# RAHUL SANDIP DESHMUKH

602-200-4347 • desrahul.200@gmail.com • www.linkedin.com/in/desrahul200 • desrahul200.github.io

## EDUCATION

**Master of Science, Computer Science** — Graduating May 2025
Arizona State University, Tempe, AZ — 3.90 GPA

**Bachelor of Technology, Information Technology** — May 2023
K.J. Somaiya Institute of Technology, University of Mumbai, India — 3.83 GPA

Relevant coursework: Data Structures and Algorithms, Statistical Machine Learning, Planning Learning methods in AI, NLP, Data-Intensive Systems for Machine Learning, Frontier Topics in GenAI, Computing for Data-Driven Optimization, Big Data Analytics, Database Management, Data Mining and Business Intelligence, Engineering Blockchain

## TECHNICAL SKILLS

**Programming Languages & Web Development:** Python, Java, C++, JavaScript, SQL, Flask
**Data Management & Big Data Technologies:** Google BigQuery, MySQL, MongoDB, DBeaver, Hadoop
**Cloud Platforms & DevOps Tools:** AWS, Google Cloud Platform, Docker, CI/CD, Git, Unix/Linux
**Machine Learning Libraries & Data Analysis:** PyTorch, TensorFlow, Keras, scikit-learn, NumPy, Pandas, BERTopic, FastAPI, Matplotlib, Seaborn, Excel, Tableau, Power BI

## EXPERIENCE

**Data Scientist – Student Assistant: ASU Enterprise Partners, Scottsdale, AZ** — January 2024 – Present
- Developed and deployed **first-touch, last-touch, and data-driven attribution models** in **Python** for **Sun Devil Athletics**, using **GA4** and **BigQuery** to identify top-performing acquisition channels and inform marketing spend strategies.
- Built and evaluated **multi-stage attribution models** for **ASU Learning Enterprise**, mapping user journeys from sign-up to purchase to optimize full-funnel performance and drive ROI-focused media investment.
- Designed a **K-means segmentation pipeline** in **R** with **Mahalanobis distance** to cluster 100K+ alumni into 5 affinity groups, enabling targeted engagement and improving outreach efficiency.
- Developed a **donor propensity model** using **Bayesian optimization**, with end-to-end **ETL**, **EDA**, and hyperparameter tuning, improving predictive accuracy by **10%** and shaping strategic fundraising efforts.

## ACADEMIC PROJECTS

**LLM-Powered Sentiment & Topic Mining – Yelp Reviews Analysis** — Spring 2025
- Fine-tuned transformer-based models (**BERT, RoBERTa**) for multi-class sentiment classification, achieving **94.2% accuracy** on 500K+ real-world Yelp reviews.
- Integrated **BERTopic** with **SentenceTransformer embeddings** to improve topic coherence, resolving cluster overlap seen in traditional KMeans-based approaches.
- Deployed the pipeline as a **RESTful API using FastAPI**, enabling real-time sentiment analysis on user-submitted reviews.

**3D VAE Developer – Vermilion: ASU's In-house Text-to-Video Generative Model** — Fall 2024
- Optimized the **Variational Autoencoder (VAE)** in the **CogVideoX pipeline**, reducing GPU memory usage by **30%** via **dynamic tiling** and **gradient checkpointing**.
- Developed **Optical Flow Consistency (OFC)** and **SSIM** metrics to enhance temporal coherence and structural fidelity in video sequences.
- Designed and tested **10+ complex prompts** to improve **spatial fidelity** and **temporal transitions**, refining model robustness.

**Harnessing Deep Reinforcement Learning for Autonomous Driving in CARLA** — Fall 2023
- Designed an adaptive **autonomous driving agent** using **Deep Reinforcement Learning** in CARLA simulation environment.
- Employed a **Variational Autoencoder (VAE)** for feature extraction and **Proximal Policy Optimization (PPO)** for decision-making, achieving a mean reward of **92.3%** in dry conditions and **78.5%** in adverse scenarios.
- Optimized model performance by refining reward functions and limiting training to 3 million steps, demonstrating **lane-keeping**, **obstacle avoidance**, and **precise navigation** across diverse scenarios.

**Identification and Classification of Plant Leaf Diseases** — Spring 2023
- Implemented **Generative Adversarial Networks (GANs)** to augment the plant leaf dataset, enhancing model performance by increasing data diversity. Used Labelimg for precise annotation, ensuring accurate **model training**.
- Developed a **YOLO v4-tiny** model for real-time detection of diseased leaves, achieving **77.0% Mean Average Precision (mAP)**. Created a mobile-based detection app in **Java** and hosted the model on a **Flask** website for instant identification.

## PUBLICATION

- Deshmukh, R., Mayekar, V., Patel, S., & Rathod, M. (2023). Identification and Classification of Plant Leaf Diseases using YOLOv4-tiny Algorithm. 2023 6th International Conference on Advances in Science and Technology (ICAST), 352–357.