Sharjeel Mustafa
LinkedIn GitHub Portfolio
Email : sharjeeliv@gmail.com
Mobile : 226-340-7840

Technical Skills

Languages: Python, SQL, Java, C/C++, JavaScript
Frameworks: Pandas, NumPy, PyTorch, Flask, Matplotlib, SpaCy, JUnit, PyTest, Selenium
Technologies: Git, Neo4j, GraphQL, REST, Linux, Docker, CUDA

Education

- **B.Sc. Honours Computer Science, Software Engineering Specialization** Sep. 2020 – Aug. 2024
  University of Windsor; CGPA: 91.7/100 *Windsor, ON*

Experience

- **Undergraduate Researcher** Jan. 2024 – Aug. 2024
  University of Windsor *Windsor, ON*
  - **Data Engineering**: Developed an end-to-end Graph Neural Network (GNN) testing platform, consisting of data processing, modelling, training, and testing with modular compoents using **PyTorch, Pandas, and Pytorch Geometric**
  - **Neural Networks**: Implemented and integrated **6 GNN models** into the platform and presented GNN mechanisms to improve heterophily performance, contributing to a poster and research paper
- **Research Assistant: Software Engineering** Jan. 2022 – Aug. 2024
  University of Windsor *Windsor, ON*
  - **Data Engineering**: Developed a rule-based pipeline to extract topic related text from over **30,000** semi-structured documents using Python, Pandas, and SpaCy with a semi-supervised approach, achieving over **90% median accuracy**.
  - **Automation**: Developed a program using **Python, PyTesseract, RestFul APIs, and SpaCy** to extract article metadata from 1000's of articles and migrate to an online database, automating a manual process and improving input time by **94%**
- **R&D Data Scientist** May. 2023 – Dec. 2023
  Swift Medical *Windsor, ON*
  - Wrangled and analyzed over **180 million** rows of time-series health data using **SQL, Python, Pandas, AWS, Snowflake, and Matplotlib**, contributing to a research paper
  - Contributed to a multimillion-dollar initiative by creating a pipeline for multiple Relation Extraction methods, such as **fine-tuning LLMs on 12,000 clinical texts**, improving performance by **96%**
  - Validated models and extracted insights by creating ROC curves and analyzing trends in computed metrics like **AUC, F1, and balanced accuracy scores**, contributing to a paper
- **Software Developer** May. 2022 – Aug. 2022
  Connecting With Technology *Windsor, ON*
  - Designed and developed a Neo4j graph data model for social networks, served by a **GraphQL API**, streamlining graph traversals compared to previous relational model and **REST API**
  - Developed an asynchronous **CRUD** backend with **JavaScript, Fastify, and GraphQL** following MVC architecture to operate efficiently on highly connected graph data

Research

1. L. Goldstone, H. Mohammed, R. Gupta, **S. Mustafa**, R. Fraser, J. Allport. "Comparing Percent Area Reduction to Objective, AI-Powered Healing Index as a Metric of Wound Healing" (Submitted)
2. M. Hashemi, **S. Mustafa**, A. Ngom, L. Rueda. "HeteroGraphNet: Advancing Cell Type Prediction in scRNA-seq with Enhanced Graph Neural Network for Heterophilic Structures" (Submitted)