Pranjal Jain

(959) 248-9670 | pranjal.jain@uconn.edu | LinkedIn | GitHub | Interactive Resume

EDUCATION

University of Connecticut, Graduate School of Business (UCONN) Hartford, CT | Open to Relocation
Master of Science (M.S.) in Business Analytics and Project Management ***May 2025***
(**Relevant Coursework**: Data Science, Predictive Analytics, Statistics, Quantitative Analysis, Data Engineering, Big Data Management)

WORK EXPERIENCE

***International Medical Crisis Response Alliance* (IMCRA)** Norwalk, Connecticut
Data Science Associate ***January 2025 – Present***

- **Spearheaded** the development of an automated data pipeline using Azure Data Factory and Databricks to process 5GB+ of oceanic research data for the SIDS 4 project, **reducing data ingestion time by 40%**.
- **Improved data quality** by implementing schema validation and deduplication in PySpark, which **cut manual cleanup efforts by 60%**.
- **Delivered actionable insights** for marine policy decisions by running optimized SQL queries in Synapse Analytics, enhancing **data retrieval efficiency by 30%**.
- **Built stakeholder confidence** through dynamic dashboards and reports, driving engagement among researchers and international delegates.
- **Secured project funding** by leading fundraising initiatives and representing IMCRA in diplomatic discussions, expanding visibility and resource backing for ocean data initiatives.

PepsiCo. Valhalla, New York
Graduate Student Consultant ***August 2024 - December 2024***

- Developed a predictive modeling framework leveraging ML models to forecast **PET bottle performance metrics**, enabling data -driven optimization of PepsiCo's blow molding processes, to save cost and reduce material waste.
- Designed **interactive Tableau Stories** to synthesize EDA insights and predictive analytics, providing real-time visibility into key manufacturing inefficiencies and enhancing data-driven decision-making for the packaging team.

Tata Consultancy Services Indore, India
Systems Engineer ***April 2021 - June 2023***

- Developed and implemented a test utility on Jenkins for **CI/CD**, reducing infrastructure costs by **$100k**. Utilized data visualization techniques to present cost-saving metrics (Matplotlib, Seaborn).
- Automated report generation via **Jenkins**, tracking application performance metrics and **saving 85% of reporting time**. Applied statistical analysis to interpret performance data (Python).
- Conducted knowledge transfer sessions for **40+** cross-functional teams on testing methodologies, script optimization, and performance improvement (Leadership, Organizational skills).
- Led **3 Proof of Concept** initiatives to enhance data privacy and quality using TCS Mastercraft DataPlus and SQL queries, managing **a team of 5** (Data Cleaning and Wrangling).

LEADERSHIP & ACTIVITIES

Modlee Artificial Intelligence and Machine Learning Society– Treasurer ***October 2024 - Present***

- Society Created in Partnership with Modlee.ai an online Machine learning and AI Platform
- Facilitated Weekly Hackathons for the University of Connecticut Student body.
- Secured 3rd place in a club-wide machine learning hackathon focused on image classification
- Organizing Training Sessions in Machine Learning and AI with Modlee Leadership Team (CEO and Co-founders).

ACADEMIC PROJECT EXPERIENCE

Healthcare Insurance & Claims Data Analysis | SQL, Lucid Chart ***December 2024***

- Engineered a healthcare claims database implementing an ERD, BeautifulSoup for data extraction, and PostgreSQL with DDL/DML scripts, which improved claim processing speed by 40% and data accuracy by 15%.
- Developed advanced SQL queries and analytical reports to derive insights on claim approval trends, insurance coverage patterns, and financial assistance eligibility, while optimizing performance through **effective indexing and query tuning.**

Optimizing Risk Prediction in Car Insurance Claims | Python ***December 2023***

- Cleaned a substantial dataset of more than 100k columns and preprocessed data using and feature engineering, encoding and outlier detection to eliminate 6 non-predictive columns and refine data with 20 final columns.
- Employed models such as linear regression, XG boost, random forest and gradient boosted tree (GBT), with GBT model at 83% accuracy, showing a significant improvement of 10% over other machine learning models.

Statistical Nutritional Analysis | *R programming (hypothesis testing, T-test, Z-test)* ***December 2023***

- Analyzed 55,000+ data points from 50+ cereal products using R (linear regression, T-test, Z-test) to predict calorie content. Achieved 80% model accuracy and visualized findings in a Tableau dashboard viewed by 200+ users, influencing healthier food decisions.

RELEVANT SKILLS

Data Analysis & Visualization: Tableau, Power BI, Excel (Pivot Tables, VLOOKUPs, Graphing), Visio, Lucid Chart
Programming & Scripting: SQL, Python(NumPy, Pytorch, TensorFlow, Pandas), R, PostgreSQL, PL/SQL, Hadoop
Data Management & Processing: Data Wrangling, Data Cleansing, ADLS Gen2
Cloud & Tools: AWS, Azure, JIRA, Git, Google Cloud, Azure Databricks, Synapse Analytics
Business Acumen: Leadership, Organizational skills, Cross Functional Strategic Planning
Certifications: SQL ,Intermediate SQL, Python Data structures, JIRA fundamentals, Agile