

# S HARINI

Data Science Intern | Machine Learning | Experiments | Analytics

✉ hariniselvakummar@gmail.com  
🌐 harini-flqo.onrender.com



📍 Chennai, India

🌐 harini-s-4b127428b

👤 Harini-win

## EXPERIENCE

### Software Engineering Intern

#### Anusoft (AI Systems)

📅 May 2025 – Jun 2025

- Developed SQL-based validation pipelines integrated into production systems to verify AI model outputs and ensure reproducibility.
- Collaborated with engineers and product stakeholders to align validation logic with real-world usage scenarios.

### Data Analytics Intern

#### Infosys (Pragathi Cohort 3)

📅 Jan 2025 – Apr 2025

- Built data preprocessing and transformation pipelines supporting analytics and reporting workflows.
- Automated recurring data checks and reporting, reducing manual effort by 35.
- Worked closely with teams to translate business requirements into data-driven solutions.

## PROJECTS

### Purchase Value Prediction from Web Session Logs

#### Machine Learning & XGBoost

- Designed an end-to-end ML pipeline to predict customer purchase value from session-level behavioral data with extensive EDA and feature engineering.
- Trained and evaluated 9+ models (inc. Random Forest, MLP); achieved  $R^2 = 0.96$  with GradientBoostingRegressor via Grid-SearchCV.

### Automated Data Analysis & Visualization Engine

#### Python & GPT-4o-Mini

- Built a Python system that automates EDA, anomaly detection, and insight narration for CSV datasets.
- Integrated GPT-4o-Mini to generate actionable Markdown reports, optimizing token usage by summarizing data pre-inference.

### GitHub User Analysis — Stockholm Case Study

#### Data Analytics & API Integration

- Collected ecosystem data via GitHub API to analyze trends in language popularity and repository engagement.
- Performed regression analysis identifying a positive correlation between bio details and follower growth.

## SKILLS

Python

SQL

JavaScript

Git

Jupyter

Pandas

NumPy

Scikit-learn

Matplotlib

Seaborn

EDA

Feature Engineering

Hypothesis Testing

Experimentation

Regression

Classification

Clustering

CNNs

Ensemble Models

LLMs

Prompt Engineering

RAG

Cross-validation

$R^2$

RMSE

Dashboards

Data Visualization

## EDUCATION

### B.Tech Computer Science & Engineering

#### VIT Chennai

📅 2023 – 2027

- CGPA: 9.14

### B.S. in Data Science

#### IIT Madras

📅 Ongoing

## CERTIFICATIONS



#### Google Cloud Arcade Ranger

Completed 100+ labs in cloud tooling, automation, and deployment workflows.



#### Building GenAI Apps — MongoDB Atlas

Hands-on experience integrating secure GenAI workflows with databases.

## ACHIEVEMENTS



#### President — Mathematics Club

Managed 150+ members and oversaw 20+ events for a wider audience across campus.



#### Topper Badge (9x) – IITM BS DS

Awarded to top 1% performers in individual modules.