

# SUJAY SHRIVASTAVA

45, Englewood Avenue, Buffalo, NY, US | (716) 617-1691 | [sujay1829@gmail.com](mailto:sujay1829@gmail.com) | [LinkedIn](#) | [Portfolio](#) | [GitHub](#) | [Kaggle](#)

---

## EDUCATION

**Master's in Engineering Sciences: Data Science**, University at Buffalo

January 2023 - June 2024

- GPA - 3.3/4.0
- Relevant course work: Statistics, Python Programming, Probability Theory, Numerical Mathematics, Data Intensive Computing, Data Models Query Language, Machine Learning

## TECHNICAL SKILLS

- **Programming Languages:** Python, R, MATLAB, SQL, C/C++, C#, Java, HTML, CSS, Javascript
- **Frameworks:** TensorFlow, Keras, OpenCV, PyTorch, scikit-learn, Orfeo Toolbox, Apache Hadoop, Apache Spark, DynamoDB, GDAL, NLTK, spaCy, Hugging Face, XGBoost, REST API
- **Tools:** GCP, Azure, Metabase, Mixpanel, Zoho, Notion, IBM Watson Studio, Docker, Excel, Microsoft Office, Power BI, Tableau, QGIS, DigitalOcean, Linux, Vertex AI

## RELEVANT EXPERIENCE

**Graduate Research Assistant | University at Buffalo, Buffalo, New York**

September 2023 - Present

- Furthered a novel Masked Auto Encoder with temporal embeddings of past 20 years for land cover analysis using multispectral satellite imagery resulting in 37% improvement in intersection over union
- Improved data pipeline efficiency by 5% by developing cloud-native solutions in Python/SQL to integrate and optimize research data processes and models
- Accelerated project progress by creating satellite image annotation and segmentation pipeline, saving \$3,000 and breaking through 6-month standstill

**Data Scientist Intern | Curelink, Gurugram, Delhi - NCR**

May 2022 - August 2022

- Reduced churn rate by 3% by developing Power BI dashboard for new customer retention department with recursive and nested SQL queries up to 8 levels deep
- Saved approx. \$0.5M by communicating insights to close non-profitable department and reallocating resources
- Developed a C++ script to initiate data synchronization between AWS and ZOHO CRM for on-site operations, ensuring double verification to minimize redundancy and enhance data accuracy by 25%

**Machine Learning Engineer Intern | IIT Kharagpur, Remote**

January 2021 - March 2021

- Improved predictions by leveraging sentiment analysis on customer satisfaction data with 50k records by 30%
- Boosted model accuracy by 15% doing A/B testing and driving stakeholder consensus on low-bias model and integrating it into backend with REST API enabling real-time predictions with 50ms average delay

## PROJECTS

**Hill Climb Racing AI Agent using RL:** Python, Pytorch, OpenAI Gymnasium

- Implemented object detection using YOLOv5, then developed a PyTorch-based reinforcement learning agent to autonomously play Hill Climb Racing in OpenAI gymnasium

**Spotify Analytics:** Python, HTML, SQL

- Created synthetic database, optimized queries by 20% using indexing, Built an ETL pipeline and Tableau dashboard

**US Accidents Prediction:** Python

- Addressed data imbalance using different sampling techniques to train model for 4-level accident severity classification on massive 8 million record dataset achieving 78% accuracy

## PUBLICATIONS

- Co-authored a paper titled "**Artificial Intelligence and Machine Learning in Marketing: A Review of Recent Advances and Future Trends**" in the European Chemical Bulletin

## EXTRA - CURRICULAR

- Led team to secure 3rd place in Blackstone Launchpad competition for leveraging AI and ML to counter natural calamities
- Elected as Social Media Coordinator in junior and senior year and Member of Cultural Organizing Committee in sophomore year and organized Fall Mixer event for 200 plus incoming freshmen

## CERTIFICATIONS

- **Microsoft Azure:** Azure AI Engineer Associate (AI - 102)