

HARSHIT SONI

New York, NY | 347-818-5361 | hsoni1@binghamton.edu | [Linkedin \(harshitsoni01\)](https://www.linkedin.com/in/harshitsoni01/) | harshitsoni.com | [GitHub \(harshitsoni01\)](https://github.com/harshitsoni01)

EDUCATION

Binghamton University, State University of New York

Expected Dec 2024

Master of Science in Computer Science, Concentration in Artificial Intelligence | **Cumulative GPA: 3.56 /4.00**

Relevant Courses: Data Structures and Algorithms, Machine Learning, Artificial Intelligence, Programming Language, Design Patterns

TECHNICAL SKILLS

Languages: Python, SQL, C, C++, HTML/CSS, JavaScript, Java

Tools: Flask, AWS (S3, EC2, IAM), GCP, Pandas, NumPy, GIT, MATLAB, Scikit-learn, TensorFlow, Keras, Cuda, Shell, ETL

Additional: Automation, Data Analysis, Image Processing, NLP, Quantitative Analysis, Statistical Modeling, Selenium, LLM, RDMS

Certifications: JPMorgan Software Engineering Virtual Internship Experience, Construct Stock Market Indices, Profit Analysis using Economic Value Added, Git for Developers, Mathematics for Machine Learning

PROFESSIONAL EXPERIENCE

Julius Education, Machine Learning Engineer | New York City, NY

Jun 2024 – Sept 2024

- Developed and deployed machine learning models for career dashboards, enhancing green labor seekers' experience through data integration. Utilized **AI**, RAGs, and prompt engineering techniques to optimize sentence structure for data dashboard renewable roles
- Led data labeling, scraping, and cloud service storage processes for government website data, ensuring accuracy and usability
- Collaborated with the **VP** on various client projects, contributing to diverse solutions such as building data pipelines including NREL

Trust Scout, Software Engineering Intern | New York City, NY

Jun 2023 – Aug 2023

- Engineered an automated email system and AI-driven prompter, reducing manual efforts by 80%, elevating client interaction by **40%**
- Led data analysis on **10M** real estate transactions and **1.6M** agents, delivering insights that directly improved decision-making and enhanced client relations, through statistical modeling, machine learning models, and big data for storage
- Collaborated with the **Founder/CEO** to enhance database and marketing strategies, resulting in a 25% increase in client satisfaction and a **33.3%** growth in the sales team. Technologies utilized: Python, AWS, Django, PostgreSQL, APIs, and OpenAI

RoboticsForSure, Software Engineering Intern | Remote

Jun 2021 – Aug 2021

- Implemented Robotics Process Automation (RPA) within ML and C++ projects, achieving a **30%** reduction in processing time by automating data preparation phases for clients
- Automated data analysis and math operations, and developed live ML solutions via the Google Cloud and AWS for data extraction

Fiverr, Freelancer | Remote

Mar 2021 – Aug 2021

- Delivered custom data mining and automation algorithms for **7+ international** clients, boosting efficiency by 10-25%
- Constructed a comprehensive database for South African election research analysis using Twitter data and news articles
- Designed and deployed a Twitter Bot for auto-replies and retweets, enhancing client social media engagement

RESEARCH EXPERIENCE

Binghamton University, Research Assistant for Human-Computer Interaction

Jan 2024 – Mar 2024

- Researched to enhance real-time ASL recognition by application of ML models, image processing, and Natural Language Processing
- Advanced models through Neural networks, and computer vision to tackle research challenges of hand recognition

Binghamton University, Research Assistant for Semiconductor Fabrication

Jan 2023 – May 2023

- Researched the feasibility of ML (Pytorch, CUDA) autoencoder for optimizing microscopic filter design in semiconductors
- Formulated an exploratory algorithm using Neural Networks to fabricate semiconductors

PROJECT EXPERIENCE

Analyzing Stock Market Sentiment and Price Movements, Team Project

Aug 2023 – Dec 2023

- Led full-stack development of predictive analytic model for social media sentiment and S&P stocks, using NLP and statistical method
- Managed real-time data collection from Reddit (6k posts, 80k comments) and Yahoo APIs using PostgreSQL
- Developed Flask dashboard with Matplotlib for sentiment visualization, utilizing Docker and Python on remote servers

Wear Your Weather, Independent Project ([link](#))

Oct 2020 – Jan 2021

- Created a dynamic web application delivering real-time weather forecasts using APIs and personalized outfit suggestions, enhancing user experience and daily decision-making. Tech stack utilized: Python, Flask, SQL, HTML/CSS, and JavaScript

Python and Arduino Integration, Independent Project ([link](#))

May 2020 – Jul 2020

- Integrated Arduino systems with Python via the Pyfirmata Library, enabling seamless communication between OS and Arduino
- Engineered a program for the Arduino to run continuously, optimizing board space without occupying memory

LEADERSHIP & HONORS

IEEE, Published Research Paper ‘Fake News Detection Using NLP and Logistic Regression’

Sept 2021

Hackathon (Secured 2nd out of the 49 teams) ([link](#))

Jul 2020

- Led a team of 3 to develop a ‘Fake News Detection Model using Machine Learning and Natural Language Processing’
- Built an ML model to classify news as "fake" or "real," providing users with verified news, enhancing digital media's reliability
- Integrated the model with a Twitter bot that replies to a tweet when tagged, a stand-out feature that elevated the project