

Keerthi Gopireddy

Data Analyst | Data Scientist | M.S. in Artificial Intelligence

✉ keerthi16101996@gmail.com

☎ 512-698-4660

📍 Texas, United States



<http://www.linkedin.com/in/lakshmi-keerthi>



https://lakshmi-keerthi.github.io/my_portfolio



<https://github.com/lakshmi-keerthi>

PROFESSIONAL SUMMARY

Data Analyst/Data Scientist with **5+ years** of experience delivering analytics and AI-enabled insights across financial services, logistics, and digital advertising domains. Strong expertise in **SQL, Python, Dataiku**, and **BI tools** to automate data workflows, perform exploratory data analysis, and build stakeholder-ready dashboards. Holds an M.S. in **Artificial Intelligence**, with hands-on experience applying machine learning, feature engineering, and Generative AI techniques, including LLM integrations via APIs, to support analytics automation and business decision-making.

EDUCATION

Master of Science, Artificial Intelligence

Jan 2023 - Dec 2024

University of North Texas | Denton, Texas | **CGPA: 4.0**

Bachelor of Technology, Electronics and Telematics

Jul 2014 - May 2018

Jawaharlal Nehru Technological University | Hyderabad, India

SKILLSET SPOTLIGHT

Programming & Query Languages: Python, SQL, MATLAB

Data Analysis & Statistical Methods: Descriptive & Inferential Statistics, Hypothesis Testing, A/B Testing, Correlation Analysis, Root Cause Analysis, Empirical Analytics, Quality Analysis

Data Visualization & Business Intelligence Tools: Power BI (DAX), Tableau Cloud, Microsoft Excel, Dataiku DSS

Data Manipulation: Data Cleaning and Preprocessing, ETL Pipelines, Data Modeling, DBMS

Python Libraries: Pandas, NumPy, Scikit-learn, Matplotlib, Seaborn, SciPy, Plotly, Statsmodels, TensorFlow, Pytorch

Machine Learning & AI: Supervised & Unsupervised Learning, Exploratory Data Analysis, Feature Engineering, Model Evaluation & Deployment, Natural Language Processing(NLP), GANs, RAGs

WORK EXPERIENCE



Capgemini

NYC, USA

Senior Consultant Data & Insights

Jun 2025 to till date

- Automated data preparation and transformation workflows for a *financial services client* using *Python, SQL*, and *Dataiku DSS*, reducing manual data processing effort by ~70%
- Performed *exploratory data analysis (EDA)* on large financial datasets, including transactional and risk-related data, identifying *data quality issues* and *anomalies* that improved downstream analytics reliability by ~30%
- Built interactive *dashboards* and analytical *web applications* in Dataiku to visualize operational KPIs and model monitoring metrics, cutting stakeholder reporting turnaround time by ~40%
- Supported data preparation, validation, and quality checks for *AI governance-related metrics* for an *in-house AI system*, enabling compliant model monitoring and reporting
- Collaborated with cross-functional teams to deliver *structured datasets*, accelerating *model development* and reporting cycles



xSignOn Corporation

Texas, USA

AI Engineer Intern

Mar 2025 to Jun 2025

- Built *Tableau* dashboards to monitor failed logins, unusual IP activity, and user access patterns, helping the team respond to issues 35% faster.
- Working with a wealth management company to automate customer query flow, improving response efficiency.

- Conducted *exploratory data analysis* (EDA) on customer financial behavior to identify investment patterns and risk profiles, enabling more targeted advisory recommendations.
- Built a *Retrieval-Augmented Generation* (RAG) based chatbot tailored to answer company-specific wealth management queries, improving response accuracy and user engagement in internal advisory use cases.



Wipro Technologies

Hyderabad, India

Data Scientist

Aug 2021 to Dec 2022

- Improved *ad click prediction* in Huawei Ads Platform through data querying and processing using HiveQL, followed by EDA, correlation analysis, and feature engineering, increasing eCPM by 1.4% and boosting advertiser revenue
- Automated *KPI tracking* in the cloud-based test environment using Python, reducing manual effort by 10%
- Identified model overfitting through *performance analysis*, implemented early stopping in a Deep Learning model, and validated improvements via A/B testing, increasing product performance, eCPM by 3%
- Built *ETL pipelines* in the *Dataiku platform* to automate data preprocessing for Citibank, enhancing data relations and processing efficiency, and reducing system load by ~45% in an Agile environment



Amazon Robotics

Hyderabad, India

Data Quality Analyst

Aug 2018 to May 2021

- Implemented *HITL strategies* and *SOPs* for human annotators across Robin, TRON, Canvas, and Proximity, improving data preprocessing, model training, and reinforcement learning as part of Go-AI team
- Analyzed inventory & FCSKU datasets queried using *SQL*, followed by *data wrangling*, identified the 'Scan while Stow' issue, and suggested preventive measures, leading to an ~8% decrease in low-confidence ML predictions
- Streamlined Dislike by *bucketizing issues*, setting up an SOP, and implementing product development based on *data-driven insights*
- Analyzed gaps in the Dislike Mechanism, identified root causes of high turnover time, presented logic changes for auto-generated action emails, leading to a 30% improvement in *turnover time* and a ~20% reduction in *effort*
- Automated weekly reports using *Microsoft Power BI*, *Python*, and *Excel* reducing the manual effort by ~70%
- Led alignment of *standard operating procedures* (SOPs) across three *machine learning models* as the technical expert, mentoring and managing team members on feature identification aligned with product and business needs

ACHIEVEMENTS AND ACCREDITATIONS

- Merit Certificate with Excellent Grade in '**Advanced Artificial Intelligence and Machine Learning Program**' from IIIT Hyderabad (Dec 2018 to March 2019)
- '**Lean Six Sigma Yellow Belt Certification**' by ACES Academy at Amazon for performing quality analysis and suggesting data-driven process changes to streamline the dislike mechanism.
- Awarded with '**Champion IDS Research Analyst**' in Q2 2020, Q3 2020, and Q1 2021 at Amazon for my expertise in root cause analytics focusing on the inventory data
- Accredited as the '**Most Valuable Player**' in the team for the year 2020 at Amazon for identifying trends in the data leading to more dislikes using statistical modelling
- Recognized as '**Outstanding Contributor**' in February 2022 by Huawei Ads for performing data analysis to identify important features to optimize ad click-through rate
- Databricks** accredited Generative AI Fundamentals Certification
- Certified as an ML Practitioner and Advanced Designer by *Dataiku*
- Completed LeetCode "**SQL 50**" and "**Intro to Pandas**"

INTERNSHIP / PART-TIME

- Worked as a **Digital Imaging Student Assistant** (Feb 2023 - Dec 2024) at the University of North Texas:
 - Captured metadata* for 3,000+ historical letters, enhancing archival preservation and accessibility
 - Analyzed* historical records, identifying **trends** and categorizing them for improved accessibility and research
 - Digitized* 20+ archival books and ledgers with precision to ensure long-term preservation
- Developed a **Mini-CNC controller** using *MATLAB* and *Arduino* during an Internship at Central Institute Of Tool Design from May 2017 to July 2017

KEY PROJECTS

- [Animal Shelter Database System](#) | Database Management Systems | Spring 2024

Developed a database management system for the animal shelter using SQL on Oracle Cloud. Analyzed data requirements, defined entities and relationships, and created an ERD based on the schema. Applied normalization for data integrity and implemented stored procedures/functions to streamline adoption, and medical records.

- [Analysis and Prediction of Email Click-Through Rate](#) | Empirical Data Analysis | Fall 2023

Analyzed customer and product data to predict email click-through rate (CTR), using exploratory data analysis (EDA), correlation analysis, and feature selection to identify key drivers of engagement. Developed machine learning models, particularly XGBoost, to optimize targeting strategies and improve email campaign performance.

- [Visualizing the Impact of COVID-19 on Air and Land Transportation](#) | Scientific Data Visualization | Fall 2023

Analyzed the impact of COVID-19 on road and air transportation safety using complex datasets to uncover key correlations. Built interactive dashboards in Tableau and Power BI, and developed a custom D3.js webpage hosted on GitHub to visualize trends. Used Python for data cleaning, transformation, and exploratory analysis.

- [Generating Synthetic Faces using Generative Adversarial Networks](#) | Machine Learning | Spring 2023

Developed a generative AI model using GAN model architecture capable of producing realistic, high-quality synthetic faces indistinguishable from real faces. GAN uses a generator and discriminator to produce fake faces. Produced synthetic faces that do not exist in the dataset ensuring privacy-preserving data augmentation.