

SARATH THARAYIL SREENIVASAN

Sheffield • iamsaraththarayil@gmail.com • (+44) 7533886825 • <https://www.linkedin.com/in/SarathTharayil>

SUMMARY

Experienced Data Scientist with over 2 years of expertise in developing and implementing advanced ML models across various domains, including computer vision, NLP, and business forecasting. Proven ability to drive innovation, enhance operational efficiencies, and deliver high-impact solutions. Proficient in Python, deep learning frameworks, and cloud-based ML pipelines, with a strong focus on translating complex data into actionable insights.

PROFESSIONAL EXPERIENCE

MU SIGMA INC.

Bengaluru, India

Data Scientist II

January 2022 – September 2022

- Developed sophisticated ensemble models combining ARIMA and LSTM networks for price and demand forecasting in the energy sector, achieving 90% accuracy and reducing forecast error by 20%
- Engineered a distributed data processing pipeline using PySpark and Apache Hive, optimizing the ingestion and transformation of 1000+ daily API calls, resulting in a 22% reduction in data processing time.
- Implemented advanced anomaly detection algorithms using isolation forests and autoencoders in RapidMiner AI Hub, reducing data anomalies by 10% and improving overall model performance.
- Designed a novel feature engineering framework using wavelet transforms and Fourier analysis to extract temporal patterns from energy market data, enhancing forecast accuracy by 15%.
- Collaborated with dashboarding team on actionable Tableau visualizations for price and demand forecasts for products in the energy market, enabling 15% faster insights.
- Contributed to user testing, iteratively improving dashboards for 20% fewer usability issues.
- Skills: Python, PySpark, Apache Hadoop, Hive, HiveQL, SQL, RapidMiner Studio, Software AG webMethods, Octoparse, Figma, Microsoft Excel

Data Scientist I

January 2021 – January 2022

- Engineered a scalable machine learning pipeline using Python and PySpark to identify root causes of product shortages, incorporating gradient boosting and random forest models, resulting in a 38% reduction in daily cuts worth \$1.5M.
- Developed an advanced Bayesian structural time series model for Market-Mix Modeling, optimizing marketing spend across channels and increasing ROI by 16% in the Mexican market.
- Created a custom loss function incorporating business-specific constraints for the Market-Mix model, leading to more actionable and realistic marketing spend recommendations.
- Designed and implemented a reinforcement learning algorithm to dynamically adjust inventory levels based on real-time demand signals, reducing stockouts by 25%.
- Quantified media effectiveness using statistical analyses and R-based predictive models, optimizing marketing allocations across portfolios.
- Developed an RShiny Market-mix modeling simulator, predicting revenue outcomes, and leading to a 16% increase in incremental revenue in Mexico.
- Skills: Python, R, RShiny, PrestoSQL, Microsoft Excel, PowerBI, DbVisualizer, Azure DevOps

Data Sciences Intern - Apprentice

September 2019 – January 2021

- Developed a multi-armed bandit algorithm for A/B testing in digital marketing campaigns, increasing click-through rates by 30% compared to traditional methods.
- Created a custom regularization technique for neural networks to handle imbalanced datasets in pharmaceutical trial analysis, improving model generalization by 40%.

- Produced persuasive sales pitch decks for multiple projects, articulating intricate problem-solving approaches and innovative solutions.

EDUCATION

THE UNIVERSITY OF SHEFFIELD

Master of Science, Major in Data Science

Sheffield, United Kingdom

January 2024

COCHIN UNIVERSITY OF SCIENCE AND TECHNOLOGY

Bachelor of Technology, Major in Computer Science and Engineering

Kochi, India

July 2020

PROJECTS

Adversarial Exploits: Classification of AI-generated human faces (Dissertation)

August 2023

- Developed a novel architecture combining CNNs and Transformer networks for robust classification of AI-generated faces, achieving 99% accuracy on challenging datasets.
- Implemented advanced data augmentation techniques, including style transfer and progressive growing of GANs, to enhance model generalization.
- Designed a self-supervised pretraining approach using contrastive learning, reducing the need for large, labeled datasets and improving model performance by 15%.

Analysis of DDoS infected IoT Devices using multiple machine learning classifiers

June 2020

- Implemented an ensemble of gradient-boosted decision trees, support vector machines, and deep neural networks for DDoS attack detection, achieving 98% F1-score.
- Developed a novel feature extraction technique using graph neural networks to capture complex network topology patterns, improving detection accuracy by 10%.

Predicting Song Popularity using Regression models: Comparative Analysis

February 2023

- Engineered a multi-modal deep learning model combining convolutional neural networks for audio processing and LSTM networks for lyric analysis, outperforming traditional regression models by 30% in popularity prediction.
- Implemented a custom attention mechanism to identify key audio and lyrical features contributing to song popularity, providing interpretable insights for music producers.

SKILLS

Machine Learning: **Deep Learning, Transfer Learning, Reinforcement Learning, GANs, Attention Mechanisms**; Advanced Analytics: **Time Series Analysis, Bayesian Modeling, Causal Inference, Graph Neural Networks**; Programming: **Python, R, Java, C++**; Data Science and Analytics: **RapidMiner, KNIME, SPSS, Keras**; ML Frameworks: **TensorFlow, PyTorch, Keras, Scikit-learn**; Big Data: **Spark, Hadoop, Hive, Presto SQL**; Cloud Platforms: **AWS, Azure, Google Cloud**; Data Visualization: **PowerBI, Tableau, Matplotlib, Seaborn**; MLOps: **Docker, Kubernetes, MLflow**; Other: **Linux, RESTful APIs, Agile methodologies**;