

Thathsara Rajapaksha

Homagama, Sri Lanka | ntratofficial@gmail.com | tel: +94706161762 | linkedin.com/in/Thathsara-Rajapaksha
github.com/NTRajapaksha | Portfolio : NTRajapaksha

Profile

Data Science graduate with **First Class Honours** and hands-on experience as an **AI Engineer**, specializing in machine learning, computer vision, NLP, and MLOps. Skilled in developing production-ready ML models, building scalable data pipelines and ETL workflows (Spark, Kafka, Hadoop), and creating interactive BI dashboards with Power BI and Tableau. Experienced in agile environments with end-to-end ML project delivery.

Skills

Programming Languages: Python, SQL, R, Java, HTML, CSS, JavaScript, Excel

ML/AI Frameworks: TensorFlow, PyTorch, Scikit-learn, MLFlow, Hugging Face, LangChain, PySpark, Keras

Data Visualization & BI: Power BI, Tableau, Looker, Matplotlib, Seaborn, Plotly

Cloud Infrastructure: AWS (S3, SageMaker, Lambda, Redshift, SQS), GCP (Vertex AI, BigQuery), Azure (Azure ML, Databricks), Docker, Kubernetes

Big Data & Distributed Computing: Apache Spark, Hadoop, Kafka, Airflow, Snowflake, MS SQL Server

Development Tools: Jupyter Notebook, Git, Docker, VS Code, Anaconda, Streamlit, Flask

Education

BSc (Honours) in Data Science - UGC Approved

National School of Business Management (NSBM) - Homagama, Sri Lanka

- **First Class Honours** (GPA: 3.76/4.0)
- **Coursework:** Machine Learning, Deep Learning, Big Data Analytics, Statistical Methods, Data Visualization, Cloud Computing, Natural Language Processing

GCE Advanced Level (Physical Science Stream - Mathematics)

Mahinda Rajapaksha College - Homagama, Sri Lanka

- Com. Maths : C , Physics : S, Chemistry: S, English : B

Experience

AI Engineer Intern

Feb 2025 - Aug 2025

Codebell PVT LTD, Homagama, Sri Lanka

- Developed end-to-end ML pipelines spanning computer vision (CNN transfer learning), NLP (BERT/DistilBERT for text classification, Sentiment Analysis), and traditional ML (XGBoost, hybrid recommenders, contextual bandits) across all production projects achieving nearly 80-90% model accuracy.
- Optimized models for production deployment using quantization (40% size reduction), distillation, caching strategies, and implemented fairness-aware ML with bias mitigation.
- Delivered production-ready models with comprehensive documentation through cross-functional collaboration, demonstrating capabilities in model development, optimization, ethical AI practices, and stakeholder communication.

Customer Relations Officer

Dec 2021 - Apr 2022

LB Finance PLC, Delkanda, Sri Lanka

- Handled financial reporting, customer account management, and client communications, developing strong stakeholder engagement and problem-solving skills in a fast-paced financial services environment.

Projects

IoT Energy Consumption Anomaly Detection with Predictive Analytics

- Developed a Lambda Architecture-based system for analyzing IoT energy consumption data, implementing batch processing with HDFS and Apache Spark for historical data analysis and real-time anomaly detection using Kafka and Spark Streaming.

- Implemented Linear Regression models with Spark MLlib for energy consumption forecasting and statistical methods (3-sigma rule) for real-time anomaly detection, with results visualized through a Flask-based web dashboard.
- Tools Used: Apache Hadoop, Apache Spark, Apache Kafka, Spark MLlib, Flask, Python, Docker

Scalable Employee Directory Application on AWS

- Built a cloud-based employee directory application with automatic scaling and high availability across multiple data centers to ensure zero downtime and handle varying user loads efficiently.
- Implemented secure architecture with database storage, cloud monitoring, and access controls to protect sensitive employee information while maintaining fast performance and reliability.
- Tools Used: AWS (EC2, Application Load Balancer, Auto Scaling, VPC, DynamoDB, S3, CloudWatch, IAM)

Credit Card Fraud Detection - End-to-End MLOps Pipeline

- Engineered a robust MLOps workflow for credit card fraud detection, utilizing MLflow for full lifecycle management including automated experiment tracking, metric comparison, and centralized model registry.
- Implemented a containerized deployment strategy using Docker and MLflow Aliases (e.g., @production), effectively decoupling model training from inference logic to ensure reproducible production environments.
- Tools Used: Python, MLflow, Docker, Scikit-learn, Flask

Colombo Apartments Pricing Data Analysis

- Conducted comprehensive exploratory data analysis on Colombo's rental apartment market, analyzing 250+ property listings to identify key price determinants including size, location, and amenities.
- Applied statistical methods to detect outliers, analyze correlation patterns, and develop data visualizations that revealed unexpected pricing trends, such as the counterintuitive relationship between furnished status and rental prices.
- Tools Used: Python, Pandas, NumPy, Matplotlib, Seaborn, Jupyter Notebook

Revenue Forecasting Tool

- Developed and deployed a Streamlit-based web application that uses a LSTM model to predict stock prices for multiple brands (Google, Apple, Samsung), providing actionable insights for financial planning and market analysis.
- Tools Used: Python, Jupyter Notebook, Scikit-learn, Streamlit

Collaborative Dashboard Development for Business Decision-Making

- Translated the business requirements into interactive dashboards using Power BI and Tableau, enabling real-time tracking of key performance indicators (KPIs). Improved decision making efficiency for stakeholders by delivering intuitive and actionable visualizations.
- Tools Used: PowerBI, Tableau

Additional Courses

MLOps Specialization - Duke University

Microsoft AI & ML Engineering Professional Certificate - Microsoft

Google Advanced Data Analytics Professional Certificate - Google

Google Business Intelligence Professional Certificate - Google

Advanced Machine Learning on Google Cloud Specialization - Google

Mathematics for Machine Learning and Data Science Specialization - DeepLearning.AI

Microsoft Power BI Data Analyst Professional Certificate - Microsoft

Deep Learning Specialization - DeepLearning.AI

Machine Learning Specialization - Stanford University

AWS Fundamentals Specialization - AWS

*References available upon request