

SHREE HARI ANBAZHAGAN

Chennai, Tamil Nadu - 600106, India

 [linkedin.com/in/a-shreehari](https://www.linkedin.com/in/a-shreehari)

 +91 7667899999

 shreeharianbazhagan@gmail.com

EXPERIENCE

Comcast

Development Engineer 1, Data Engineer

July 2023 - present
Chennai, India

- Led the development and maintenance of ETLs for internal and external teams, integrating Kafka with AWS Redshift, ensuring near-zero downtime and improved system reliability.
- Architected and deployed multiple solutions using **infrastructure as code (IaC)** and **draw.io**, following best practices for **scalability, security, and automation**, resulting in a **30% reduction in deployment time and effort**.
- Designed and deployed data synchronization tools (**EnvSyncer**) to ensure consistency across **Data Warehouses**, implementing anomaly detection with AWS Lookout Metrics, and building **reconciliation frameworks** that improved data accuracy by **40%**.

Comcast

Intern

Jan 2023 - June 2023
Chennai, India

- Developed a **Slack notification system** to monitor ETL runs and job status, providing real-time alerts and daily cost breakdowns to enhance operational efficiency.
- Assisted and learnt of a Data Warehouse (DWH) application to create a single source of truth for organizational data.

Lasso and Marquee Design Solutions

Intern

June 2019 - July 2019
Chennai, India

- The internship provided valuable insights into IT operations, teamwork, and professional communication.

EDUCATION

Indian Institute of Technology Madras

Diploma in Data Science

Sep 2023 - present
Chennai, India

Gaining expertise in analytics, machine learning, and big data technologies through the program.

SASTRA Deemed University

Bachelor of Technology in Computer Science & Engineering,

Graduated First Class with Distinction.

July 2019 - May 2023
Thanjavur, India

PROJECTS

IITM Machine learning project

Sep 2024 - Nov 2024

- Built models to predict bank telemarketing success.
- Performed **EDA and preprocessing** based on insights, **trained 4 models**, and achieved the best results with an **accuracy of 87%** and a macro F1 score of 0.75, with notable performance in precision and recall for the 'no' class.

Sentiment Analysis and Topic Modeling for Tweets

Dec 2023 - Feb 2024




- Contributed to the ML project "Inside Out – Understanding User Emotions," focusing on sentiment and topic analysis of tweets, with findings presented at the **G20 Library Summit 2024**.
- Developed an Excel sheet with base data and analysis results, covering sentiment analysis, topic modeling, and engagement metrics. incorporating visualizations like heat maps and graphs to highlight data trends.
- Used **pySentimiento** for sentiment analysis and **cardiffnlp/tweet-topic-21-multi** for topic modeling, achieving **94.53% accuracy** in extracting insights from tweets.

Diabetes prediction

Sep 2022 - Dec 2023

- Developed a **Django web application** for **diabetes prediction**, leveraging multiple **machine learning algorithms** (**Logistic Regression, Random Forest, Adaboost**, etc.) to provide accurate predictions based on user health metrics.
- Integrated **PDF report generation** and **SMS notification features** within the application, with the model achieving an **accuracy score of 0.82 for Random Forest** and **0.80 for Adaboost**, enhancing user experience by delivering prediction results in both **downloadable and mobile-friendly formats**.

Other projects

- Helper ** Created a personal project involving **reusable helper classes** for **AWS, API integration, Kafka, and Slack**.
- IITM Business Data Management project ** Aims to tackle business challenges by leveraging data management and analysis techniques.
- Resume ** My LaTeX resume.

TECHNICAL SKILLS

Programming: Python, SQL, Terraform (Infrastructure as Code), shell commands (basic), YAML, JSON.

Tools and Frameworks: AWS (AWS Lambda, Glue, S3, Redshift, Athena, IAM (Roles, Policies), Secrets Manager, CloudWatch), Boto3 (AWS API), REST API, PostgreSQL, Docker, Kafka, Machine Learning (ML), Git, VSCode, JIRA, Excel.