**Laxmi Thodupunuri**

* Hyderabad , Telangana • [thodupunurilaxmi2003@gmail.com](mailto:thodupunurilaxmi2003@gmail.com) • +91 9398528454• [LinkedIn](https://in.linkedin.com/in/thodupunuri-laxmi-she-hers-1980b0205)• [Github](https://github.com/thodupunurilaxmi2003) • Open to Relocate

“Software Engineer (**2 yrs**) specializing in high‑volume **cloud data migration** & **lake‑house optimization** on **AWS**; proven record cutting costs **40 %+** and boosting data quality to **96 %**”

**SKILLS**

**Programming Languages:** Python (NumPy, Pandas, TensorFlow, PyTorch, Seaborn, Matplotlib), SQL

**Databases & Storage:** Snowflake, Redshift, S3, MySQL, MSSQL, MongoDB (NoSQL)

**ETL & Visualization:** Databricks, Athena, Power BI, Tableau, Splunk , Grafana

**Big Data Technologies**: Spark, Kafka, YARN, HDFS, Hive

**Cloud Technology**: Amazon Web Services (Lambda , DynamoDB, S3 ,AWS Glue ,Step Functions ,VPC) , Terraform

**Other Tools**: Git, Agile, JIRA, Jenkins, IntelliJ

**WORK EXPERIENCE**

**Software Engineer & Data Jan 2024 – Present**

**JPMorgan Chase & Co. Hyderabad**

* Worked on building scalable, self-service data migration utilities to seamlessly move **250 TB+** datasets from on-premise **(Teradata**, **Oracle)** to the **AWS** cloud, increased data discoverability and management by **80%**
* **Implemented** PySpark‑based ingestion pipelines with automated schema validation (**Apache Deequ**) and file‑format checks (Delta Lake utilities), boosting data‑quality pass rates from **85 % to 96 %** and keeping multi‑environment datasets in sync.
* **Implemented** a fully serverless data‑processing stack on **AWS (S3|Lambda|Step Functions |VAC)**, replacing batch servers with event‑driven orchestration and automated retries—cutting monthly compute spend by **≈40 %** and shrinking incident recovery time from **hours to minutes**.
* **Built** a **self‑service ingestion framework**—Terraform‑provisioned endpoints backed by **PySpark**—that lets 50 + internal publishers onboard data in **< 1 hour** (down from multi‑day ticket cycles).
* **Optimized** the Lakehouse by migrating to **Apache Iceberg** open‑table format and scheduling size‑aware compaction jobs, trimming S3 scan costs **25 %** and cutting query latency **≈35 %**.
* **Modernized** governance in **AWS Glue Catalog**, automating lineage, auditing, and schema‑evolution rules, while rolling out point‑in‑time **DynamoDB backups** to hit **100 % RPO/RTO compliance**.

**Software & Data Engineer Intern June 2023 –July 2023**

**JPMorgan Chase & Co. Hyderabad**

* **Implemented** an **autoencoder‑based anomaly‑detection pipeline** (PyTorch | SMOTE) that learns “normal” card‑payment behavior from 1 B+ legitimate transactions and flags outliers as fraud with **95 % overall accuracy** and **92 % precision**.
* **Bench‑marked** six candidate models (KNN, LOF, Isolation Forest, One‑Class SVM, Autoencoder, Robust Covariance), running a full precision‑recall analysis to select the autoencoder—cutting false‑positive alerts **≈40 %** versus the next‑best model.
* **Automated data balancing** via **SMOTE** to synthesize minority‑class fraud events, lifting recall from **63 % → 89 %** without sacrificing latency in the real‑time scoring path.
* **Delivered** a **Streamlit front‑end** for live transaction scoring and drill‑down analytics, enabling risk analysts to investigate suspicious activity in seconds and accelerating incident response by **3 ×**.

**EDUCATION**

**Chaitanya Bharathi Institute of Technology, Hyderabad** **Hyderabad, IN**

Bachelor of Engineering in Information Technology

Grade : 9.8

**ACADEMIC PROJECTS**

**Face Emotion Recognition. Key Skills**: **python | Deep Learning |Open CV**

* Developed a CNN-based facial emotion recognition system using OpenCV with **~96%** accuracy. Visualized real-time emotion percentages through dynamic bar graphs.

**Future Energy Prediction Key Skills: Deep Learning**

* The main of project is to create a Neural network which can remember the pattern of energy consumption over past 10 years which is made using LSTM(Long Short Term Memory) and predict the energy consumption based on the pattern learned

**Interpretable Anomaly Detection in Industrial Control Systems (**[Link](https://ieeexplore.ieee.org/document/10677782)**) Key Skills: Deep Learning , Federated Learning , XAI**

* **Engineered** a federated‑learning anomaly‑detection framework for IoT: FedAvgM‑aggregated autoencoders trained across edge nodes with SHAP/LIME feature explainability, outperforming centralized baselines while keeping all raw device data local