

Sanskruti Raut

Portfolio: sanskrutiraut14.wixsite.com/mysite

Github: github.com/sanskruti-raut

Email: sanskrut@usc.edu

Mobile: +1 (213)-913-7696

LinkedIn: <https://www.linkedin.com/in/sanskruti-raut/>

EDUCATION

- **University of Southern California, Los Angeles, USA** Los Angeles, USA
Master of Science, Electrical Engineering (Machine Learning and Data Science) Aug 2023 - May 2025
Courses: Linear Algebra, Probability, Deep Learning, Machine Learning, Natural Language Processing, Cloud Computing
- **Maharashtra Institute of Technology** Pune, India
Bachelor of Technology, Electronics and Communication Engineering Aug 2017 - July 2021

SKILLS

Languages: Python, MATLAB, JAVA, SQL, C++, JavaScript

Libraries/Frameworks: Pandas, Sklearn, Matplotlib, Seaborn, PyTorch, NumPy, SciPy, Scikit-learn, Tensorflow, Selenium, MONAI, NiLearn, spaCy, BeautifulSoup, Node.js, Express.js, LangChain, ChromaDB, Pinecone

Tools: Eclipse, Jupyter Notebook, VSCode, JIRA, JAMA, ServiceNow, AWS (SageMaker, EC2, S3), Git, Docker, IntelliJ, Postman

EXPERIENCE

- **Lee Lab, Keck School of Medicine, USC (Research)** Los Angeles, USA
Natural Language Processing Researcher - Advisor: Prof. Dr. Hayoun Lee Jan 2025 - Present
 - Implemented a Medical Retrieval-Augmented Generation (RAG) MVP to analyze HIV guidelines and research, generating regimen recommendations for Physicians.
 - Developed a LangChain-based system to convert complex medical PDFs into structured text and integrated Chroma DB for retrieval.
 - Leveraged Qwen-2.5-7B for LLM and achieved accuracy of 85.7%.
- **ViyaMD (Internship)** Los Angeles, USA
Machine Learning Engineer May 2024-July 2024
 - Conducted a Virtual Product Review (VPR) on multiple Data Ingestion platforms (Azure Data Explorer, Adobe Experience Platform Data Ingestion, PDFMiner, LangChain data loaders, etc.) for Medical RAG use-case.
 - Compared extracted PDFs from the VPR with ground-truth PDFs by designing a custom Data ingestion-evaluation pipeline.
 - Evaluated pipeline accuracy by measuring precision (92.3%), recall (89%), and F1(90.12%) scores to quantify extraction and conversion reliability.
 - Visualized the above performance analysis using CometML and found Adobe Experience Platform Data Ingestion to be the best.
- **Biomedical Imaging Group, USC (Research)** Los Angeles, USA
Machine Learning Researcher - Advisor: Prof. Dr. Anand Joshi Jan 2024 - April 2024
 - **Project1:** Adapted DeepBet v1.0 tool with U-Net architecture for macaque MRI dataset achieving high precision (99%) for skull-stripping.
 - **Project2:** Trained a PyTorch+MONAI-based SwinUNETR pipeline on NVIDIA A100 GPU, reaching an 80% Dice score for macaque brains.
- **Deloitte USI (Full Time)** Mumbai, India
Software Engineer (Analyst) - State of North Dakota, Health and Human Services Sept 2021-Jun 2023
 - Created and implemented 150+ automation scripts using Keyword and Data-driven frameworks with Java, Selenium, TestNG, Jenkins, and Maven to test AngularJS based web application with 2026 web pages and performing backend validations with SQL.
 - Spearheaded the automation of the regression suite for SIT sign-offs and identifying 200+ defects and reducing manual effort by 60%.
 - Collaborated cross-functionally by managing project workflows with JIRA and JAMA, adhering to Agile methodologies and SDLC.
 - Recognized with a '**Spot Award**' for outstanding automation expertise and contributing to 3 major releases.

PROJECTS

- **Code Detective:** Fine-tuned GraphCodeBERT on NVIDIA A40 for detecting semantically similar Python snippets (F1: 0.96) (Dec'25)
- **Transfer Learning for Image Classification:** Built a multi-class classifier (ResNet50/101, EfficientNetB0, VGG16), achieving 91.21% F1 and 99.19% AUC with EfficientNetB0. (Dec'25)
- **Laptop Price Prediction using Machine Learning Algorithms:** Compared Linear Regression, Random Forest, Support Vector Regression, K-nearest neighbors, and Neural Networks, scoring $R^2 = 0.844$ with Random Forest.(May'24)
- **Computer Vision Super Resolution using SRGAN:** Achieved 3x enhancement on DIV2K images with SRGAN (Adversarial Loss: 0.0010, VGG Loss: 0.0014, Pixel Loss: 0.0625) (Dec'23)

PUBLICATIONS

- Raut, S., et al., "Real Estate Based Recommender System Using ML", GIS Science Journal, 2022. [Paper]
- Raut, S., Motade, S., "IOT Based Smart Irrigation System using Cisco Packet Tracer", IJCSE, 2021. [Paper]
- Raut, S., Naware, S., Tank, V., "Vehicle Cluster Development", ICTCS, 2021. [Paper]