SRIVALLI LANKA

Email: sslanka2@illinois.edu | Tel: +1(331)-289-9444 | linkedin.com/in/srivallilanka

EDUCATION

University of Illinois at Urbana-Champaign, *Master's in Information Management* | IL, USA   **GPA: 4.0/4.0**   Aug 2024 – May 2026

Jawaharlal Nehru Tech University, *Bachelor's in Electronics & Telematics* | Telangana, India   **GPA: 8.77/10**   Jul 2018 – Jul 2022

Relevant Coursework: Data Warehousing & Business Intelligence, Data & Statistical Models, Machine Learning in Cloud, Information Modelling, Information Consulting, Advanced Business Research, Introduction to Generative AI

WORK EXPERIENCE

Research Assistant, Data Science – RailTEC (Rail Transportation and Engineering Center)   May 2025 – Present

- Developed statistical models of freight derailment risk using **Poisson** and **Negative Binomial regression**, enabling **quantitative safety diagnostics** and data-driven prioritization of infrastructure investments across **regional rail systems**.
- Designed a **text-mining workflow** for FRA accident narratives (tokenization, TF–IDF, topic modeling, early severity classification), turning unstructured text into **structured features** that **enhance quantitative risk models**.
- Gained hands-on experience with large-scale, university-led transportation research by developing **analytical workflows** that support **system-level risk analysis and decision support** for **large-scale infrastructure planning** across **North America.**

Data Analyst - CHIME-in, University of Illinois   Feb 2025 – May 2026

Extreme Weather and Health Project (Illinois Department of Public Health)

- Engineered a statewide county-day heat exposure dataset (Air Temperature, Humidex, Heat Index) for all Illinois counties, building **reproducible Pandas pipelines** that form a foundational data asset for **public-health** and **epidemiological analysis**.
- Conducted **time-series correlation** and **lag analysis** linking extreme heat events to health outcomes (e.g., hospital admissions), delivering analysis-ready datasets that **support predictive modeling** and **heat-risk assessment** by **state health researchers**.

Analyst - Deloitte   Sep 2022 – Aug 2024

- Analyzed network performance data using SD-WAN telemetry and statistical summaries to ensure **99.9%** uptime across **300+** sites; automated documentation with **Python** and **Power Query**, improving reporting accuracy by **40%**.
- Built and maintained **Power BI dashboards** to track **SLA** compliance and detect anomalies in real-time; co-designed alert systems, cutting ticket backlog by **89%** and accelerating response time.

TECHNICAL SKILLS

Programming & Data Tools: Python (Pandas, NumPy, Seaborn, Scikit-Learning), R (Tidy verse, Caret, Metrics), SQL, MS Excel, Git

Data Engineering & Cloud: AWS (S3, EC2, Glue, Lambda, Sagemaker, Athena, QuickSight), Apache Airflow, REST APIs, ETL

Visualization & Analytics: Power BI, Tableau, QuickSight, ggplot2, Matplotlib

Statistics & Machine Learning: Poisson, Negative Binomial, Linear, Logistic Regression; Random Forest, EDA, NLP

Certifications: Azure (AZ-900), ISC² (CC), Data Wrangling in Excel Power Query, AWS Data Engineer Associate (in progress)

PROJECTS UNDERTAKEN

Carely – Generative AI Companion for Elderly Care   Aug 2025 – Dec 2025

- Designed and built *Carely*, an AI-powered elderly-care companion that supports medication management, daily wellness check-ins, memory cues, and risk-aware conversations, grounded in senior-care safety and usability principles.
- Implemented a **multi-layer memory system** (short-term conversational buffer + long-term memory using **ChromaDB** + structured **SQLite** health data) with risk-aware logic and caregiver alerts, improving personalization and reducing missed health cues.

RAG Chatbot for Billing & Support -- Business Intelligence Group, UIUC   Aug 2025 – Dec 2025

- Built a **retrieval-augmented** chatbot for answering user questions from internal billing and policy documents by designing ingestion, preprocessing, and embedding pipelines that turned PDFs into a searchable knowledge base.
- Implemented **retrieval** + **LLM** response logic with out-of-scope detection and safe escalation, improving answer accuracy while preventing hallucinations and routing unclear queries to appropriate support channels.

Automated Weather Data ETL Pipeline   May 2025 – Aug 2025

- Built a robust **ETL pipeline** using Apache Airflow on **EC2** to ingest real-time weather data from OpenWeatherMap API, aimed at enabling downstream **forecasting models**, with Python **DAGs** ensuring retry logic and fault-tolerant orchestration.

YouTube Analytics Data Engineering Pipeline   May 2025 – Aug 2025

- Engineered an AWS-native pipeline to process YouTube metadata via **S3**, **Glue**, **Lambda**, and **Athena**; delivered interactive **KPI dashboards** in **QuickSight** to track regional trends.

IoT-based Automated Hydroponic System   Sep 2021 – Jul 2022

- Built an **IoT-driven** hydroponics system leveraging sensor data (pH, humidity, temperature) to improve **monitoring accuracy** by **35%**, enabling efficient resource usage; work published in IJRASET journal (link).

LEADERSHIP

- Led weekly performance review meetings for the Network Managed Services team, collaborating with client managers to streamline delivery workflows; delivered **KPI** reports on SLA compliance, incident trends, and system health to drive data-informed decisions.
- Implemented enhancements to a resource tracking application, coordinated senior leads, enabling real-time visibility into team capacity and improving resource allocation efficiency by **35%** across a **300+** member operations unit.