

Aryan Jain

New York City | (267) 315-9768 | aj3246@drexel.edu | <https://github.com/aryanj10> | <https://linkedin.com/in/aryanj10>

EDUCATION

MS Data Science <i>Drexel University Philadelphia, PA</i>	Sept 2023 – June 2025 GPA: 3.9
BS Computer Science with Business Minor <i>NMIMS University Mumbai, India</i>	July 2019 – May 2023

WORK EXPERIENCE

Data Science Intern MitraQSR Dallas, TX	Sept 2024 – Dec 2024
<ul style="list-style-type: none">• Extracted P&L data for 178 stores from R365 and built dashboards to compare revenue, costs, and EBITDA across stores• Built a RAG tool with FAISS + LLM API to answer queries like Which Region X store had highest revenue in Period Y?• Identified \$12K+ EBITDA gaps from high maintenance, overtime, helping execs focus on underperforming locations	
Data Analyst Intern Barrier Breakers Inc. Princeton, NJ	June 2024 – Sept 2024
<ul style="list-style-type: none">• Migrated 1000+ student records from Trello, Evernote, and Zoho to Monday.com with 99% data accuracy• Built automation scripts to trigger workflows from form fills, cutting manual task handling by 48%	
Data Scientist, Junior Associate Celebal Technologies India	Jan 2021 – Sept 2022
<ul style="list-style-type: none">• Analyzed prediction gaps in revenue and headcount models for 20M+ US/Canada business records; developed industry-specific XGBoost solutions to improve accuracy in weak-performing sectors• Presented EDA and modeling insights to the client, highlighting where the model's performance dropped across industries• Led P&L dashboard build; developed forecasting APIs (SARIMA, Prophet) and deployed on Azure for BI integration	

PROJECTS

LLM-Powered Payroll Migration Tool Personal Project	July 2025
<ul style="list-style-type: none">• Parsed 1000+ payroll PDFs into structured JSON using Gemini LLM; saved 10+ weeks of manual data entry• Built Playwright bot to automate payroll uploads to Accountant World's non-API portal across 85+ clients• Saved \$20K+ in labor costs; generated \$2K+ in revenue via tool reuse for new client payroll migrations	
CW-RIS: Child Wind Risk Intelligence System UN Tech Over Hackathon	June 2025
<ul style="list-style-type: none">• Built Streamlit app that overlays ERA5 wind data with child population exposure to map high-risk zones• Computed zonal statistics and risk scores using raster and vector geospatial data for humanitarian use	
Comcast AI Challenge Winner Philly Codefest 2025	Mar 2025
<ul style="list-style-type: none">• Created a generative AI-based animated video using tools like META AI, Runway ML for character and scene generation• Integrated GPT-4 and ElevenLabs for script generation and voice synthesis, creating expressive and human-like dialogues	
High-Speed Inference with Triton and PyTorch Drexel University	Jan 2025 – June 2025
<ul style="list-style-type: none">• Deployed ResNet-based classifier on NVIDIA Triton Inference Server to batch infer 10,000 images in 5s per request• Optimized PyTorch training pipeline using CUDA kernels, cutting training time by 40% on A100 GPUs	

LEADERSHIP

Teaching Assistant Data Science Programming Drexel University	Mar 2025 – June 2025
<ul style="list-style-type: none">• Assisted instruction for 70+ students in Python, Linux, Github, and efficient data processing with NumPy and Pandas• Designed and graded coding tasks on NumPy, data wrangling, file I/O, and high-performance computing techniques	
Course Assistant Human-AI Interaction Drexel University	Dec 2024 – Mar 2025
<ul style="list-style-type: none">• Guided 30+ students on ethical AI frameworks, algorithmic fairness, explainable AI, and responsible data ownership• Provided project mentorship and feedback aligned with Responsible AI curriculum	

CERTIFICATIONS

Azure Data Scientist Associate Microsoft	May 2022
Azure AI Engineer Associate Microsoft	Mar 2022

TECHNICAL SKILLS

Programming: Python, R, SQL, MATLAB, JavaScript, C/C++, HTML, CSS

ML and AI: Scikit-learn, TensorFlow, PyTorch, XGBoost, LightGBM, OpenCV, SpaCy, BERT, Langchain, Time Series

Cloud and DevOps: AWS (S3, RDS, EC2), Azure, Databricks, PySpark, MySQL, Docker, MLflow, CI/CD Pipelines

Visualization and APIs: Tableau, Power BI, Plotly, Dash, Matplotlib, Seaborn, Streamlit, Flask, FastAPI, RESTful APIs