

SARATH THARAYIL SREENIVASAN

+91 7012604990 | iamsaraththarayil@gmail.com | www.linkedin.com/in/SarathTharayil | www.saraththarayil.com

Senior Data Scientist with ~4 years of experience developing forecasting models and scalable data solutions that support commercial decision-making. Experienced in time-series modelling, demand forecasting and statistical analysis, with strong expertise in Python, SQL and cloud platforms including AWS and Azure. Proven track record of leading end-to-end analytical initiatives, from data engineering and model development to stakeholder communication and deployment of production forecasting systems. Adept at translating complex modelling outputs into actionable insights that support pricing strategy, operational planning and risk management.

PROFESSIONAL EXPERIENCE

Google – Data Analyst (Contract)

London, Oct 2025 – Jan 2026

- Designed and implemented 15+ conceptual and logical data models for ML evaluation frameworks, supporting structured comparison of model variants and reducing evaluation cycle time by 35%
- Collaborated with 8-member cross-functional agile team, participating in bi-weekly sprint planning and retrospectives to deliver iterative data model improvements
- Maintained high-quality data standards through systematic validation and review processes supporting production ML evaluation workflows

Upwork – Freelance Data Scientist

Remote, Jan 2024 – Oct 2025

- Applied NLP techniques (text classification, sentiment analysis, named entity recognition) to process customer feedback and support ticket data
- Designed and implemented feature engineering pipelines on GCP BigQuery, creating multiple engineered features from raw data and improving model performance by 18%
- Applied advanced statistical methods including Bayesian inference, hypothesis testing, and experimental design to validate model performance and business impact
- Built MLOps infrastructure using MLflow, Docker, and CI/CD pipelines (GitHub Actions), enabling automated model training, versioning, and deployment with 98% deployment success rate

MuSigma – Senior Data Scientist (Commodity Price & Demand Forecasting Platform) Bengaluru, Sep 2020 – Sep 2022

- Led a data engineering team of 10 for the design and implementation of scalable data transformation and ingestion pipelines, converting raw market data into structured, ML-ready datasets
- Designed and deployed an end-to-end forecasting system to predict demand and pricing for 60+ petrochemical products across multiple regions at daily, weekly and monthly horizons
- Developed hybrid ML solutions combining statistical methods (ARIMA, exponential smoothing) with tree-based models (XGBoost) for time-series forecasting, improving forecast accuracy by 18%
- Engineered time-series features including seasonality, lagged variables and market indicators to improve forecasting performance
- Built production-grade ML pipelines on AWS using Python, PySpark, and Airflow, processing 3M+ daily transactions with automated retraining and monitoring capabilities
- Deployed ML models using Docker containers on AWS (SageMaker, ECR, ECS), implementing CI/CD pipelines for automated testing and deployment
- Implemented forecast monitoring and backtesting frameworks to continuously evaluate model accuracy and improve forecasting performance
- Led and mentored team of 23 data scientists on ML best practices, conducted code reviews, and established model development standards, improving team delivery velocity by 30%
- Worked directly with senior leadership and commercial stakeholders to translate model outputs into pricing strategy insights that improved margin optimisation

EDUCATION

The University of Sheffield - Master of Science, Data Science

Sep 2022 – Jan 2024

- Relevant Coursework: Data Analysis and Statistics, Data Visualisation, Data Mining, Database Design, Business Intelligence, Artificial Intelligence, Big Data Analytics, Data Ethics, Inferential Statistics, Classification of Real and AI-generated human faces (Dissertation)

Cochin University of Science and Technology – B. Tech, Computer Science

Aug 2016 – Aug 2020

- Coursework: Analysis of DDoS attacks by Botnet infected IoT devices using multiple machine learning classifiers

SKILLS, ACHIEVEMENTS & CERTIFICATIONS

Tools & Platforms: Python and R Programming, SQL, NoSQL, JavaScript, Power BI, SPSS Modeler, AWS (S3, Redshift, Glue, Lambda, SageMaker), GCP, Salesforce, MCP, OCR, Neo4j, LanGraph, Langchain, LLM, Generative AI, Docker, Postman, Apache Airflow, PySpark, dbt, Databricks, Snowflake, PowerBI, Looker, Tableau, Streamlit, Matplotlib, Spark, Hadoop (HDFS, Hive, H-Base, Oozie), Apache Nifi, Azure (ADLS, DataBricks, Azure Data Factory), Elasticsearch, AVRO / PARQUET file formats, Snowflake, SQL, Data Vault 2.0, MicroStrategy, Informatica, Kafka, Bitbucket, Jenkins, Azure DevOps, Kubernetes, docker, SonarQube

Skills: Data Modelling & Visualisation, Statistical Analysis, Time Series Forecasting, Forecast Evaluation, Backtesting, Scenario Analysis, Demand Forecasting, Model Monitoring, Price Forecasting, RAG, Agentic AI, Prompt Engineering, Data Strategy and Governance, Leadership, Stakeholder Management, Design Thinking, Agile Project Management, Storytelling.