

# HARSHIT SONI

New York, NY | (347) 818-5361 | hsoni1@binghamton.edu | LinkedIn (harshitsoni01) | harshitsoni.com | GitHub (harshitsoni01)

## EDUCATION

**Binghamton University, State University of New York**

May 2025

*Master of Science in Computer Science, Artificial Intelligence Track*

| **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, MATLAB, Shell

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

**Additional:** Automation, Data Analysis, Business Planning, NLP, Quantitative Analysis, Statistical Modeling, Selenium, LLM, RDMS

**Certifications:** SUNY Startup Summer School Class of 2024, JPMorgan Software Engineering Virtual Internship Experience, Mathematics for Machine Learning, Construct Stock Market Indices, Profit Analysis using Economic Value Added, Git for Developers

## PROFESSIONAL EXPERIENCE

**Koffman Southern Tier Incubator, Founder** | *Entrepreneur in Residence* | Binghamton, NY

May 2024 – Present

- Developed AI-driven storytelling solutions & eco-friendly consumer products, refining business strategies with industry leaders
- Selected for SUNY Startup School '24, completing formal training in business development, fundraising, and commercialization
- Led multiple early-stage ventures, focusing on digital marketing strategies, website optimization, and go-to-market planning
- Conducted in-depth market research and competitive analysis to validate business concepts and identify target customer segments

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

Jun 2024 – Sept 2024

- Developed and deployed machine learning models for career dashboards, leveraging AI, RAGs, and prompt engineering techniques to optimize sentence structure for dashboard renewable roles, improving data clarity for green labor seekers
- Led data labeling, web scraping, and cloud storage & tableau integration for government data, ensuring accuracy for analytics
- Collaborated with the **VP** on client projects, such as developing data pipelines (including for NREL) and Google Analytics dashboards

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

Jun 2023 – Aug 2023

- Engineered an automated email marketing system, reducing manual efforts and increasing client replies by **40%**
- Led data analysis on **10M** real estate transactions and **1.6M** agents using statistical modeling, machine learning models, and big data storage, delivering insights that directly improved decision-making and enhanced client relations
- Collaborated with the **Founder/CEO** to optimize business database and marketing workflows, increasing client satisfaction and expanding the sales team by **33.3%**. Tech stack: Python, AWS (S3, EC2), Django, PostgreSQL, APIs, OpenAI

**RoboticsForSure, Software Engineering Intern** | Remote

Jun 2021 – Aug 2021

- Implemented Robotics Process Automation (RPA) into ML and C++ projects, automating data preparation phases for clients and reducing processing time by **30%**
- 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 process 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's data-driven social media engagement strategy

## RESEARCH EXPERIENCE

**Binghamton University, Research Assistant – Human-Computer Interaction**

Jan 2024 – Mar 2024

- Researched ML model integrating computer vision and natural language processing (NLP) to translate real-time ASL into English text
- Improved hand recognition accuracy using neural networks and image processing for more precise ASL interpretation

**Binghamton University, Research Assistant – Semiconductor Fabrication**

Jan 2023 – May 2023

- Researched ML autoencoder models to develop microscopic filter designs for faster semiconductor manufacturing
  - Trained an adversarial model (PyTorch, CUDA) to predict and generate filter shapes based on light transmission across wavelengths
- IEEE, Published Research Paper Title '**Fake News Detection Using NLP and Logistic Regression**'

Sept 2021

## PROJECT EXPERIENCE

**Zillow Page Summarizer: Chrome Extension Assistant, Independent Project** ([link](#))

Jul 2024 – Sept 2024

- Developed an extension that summarizes Zillow property listings, providing real-time insights using web scraping and NLP
- Built the extension's backend using the Flask framework and integrated LLM for sentence processing and APIs for data retrieval
- Tech stack: LLM, Python, Flask, Pyzill, API, HTML/CSS, JavaScript, Chrome Extension Framework, Natural language Processing

**Analyzing Stock Market Sentiment and Price Movements, Team Project**

Aug 2023 – Dec 2023

- Built a full-stack predictive model using NLP and stats to analyze social media sentiment and S&P 500 price trends
- Collected Reddit & Yahoo (6k posts, 80k comments), visualized data via Flask dashboard with Matplotlib, deployed using Docker

## LEADERSHIP & HONORS

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

Jul 2020

- Led a team of three 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