Science Research Assistant**

Sep. 2020 – Aug. 2021

*New York University Abu Dhabi*

Abu Dhabi, UAE

- Led coding and data analysis efforts for research group working with Assistant Professor Andy Harris (Political Science)
- Cleaned up 10+ datasets and conducted quantitative analysis (e.g., linear regression models, convex hull construction) by using **Python** (particularly **Pandas**, **NumPy**, and **Beautiful Soup**) and **R** (particularly **dplyr**, **ggplot2**, and **Shiny**)
- Fuzzy-matched and geo-mapped election polling stations in Malawi to identify voting irregularities and incidence of fraud

## PROJECT WORK

### **Extractive Text Summarizer for COVID-19 Medicinal Articles** | [bit.ly/3IRR9KO](https://bit.ly/3IRR9KO)

Nov. 2021 – Dec. 2021

- Developed **Python** program that extracted machine-produced summaries of COVID-19 research articles through modified implementations of **BERT** and **Word2vec** vectors built using **NLTK** and **Scikit-learn**
- Designed and implemented custom **TF-IDF** vector generation technique that gave greater weightage to medical terms produced through the open-source **Termolator** tool
- Achieved precision and recall improvement of ~6 % over established baseline approaches (e.g., random sampling)

## AWARDS AND HONORS

**Intern of the Month (July 2021)** *Frich*

Jul. 2021

**Top 10 in Pakistan, National Mathematics Talent Contest** *Higher Education Commission*

Dec. 2018