SHREE HARI ANBAZHAGAN

Chennai, Tamil Nadu - 600106, India

 linkedin.com/in/a-shreehari · +91 7667899999 · shreeharianbazhagan@gmail.com

👨‍💼 EXPERIENCE

Comcast 🔗 *July 2023 - present*
Development Engineer 1, Data Engineer *Chennai, India*

- Developed ETLs integrating Kafka with AWS Redshift, achieving near-zero downtime and improving reliability.
- Architected scalable, secure solutions with **IaC** tools and **draw.io**, reducing deployment time by **30%**.
- Designed **EnvSyncer** to sync data between environments, improving data availability.
- Built **reconciliation frameworks** to validate data integrity between the source and DWH, improving data accuracy by **40%**.

Comcast 🔗 *Jan 2023 - June 2023*
Intern *Chennai, India*

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

Lasso and Marquee Design Solutions 🔗 *June 2019 - July 2019*
Intern *Chennai, India*

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

🎓 EDUCATION

Indian Institute of Technology Madras 🔗 *Sep 2023 - present*
Diploma in Data Science *Chennai, India*
Gaining expertise in analytics, machine learning, and big data technologies through the program.

SASTRA Deemed University 🔗 *July 2019 - May 2023*
Bachelor of Technology in Computer Science & Engineering, *Thanjavur, India*
Achieved First Class with Distinction.

💻 PROJECTS

IITM Machine learning project 🔗 *Sep 2024 - Nov 2024*

- Built models to predict bank telemarketing success.
- Performed **EDA and preprocessing**, **trained 4 models**, achieving **87% accuracy** and 0.75 F1 macro score.

Sentiment Analysis and Topic Modeling for Tweets 🔗 *Dec 2023 - Feb 2024*

- Contributed to "Inside Out – Understanding User Emotions" ML project, analyzing tweet sentiments and topics, presented my findings at **G20 Library Summit 2024**.
- Compiled findings in Excel sheets with sentiment analysis, topic modeling, engagement metrics, heat maps, and graphs to visualize data trends.
- Used **pysentimiento** and **cardiffnlp/tweet-topic-21-multi** to achieve **94.53% accuracy** in tweet insight extraction.

Diabetes prediction 🔗 *Sep 2022 - Dec 2023*

- Developed **Django web application**, using **machine learning algorithms** (Logistic Regression, Random Forest, Adaboost) to predict diabetes using **basic health metrics**
- Integrated **PDF report generation** and **SMS notification** features, with **Random Forest achieving 0.87 accuracy**, delivering results in downloadable and mobile-friendly formats.

Other projects

- **Helper** 🔗 Created reusable helper classes for AWS, API integration, Kafka, and Slack.
- **IITM Business Data Management project** 🔗 Tackles business challenges through data management and analysis.
- **Resume** 🔗 using LaTeX.

🛠 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.