Samay Shetty

+1 585 230 7406 • samayshetty7103@gmail.com • linkedin.com/in/samay-shetty • https://samay-shettyyyy.github.io/portfolio/

Education:

Rochester Institute of Technology, Rochester, New York **August 2024 - Present**
Master's in computer science

- 3.33 / 4 GPA

Vidyalankar Institute of Technology, (University of Mumbai), Mumbai, India. **August 2020 – May 2024**
Bachelor's in engineering (Electronics and Telecommunication)
Minor in Data Science

- 9.45 / 10 CGPA
- Techno-Xian- World Robotic Championship 2023-24 - 2nd Runners Up- July - (Team Lead)

Skills:

Languages: Python, SQL, HTML/CSS, Java, C, C++
Tools: MATLAB, CAD, PolsarPro, Fusion360, Google Cloud
Technologies: OpenCV, NumPy, Pandas, PyTorch, TensorFlow, Sci-kit, PowerBI, SQL, MongoDB
Expertise: Machine Learning (Deep Learning, Computer Vision, Generative AI, Large Language Models, LangChain, Retrieval-Augmented Generation (RAG)), Data Science, Natural Language Processing (NLP), Internet of Things (IoT), Big Data.
Soft Skills: Leadership and Team Management, Effective Communication and Collaboration, Problem-Solving, Project Management, Time Management.

Research and Publications:

- Research and Development of a Dual Port solar charger and published a research paper titled " *Solar Charger: A Green Way of Synthesizing & Using Energy*" in the International Journal of Scientific & Engineering Research (IJSER) Volume 12, Issue 11, November 2021 Edition.
- Review paper on One hot encoding for diabetes prediction published in TANZ Journal titled "*DIABETES PREDICTION USING MACHINE LEARNING: ENHANCING ACCURACY USING HOT ENCODING TECHNIQUE* "Volume 18, Issue 12, December 2023
- Research paper titled *"PolSAR Hub: A Large-Scale Repository for PolSAR Data"* selected for presentation in IEEE-INGARSS-2024 held in Goa, India

Experience:

Research Assistant (Vidyalankar Institute of Technology) August 2023 - December 2023

Python | TensorFlow, Sci-Kit, NumPy, Pandas
Developed the backend for a machine learning model predicting diabetes using Random Forest, implementing optimized one-hot encoding for improved accuracy. Designed and refined data processing pipelines, ensuring efficient model training and deployment. Contributed to research published in TANZ Journal, applying AI to enhance healthcare predictions.

Projects:

BizzRizz: - An Agentic AI Based solution for Restaurant Consultation. **https://bizrizzai.onrender.com**
Python | Gemini, Google Place API, Paychex API | HTML, CSS, JS | Git, Render

- Developed an Agentic AI for restaurant owners, providing competitor analysis, customer feedback insights, and business recommendations.
- Integrated Google Places API for competitor data and reviews, Gemini AI for NLP-based insights, and Paychex API for financial analysis.
- Built the frontend (HTML, CSS, JS) and backend (Flask); deployed on Render for scalability.

MatSAR - ML Based tool for PolSAR data classification. **https://tinyurl.com/4h3s6jaw.**
C++ | MATLAB, MATLAB App Designer

- Led the design and development of MatSAR, an ML-driven tool for PolSAR data classification, improving upon PolSARPro.
- Implemented ANN, SVM, Decision Trees, Random Forest, and XGBoost for advanced classification on decomposed PolSAR data.
- Utilized MATLAB for computational processing and developed a user-friendly GUI with MATLAB App Designer.

Gesture AI – Motion and Object Captioning via Hand Gestures **https://github.com/Samay-Shettyyyy/Video_captioning.git**
Python | OpenCV | Media Pipe | Hugging Face Transformers (BLIP) | NumPy | Pillow

- Designed a vision-based system that detects hand gestures to identify objects and generate dynamic descriptions using generative AI models.
- Integrated advanced computer vision techniques with Transformer-based models to produce real-time, context-aware captions.
- Enhanced user interaction by combining motion detection with natural language generation, ensuring high adaptability across diverse scenarios

Amazon Packaging Size Prediction
Data Cleaning | NLP

- Developed an ML model predicting optimal packaging sizes by training over 200k data points using Random Forest.
- Utilized NLP techniques to extract and process product specifications, improving prediction accuracy significantly.

Additional information

- Google Cloud Certification
- Founder and Head of the 3D Printing Club at Vidyalankar Institute of Technology, Mumbai, India