

Lakshmi Sudini

Data Engineer

Email: lakshmisudini14@gmail.com | **Address:** 127 chestnut crossing dr, Apt D |

Mobile: 3024825272 | **LinkedIn:** [lakshmisudini](#) | **Portfolio:** [lakshmisudini](#) | **GitHub:** [lakshmisudini](#)

PROFESSIONAL EXPERIENCE

University of Delaware, Newark, DE / Research Data Engineer

September 2024 – Present

- **Data Management System (DMS) Development:** Developed a secure, scalable DMS using AWS (S3, RDS, IAM) to support real-time environmental monitoring.
- **Serverless Computing:** Utilized AWS Lambda to automate workflows, reducing processing time by 40% and enhancing system efficiency for handling over 1TB of data monthly.
- **Database Management and Cloud Security:** Configured a PostgreSQL database hosted on RDS, improving query performance by 30% and managing over 50 SQL tables for datasets exceeding 5 million rows. Implemented IAM roles to enforce secure and role-based access control across AWS.
- **Data Integration and Processing:** Built ETL pipelines for 10+ datasets, reducing data preparation time by 50% and improving model accuracy by 20%.
- **UI/UX Design:** Currently designing a user-friendly interface to improve accessibility and enhance stakeholder interaction with the DMS.
- **Data Visualization:** Created 20+ dashboards using Python and Power BI, reducing decision-making time by 30%.

University of Delaware, Newark, DE / Data Science Research Assistant

September 2023 – December 2023

- **Excel Mastery and Data Warehousing:** Utilized Excel with AWS tools (S3, Redshift) for secure data storage and efficient warehousing. Handled missing values, removed duplicates, and merged datasets to ensure data integrity.
- **Data Manipulation:** Analyzed 120M+ rows using Python with AWS Lambda and Data Pipeline for serverless processing. Transformed data using AWS Glue to drive informed decisions.
- **Machine Learning Pipeline Development:** Designed a Python-based pipeline with Docker and Apache Airflow, automating data ingestion, feature engineering, model training, and deployment. Reduced manual effort by 80% and implemented K-means clustering models successfully.
- **Predictive Analytics and NLP:** Conducted sentiment analysis with AWS Sage Maker, improving sentiment scores by 0.3. Used Amazon Comprehend for NLP to extract actionable insights from unstructured data.
- **Model Assessment:** Streamlined ad-hoc analysis and reporting using AWS Quick Sight, aggregating insights using pivot tables and charts. Stored results securely in S3, ensuring scalability and accessibility.

Temenos The Banking Software, Chennai, India / Data Analyst

July 2021 – January 2023

- **Data Integration and ETL:** Designed and enforced ETL pipelines using SQL to extract, transform, and load data from multiple servers, ensuring data consistency and accuracy across systems.
- **Database Management:** Created and refined relational database structures, reducing query time by 25%.
- **Data Cleaning and Preprocessing:** Processed large datasets by handling missing values, removing duplicates, and ensuring data quality to support accurate analysis.
- **Data Analysis and Insights:** Conducted exploratory data analysis (EDA) using SQL and Excel, identifying trends in financial and transactional data to support decision-making.
- **Automation, Efficiency and Client support:** Automated recurring data processing workflows using SQL scripts and scheduled jobs, reducing manual intervention and improving operational efficiency by 20%. Analyzed and resolved over 20 data-related client issues in Jira monthly, providing timely insights and updates to stakeholders.

Cognizant, Hyderabad, India / Data Analyst Intern

July 2020 - June 2021

- **Data Cleaning and Preprocessing:** Leveraged Pandas to clean and preprocess a 50,000-row dataset, reducing missing values by 30% and ensuring data readiness for modeling.
- **Machine Learning Implementation:** Built machine learning models using scikit-learn, achieving a 25% increase in prediction accuracy compared to baseline models.
- **Advanced Data Transformation:** Applied techniques like Min-Max scaling and feature engineering, boosting predictive accuracy by 15%.
- **Predictive Modeling:** Refined and improved Linear Regression models, achieving an overall accuracy rate of 85% for predicting key outcomes.

TECHNICAL SKILLS

Programming Languages: SQL, Python, R, MATLAB, SAS, C.

Machine Learning and ETL tools: Linear Regression, Logistic Regression, Decision Trees, K-NN, K-means Clustering, Apache NiFi, Alteryx, Apache Airflow

Libraries and Cloud Technologies: Libraries and Cloud Technologies: Scikit-learn, TensorFlow, Keras, PyTorch, Pandas, NumPy, Matplotlib, ggplot2, AWS (S3, RDS, Lambda, IAM roles, Cloud watch, Kinesis, Redshift, Glue, SageMaker), GCP, PostgreSQL, PgAdmin

Data Visualization Tools: Power-BI, Tableau, MS-Excel, Jupyter Notebook, DeepNote, SAS, Qlik View, Qlik Sense

General Skills: Data Cleaning and Preprocessing, Data Visualization, Machine Learning, Deep Learning, Pattern Recognition, Database Structures and Algorithms, Statistical Analysis, Exploratory Data Analysis, Data Preparation, Workflow Automation.

UI Design: HTML, CSS, JavaScript, Django.

EDUCATION

Master of Science in Data Science | University of Delaware, Newark, DE

February 2023 – December 2024