Satwik Mylavarapu

📞 +1 716-907-8090 | 📍 New York | ✉ satwikmylavarapu@gmail.com | 🔗 linkedin.com/in/Satwik | 🔗 github.com/Satwik

Summary

Data Scientist with 1+ years of experience in ML, AI, and MLOps, specializing in LLMs, NLP, and scalable ML pipelines. Proficient in AWS SageMaker, Kubernetes, and deploying ML models in production. Experienced in optimizing model performance and building AI-driven solutions that drive business impact.

Experience

Data Science Research Assistant
August 2024 – November 2024

University at Buffalo — *Data Science and Machine Learning*
Buffalo, NY

- Increased forecasting accuracy from 78% to 94% by developing ML models (Python, Scikit-learn, TensorFlow) on AWS SageMaker.
- Built an MLOps pipeline for automated model retraining using SageMaker, and Kubernetes, ensuring scalable deployment.

Data Science Analyst
October 2022 – June 2023

Data Beat — *Data Visualization, Data Analysis*
Hyderabad, India

- Developed A/B testing models for marketing teams, leading to a 30% increase in revenue and a 65% boost in engagement by optimizing campaign targeting.
- Automated SQL-based financial reporting, reducing manual effort by 70% and cutting reporting time from 3 days to 6 hours, ensuring faster decision-making.
- Designed and deployed real-time Power BI dashboards for KPI monitoring, enabling data-driven insights and reducing business reporting delays by 40%.

Data Scientist
January 2022 – September 2022

Saboo Diagnostics — *Healthcare Analytics*
Hyderabad, India

- Developed and deployed ML models on 50,000+ patient records to identify high-risk disease factors, enhancing early detection and clinical decision-making.
- Optimized an ML pipeline, improving F1-score from 72% to 89% for post-COVID complication classification, leading to a 25% reduction in misdiagnosis rates.
- Developed a Google Analytics dashboard to track patient demographics, disease progression trends, and treatment efficacy, enabling data-driven insights for physicians and reducing patient readmission rates by 25%.

Data Science Intern
October 2021 – December 2021

NSIC — *Data Science and Analysis*
Hyderabad, India

- Analyzed 200,000+ real-time sensor data in PostgreSQL, improving fault detection in generation systems by 25%.
- Extracted and processed large datasets using SQL and Python, enabling efficient forecasting and operational analytics.

Projects

Big-Star Airbyte ETL Project
GitHub

- Developed an ETL pipeline with Airbyte, PostgreSQL, BigQuery, and Redshift, automating data transformations with dbt.

USA Unemployment Analysis Dashboard
GitHub

- Built a Tableau dashboard to visualize unemployment trends across the USA, enabling insights into job market fluctuations.

Real-time Cryptocurrency Monitoring Dashboard
GitHub

- Built a real-time Streamlit app for crypto tracking using REST APIs and web scraping, improving data retrieval speed.

Credit Card Fraud Detection
GitHub

- Processed 300,000+ transactions for fraud detection using AWS SageMaker and S3, deploying an XGBoost model via a SageMaker Endpoint. Achieved 92.14% accuracy and a 94.05% ROC-AUC, with real-time predictions via a Streamlit app.

Technical Skills

Programming: Python, SQL, R, PySpark
Data Engineering: Apache Spark, Hadoop, Hive, Airbyte, dbt, Snowflake, Microsoft SQL Server, MongoDB
Cloud & MLOps: AWS (EC2, S3, ECR, SageMaker), GCP (BigQuery, Cloud SQL), Kubernetes (EKS), Docker, Azure AI
Machine Learning & AI: TensorFlow, PyTorch, Transformers, Deep Learning, NLP, Generative AI, Model Evaluation (F1-score, RMSE, AUC), AutoML, Reinforcement Learning
MLOps & Model Deployment: AWS SageMaker Pipelines, MLflow, FastAPI, Flask, Feature Engineering, A/B Testing, Streamlit
Data Analytics & Visualization: Tableau, Power BI, Looker, Plotly, D3.js, Statistical Modeling, Hypothesis Testing

Education

University at Buffalo - SUNY
August 2023 – January 2025

Master of Science in Data Science
Buffalo, NY

Certifications

AWS Cloud Practitioner Certified
Power BI Job Simulation – PWC Forage
Oracle Generative AI Certified
Microsoft Azure AI Essentials Professional Certificate by Microsoft and LinkedIn
Data Science Professional Certificate by KNIME - LinkedIn