

# SAI PREETHI POKA

Los Angeles, CA  
562-572-2969  
pokasaipreethi@gmail.com

[linkedin.com/in/saipreethipoka](https://www.linkedin.com/in/saipreethipoka)  
[github.com/pspreethi](https://github.com/pspreethi)  
[medium.com/@pokasaipreethi](https://medium.com/@pokasaipreethi)  
[pspreethi.github.io/portfolio](https://pspreethi.github.io/portfolio)

## EDUCATION

### Master of Science in Information Systems

California State University, Long Beach - **GPA: 4.0/4.0**

Graduation: May 2025

Long Beach, CA

**Relevant Coursework:** Machine Learning for Business, Responsible AI, Database Management Systems, Systems Analysis & Design

## SKILLS

**Languages:** Proficient – Python and SQL

Experience – PySpark and R

**Tools & Tech:** Microsoft Excel, Tableau, Power BI, Streamlit, DataBricks, Git, Azure, AWS, Airflow, Jupyter, VS Code

**Analytics & Modeling:** A/B Testing, Experimentation & KPI Design, Funnel Analysis, Churn Prediction, Statistical Modeling

**Data Science & ML:** Supervised Learning, Classification, Regression Analysis, Time Series Analysis, Forecasting, LLMs

## WORK EXPERIENCE

### Associate Engineer II – Data Science, Innovation

Harman International

Sep 2022 – Jul 2023

Bangalore, India

- Built a forecasting model (R) to optimize demand planning, cutting planning costs by 40%.
- Redesigned the Revenue Growth Management tool (Python, SQL, Azure) to integrate the Australia region, automating workflows and saving 6 analyst hours/week while reducing delivery errors by 30%.
- Streamlined PySpark data workflows on distributed clusters, reducing query time by 30%.
- Automated testing pipelines (PySpark) for analytics solutions like Market Structure & Hispanic Insights, reducing validation time by 50% and accelerating feature deployment.
- Designed a Python-based reporting tool, reducing report time from 2 hours to 30 minutes.
- Wrote clean, user-first documentation that halved onboarding questions for new devs and led mentorship initiatives for team members.

### Associate Engineer I – Data Science, Delivery

Harman International

Sep 2021 – Aug 2022

Bangalore, India

- Processed 100M+ POS records (PySpark, SQL) to enable predictive modeling for retail analytics.
- Built regression and classification models (Python) with 15% improvement in forecast accuracy.
- Automated data ingestion from Unify platform (Python, API), tripling execution speed.
- Developed dashboards (Streamlit, Dash) for brand reversal analysis, ML outputs, and transaction trends, empowering stakeholders to monitor behavioral shifts and adjust product strategies in near real-time.

### Product Analyst Intern

Cognizant Technology Solutions

Feb 2021 – Jul 2021

Hyderabad, India

- Converted stakeholder requirements into clear user stories for product dev team, reducing scope changes by 3–4 iterations per feature and speeding up delivery cycles.
- Designed SQL schemas that improved key query performance by 25%+.
- Built and tested RESTful APIs, keeping response times under 500ms.
- Integrated backend APIs to improve frontend load time by 20-30%.

## PROJECTS

### Air Quality Forecasting & Analytics

[GitHub](#)

- Developed a pipeline to fetch and clean AQI data from the OpenAQ API.
- Performed EDA to identify trends and seasonal patterns in air quality.
- Built a time-series forecasting model using Prophet for 90-day AQI predictions.
- Deployed an interactive dashboard to visualize real-time data and forecasts.

### A/B Testing Automation & ML Analysis

[GitHub](#)

- Built automated A/B test pipeline (Python, XGBoost) to analyze test results and predict conversions (96.8% accuracy)
- Integrated with PowerPoint for instant reporting; system dynamically adapts to new datasets.