Lakshmi Swarna Durga Nallam

(+1) 7133596070 | sn103@rice.edu | Swarna Nallam

Education

Rice University — Houston, TX
Masters in Computer Science 3.79 GPA — *August 2024 – December 2025*
- *Coursework:* Information Retrieval, Machine Learning, Introduction to Deep Learning, Artificial Intelligence (AI)

SRM University — Andhra Pradesh, India
Bachelors in Computer Science(Specialization: Big Data Analytics) 8.55 GPA — *August 2019 – May 2023*
- *Coursework:* Big Data, Applied Data Science, Data Warehousing and Mining, Applied Mathematics

Experience

FullStack Developer - Summer Engineering Innovation Program — June 2025 – Present
Rice University — *Houston, TX*
- Founded dmv-agent.ai, an AI-driven web platform automating vehicle registration for Houston DPS, reducing manual processing time from over 2 hours to less than 10 minutes per transaction.
- Implemented OCR pipelines using TrOCR and Phi Mini LLM to extract and structure driver's license and vehicle title data into JSON with 95% accuracy reducing manual intervention.
- Built responsive frontend interfaces using HTML, CSS, TypeScript, JavaScript, and React.js, and designed scalable backend services integrating MongoDB for efficient data storage and retrieval.

Teaching Assistant – Natural Language Processing (NLP) — Jan 2025 – May 2025
Rice University — *Houston, TX*
- Assisted in delivering a graduate-level NLP course to 30 students, covering probabilistic models and neural representations.
- Held weekly office hours, mentoring students on sentiment analysis, chatbot design, and information extraction.
- Designed and implemented detailed grading rubrics for all the assignments, delivering targeted feedback that boosted project scores.

AEM Developer — July 2023 – July 2024
Barclays — *Chennai, India*
- Developed and configured scalable Adobe Experience Manager (AEM 6.5), including creation of custom components, templates, and workflows, leveraging Java (Sling Models, OSGi Bundles), HTML, JavaScript, and web services for 10+ US co-branded credit card marketing pages.
- Collaborated within an agile team alongside UX designers, QA analysts, DevOps engineers, business stakeholders and CI/CD practices (Jenkins, Git, Maven).
- Conducted 50+ detailed code reviews, significantly enhancing code quality, to reduce prodcution defects, managed Digital Asset Management (DAM) for streamlined asset storage and reuse.

Advanced Data Science Intern — March 2022 – August 2022
APSSDC, Indo-Euro Synchronization — *Remote / India-Germany Collaboration*
- Analyzed customer transactions and behavior using pandas, and scikit-learn for segmentation, KPI tracking, and marketing strategy.
- Modeled Customer Lifetime Value (CLTV) and churn risk using survival analysis and logistic regression to identify high-value segments.
- Designed frequency-based rewards and next-visit predictors with SQL and ARIMA to enhance loyalty and store-level planning.

Student Researcher — June 2021 – July 2021
SRM University, Prof. Hiren Kumar Thakkar — *Andhra Pradesh, India*
- Analyzed COVID-19 epidemiological data comparing India's first and second waves; observed daily cases increased significantly from 30,000 (first wave) to 120,000 (second wave) due to the Delta variant.
- Created a multivariate linear regression models to evaluate how population density and policy measures influenced case fatality rates, finding a reduced fatality rate of 1% (second wave) compared to 4% (first wave).
- Presented key insights with visualizations and co-authored a research paper published in Springer.

Projects

Travel Chatbot with RAG Framework | *Python, LLM, SentenceTransformers* — February 2025 – April 2025
- Developed a dense semantic retrieval system for more than 30,000 UNESCO heritage site documents and travel attractions, improving query accuracy by 92%.
- Upgraded a 4.3 billion-parameter LLM (Phi-3.5-4B) model integrated with optimized FAISS vector search, resulting in query response times under 1.2 seconds and an embedding storage reduction of 78% (from 12GB to 2.7GB).
- Achieved 89% top-3 retrieval accuracy on paraphrased queries, outperforming TF-IDF baselines by 41%.

High-Resolution Colorization Using GANs | *Python, DeOldify, ESRGAN* — January 2025 – March 2025
- Combined DeOldify (for semantic colorization) and fine-tuned Real-ESRGAN (for 4× super-resolution) to deliver grayscale-to-4K colorization with sharp, vivid results.
- Fine-tuned Real-ESRGAN on the DIV2K dataset using perceptual loss, L1 loss, and adversarial loss with a Relativistic average GAN (RaGAN) discriminator.
- Enhanced model performance to 28.5 PSNR / 0.89 SSIM.

Multi-Class URL Classification with Bloom Filters | *Python, TensorFlow, XGBoost* — September 2024 – December 2024
- Designed and implemented a hybrid detection system combining logistic regression, LSTM classifiers, and dual-layer Bloom filters to identify malicious URLs with high precision in real-time traffic.
- Reduced false positives by 15% through a dual-layer Bloom filter augmented with logistic regression for early URL classification on Kaggle DMOZ dataset.
- Implemented neural hashing to dynamically optimize Bloom Filter mappings.
- Leveraged XGBoost models as secondary classifiers for ambiguous URLs, boosting overall precision to 94% on the Kaggle DMOZ dataset.

Technical Skills

Programming Languages: Python, R, SQL, C/C++, Java, JavaScript, TypeScript, HTML/CSS, Ruby
ML and Data: pandas, NumPy, scikit-learn, TensorFlow, PyTorch, Transformers, Matplotlib, Tableau
Platforms: Hadoop, Hive, Google Cloud Platform
Frameworks and Tools: React, Node.js, Rails, AEM 6.5, Git, Docker, VS Code, IntelliJ, JUnit, WordPress