

# RIJUL H. SHERATHIA

rsherathia@ucsd.edu | San Diego, California | +1 (619) 953-7403 | [linkedin.com/in/rijul-sherathia](https://www.linkedin.com/in/rijul-sherathia)

## EDUCATION

### UNIVERSITY OF CALIFORNIA, SAN DIEGO

Master of Science in Data Science, GPA: 3.9/4.0

September 2022 - December 2023 (Expected)

- Coursework: Statistical Models, Machine Learning, Deep Learning, Scalable Data Systems, Big Data Analytics & Applications, Advanced Data Mining & Visualization, Data Ethics & Fairness, Advanced Optimization, Fraud Analytics, Computer Vision.

### MIT WORLD PEACE UNIVERSITY, INDIA

Bachelor of Science in Computer Science and Engineering, GPA: 9.21/10

July 2017 - June 2021

- Coursework: Data Science, Database Management, Data Warehousing, Data Structures, Python Programming, Object Oriented Programming, Design & Analysis of Algorithms, NLP, Advanced Java, Software Modelling and Design, Business Intelligence.
- Certifications: Data Science A-Z (Udemy), Machine Learning A-Z (Udemy), Artificial Intelligence (NPTEL)

## WORK EXPERIENCE

### NATIONAL CENTER FOR MICROSCOPY AND IMAGING RESEARCH (NCMIR), UC SAN DIEGO, CA

Research AI/ML Scientist

June 2023 – Present

- Implementing cutting-edge image segmentation techniques (**U-Net, Mask R-CNN, FPN**) on microscopy brain cells for precise analysis.
- Explored novel unsupervised learning approaches to uncover brain cell patterns, resulting in valuable insights for neuroscientific research.

### INSTITUTE FOR NETWORK MEDICINE (iNetMed), UC SAN DIEGO, CA

Research Data Scientist

June 2023 – Present

- Collaborated with researchers to translate genomics findings into actionable insights for precision medicine.
- Leveraging **statistical modeling** and **machine learning** to unveil patterns and facilitate breakthroughs in genomics research.

### ZS ASSOCIATES, PUNE, INDIA

Business Technology Solution Associate

July 2021 – July 2022

- Engaged with BMS healthcare client to develop cutting-edge **ML techniques** and **statistical analytics algorithms**, resulting in **optimized operational workflow, data processing**, and cost savings of **\$1 million** in client revenue.
- Liaised with **10+ internal & client teams** to aid in ML-driven strategic solutions and performance, streamlining various business challenges using SQL, Python, and AWS Services, **yielding 20% sales improvements**.
- Worked on the enhancement of **AWS Redshift** to cut space utilization and enhance performance thereby **reducing cost by 15%**.
- Mentored 4 interns** in ETL job enhancements and familiarized them with orchestration and standard ETL procedures.
- Orchestrated 50+ analytical/ETL jobs on SnapLogic, with custom features like **task retry** and **anomaly detection** based on duration.

Business Technology Solution Associate – Intern

November 2020 – June 2021

- Developed **generic Python script** automating file movement from Email to AWS S3, **reducing 20+ hours** of weekly manual effort.
- Constructed weekly operational **Tableau** dashboard analyzing meal expenses on Redshift views, **reducing expense overhead by 25%**.
- Deployed **15+ ETL pipelines** with Data Architecture engaging AWS Services, SnapLogic, Python, and SQL for various data sources.
- Acquired understanding of the **business healthcare domain** with the ability to incorporate and translate into technical design.

## TECHNICAL SKILLS

|                             |  |
|-----------------------------|--|
| Programming Languages       | : Python(Pandas, NumPy, Sklearn, PyTorch), SQL, R, MATLAB, Java, XML, Shell Scripting              |
| Databases                   | : Amazon Redshift (PostgreSQL), MySQL, NoSQL, MongoDB, Neo4j, Redis, Apache Cassandra              |
| ML/DL Application           | : Predictive Modeling, Classification, Clustering, Statistical Analysis, CV, NN Architectures, NLP |
| Data Science Application    | : PySpark, MapReduce, Hypothesis Testing, Tableau, Power BI, Advanced Excel                        |
| Cloud & Scheduler Platforms | : ETL, AWS (RDS, Redshift, S3, Lambda, EC2), Azure, Erwin Modelling, SnapLogic, GitHub, SVN        |

## RELEVANT PROJECTS

### DynamicVision: Real-Time Image-to-Text Generation (Ongoing)

July 2023

CNN, PyTorch, Transformer, Multi-head Attention, Word Embedding, Positional Encoding, Resnet

- Developing a Generative Transformer model using Attention, to train Flickr8k images to produce descriptive and accurate text.
- Empowering users with a seamless solution for instant image-to-text transcription, enhancing accessibility and information retrieval.

### American Sign Language Detection and Recognition

May - June 2023

Python, CNN, TensorFlow, OpenCV, Predictive Modeling, Classification Algorithm

- Performed data cleaning and preprocessing techniques (skin masking and canny edge detection) to enhance model performance.
- Implemented classification and CNN models, achieving an accuracy of 97%, and successfully integrated real-time prediction capabilities.

### Credit Card Fraud Detection

May - June 2023

Python, Data Visualization, Logistic Regression, SVM, K-Nearest Neighbors, Decision Tree, Random Forest, XGBoost, CatBoost, LGBM, MLP

- Conducted extensive **data cleaning, variable creation, and feature selection** techniques to train and explore various models.
- Successfully implemented selected model that resulted in savings of \$21 million/year and achieved an FDR@3% of 56.98% for the project.