

Charitha Sree Sakhamuri

+16787391251 / www.linkedin.com/in/charitha-sakhamuri / [LinkedIn](#)

SUMMARY

Highly motivated Data Science graduate (May 2025) with a strong foundation in machine learning, data engineering, and cloud technologies. Proven ability to develop and deploy ML models, build ETL pipelines using Python, SQL, and AWS, and extract actionable insights from large datasets. Seeking a challenging role where I can leverage my skills to solve complex business problems.

EDUCATION

- Master of Science — Data Science** / The University of Texas at Arlington *May 2025*
Coursework: Probability and Statistics, Data Science, Artificial Intelligence, Machine Learning, Neural Networks.
- Bachelor of Science — Data Science and Artificial Intelligence** / ICFAI University, Hyderabad *June 2023*
Coursework: Higher Calculus, Data Structures, Numerical Methods, Operating Systems, Data Warehousing and mining, Fundamentals of Blockchain Technology, Design and Analysis of Algorithms.

SKILLS

- Programming & Scripting:** Python, SQL, Bash
- Data Engineering & ETL:** Data Pipelines, Data Modeling, ETL Development, Data Warehousing
- Databases & Cloud:** PostgreSQL, MySQL, AWS (S3, Redshift, Lambda), Snowflake, Google Big Query
- Big Data & Processing:** Apache Spark, PySpark, Pandas
- Orchestration & Workflow Automation:** Apache Airflow, Prefect, Dagster
- Machine Learning & Analytics:** Feature Engineering, Model Deployment, Scikit-Learn, TensorFlow
- Generative AI:** Prompt Engineering, OpenAI API, LangChain, Hugging Face Transformers, LLM Fine-Tuning
- Data Visualization & BI:** Tableau, Power BI, Looker, Visual Studio

EXPERIENCE

- Data Science Intern — RecVue** / Hyderabad, Telangana *July 2022 - May 2023*
 - Automated testing workflows using Selenium, reducing manual testing by 30%
 - Validated REST APIs with Postman/SoapUI, ensuring performance.
 - Contributed to Agile sprints to optimize software performance.
- Data Analyst — Vijeta High School** | Guntur, Andhra Pradesh *June 2020 – June 2022*
 - Analyzed datasets using SQL/Python, enhancing report accuracy by 20%.
 - Developed Tableau dashboards for real-time data insights.
 - Applied ETL processes to clean and transform data
- Software Developer Intern — Skill Banc** | Hyderabad, Telangana. *May 2021 – July 2021*
 - Built software solutions, optimizing processing time by 15%.
 - Automated unit/integration tests, improving software quality.
 - Maintained documentation to improve troubleshooting

ACADEMIC PROJECTS

1. Railway Ticket Booking System (Python)

- Developed a Python-based railway ticket booking system with authentication and payment integration.
- Designed and optimized a relational database (PostgreSQL/MySQL) for efficient data storage and retrieval.
- Analyzed booking trends and peak times using Pandas and Matplotlib for data-driven insights.

2. Email Spam Detection (Machine Learning)

- Built a spam classification model using Scikit-Learn to filter out unsolicited emails.
- Processed and cleaned a dataset of 50,000+ emails using NLTK for tokenization and stop word removal.
- Trained multiple machine learning models (Logistic Regression, Naïve Bayes, Random Forest) to achieve an accuracy of 98%.
- Deployed the model using Flask and AWS Lambda, integrating it into an email filtering system.

3. Breast Cancer Prediction (Deep Learning)

- Designed a deep learning model using TensorFlow/Keras to predict breast cancer based on medical imaging and tabular data.
- Collected and pre-processed patient data, handling missing values and feature scaling using Pandas and Scikit-Learn.
- Trained a CNN (Convolutional Neural Network) and ANN (Artificial Neural Network) achieving an F1-score of 94% on unseen data.
- Deployed the model using Docker and Fast API, enabling real-time cancer prediction for medical applications.

4. Data Science Job Market Analysis

- Scraped job postings from multiple sources using BeautifulSoup and Selenium, collecting 100,000+ listings across different job roles, locations.
- Designed an ETL pipeline to clean, transform, and store job market data in AWS Redshift/Google Big Query.
- Used Apache Spark for large-scale data processing, reducing data analysis time by 40%.
- Created interactive Tableau/Power BI dashboards to visualize salary trends, most in-demand skills, and regional hiring patterns.

Certifications

- AWS Educate – AWS Machine Learning Foundations**
- Duke University--Python Essentials for MLOps**
- Microsoft—Microsoft Python Development**