

Imalsha Jathunarachchi

Nugegoda, Colombo, Sri Lanka | +94 76 238 1520 | imalsha.contact@gmail.com

LinkedIn: in/imalsha-jathunarachchi/

GitHub: Imalsha-GitHub

Profile

Data Engineer with hands-on experience in designing, building, and optimizing scalable ETL/ELT pipelines using Databricks, PySpark, SQL, and Delta Lake. Skilled in data ingestion, transformation, and data quality automation. Strong foundation in Python and full-stack development, with a focus on cloud-native, data-driven solutions.

Education

University of Colombo School of Computing B.Sc. in Computer Science	2023 – 2026
Ruhunu Vijayaba College G.C.E A/Ls 2021 - Combined Mathematics: A , Chemistry: B, Physics: C G.C.E O/Ls 2017 - 7A, 1B, 1C	2019 – 2021

Professional Experience

Associate Data Engineer – Algospring 2026 June – 2026 September

Data Pipeline Development & Optimization

- Designing and maintaining scalable ETL pipelines using Databricks to process and transform large-scale datasets.
- Leveraging PySpark and SQL within the Databricks environment to perform data cleaning, integration, and performance tuning for downstream analytics.
- Implementing automated data quality checks and monitoring workflows to ensure high data reliability and availability. **(Databricks, PySpark, SQL, Python, ETL)**

Software Engineering Intern – Altria Consulting 2025 November – 2026 May

Document Management & Admin Dashboard

- Developed a full-stack Next.js and React application to manage secure document uploads and back-office data validation.
- Managed version control via GitHub and deployed the system to Azure Web App Service.

Insurance Portal RPA Development

- Engineered web automation scripts using Python and Playwright to extract data and handle workflows across multiple major insurance platforms.
- Managed complex XPath locator extraction and executed comprehensive test scenarios to ensure reliable automated data entry.

Healthcare & Motor Endorsement Automation

- Automated medical and motor insurance endorsement processes, streamlining repetitive data entry tasks for healthcare portals.
- Fixed critical frontend validation errors and resolved production bugs to maintain stability across various medical portal environments.

(Python, Playwright, TypeScript, React, Next.js, GitHub, Azure Web App Service, SQL)

Personal Projects

CSE Data Pipeline – Automated Market & Financial Data Platform 2026 July - 2026 August

- Built an automated data pipeline that ingests daily trading data for 283 Sri Lankan listed companies via scheduled Databricks jobs using a medallion architecture, and extracts historical financial statements from annual report PDFs using AI-powered extraction (Claude) into structured JSON, loaded into Delta tables for analysis. **(Python, Databricks, Delta Lake, PySpark/Pandas, Claude AI, PDF/md processing)**

Project Intelligate – AI Assistant Platform

2025 July - 2025 August

- Built a cost-effective AI assistant platform that allows organizations to provide ChatGPT-like functionality to all users through a single shared API subscription. Features role-based access, personalized workflows, and centralized knowledge management, eliminating the need for individual premium AI subscriptions.
(React, TypeScript, Python FastAPI, OpenAI GPT-4o API, Firebase, Tailwind CSS)

Digifine – Traffic Fine Management System

2024 May - 2025 April

- Developed DigiFine, a digital platform that simplifies how fines are issued and managed between law enforcement and drivers. It enables officers to issue fines digitally, notifies drivers via email, and supports secure online payments, ensuring accurate tracking, real-time updates, and reduced paperwork and errors.
(PHP, HTML, CSS, JavaScript, MySQL)

Licenses and Certifications

Get Started with Databricks for Data Warehousing - Databricks

Completed a foundational Databricks course covering data warehousing concepts and core platform capabilities.

Databricks Fundamentals Accreditation - Databricks

Earned accreditation demonstrating core knowledge of the Databricks platform and its fundamental features.

Python for Beginners - University of Moratuwa

Completed a foundational Python course covering core programming concepts, syntax, and basic problem-solving.

Skills

Cloud Warehouses & Databases: Databricks, Snowflake, PostgreSQL, MySQL, MS SQL Server, MongoDB

Programming Languages: Python, Advanced SQL, Scala, Java, C++

Data Engineering & Processing: ETL / ELT Architecture, PySpark, Pandas, Data Ingestion, Data Transformation

Data Architecture & Concepts: Data Warehousing, Data Modeling (Star Schema), Query Optimization, Data Validation

Tools & Infrastructure: Git, GitHub Workflows, Docker, Linux (Bash)

Extra Curricular Activities

Exploration Club of UCSC

2024 – 2026

- Served as President (2025–2026)
- Organized Book Donation Projects
- Organized seminars for schools
- Organized an ICT seminar series for Grade 10–11 students, covering Arduino and ICT sessions

Rotaract Club of UCSC

2023 – 2026

- Served as Sergeant-at-Arms (2025–2026)
- Member – Sports and Recreational Avenue (2023–2024)
- Organized Basketball Champion League 2024

Volunteer Experience

2024

- Volunteer – UCSC Career Fair 2024

Leadership Experience (School)

2016 – 2017

- Senior Prefect
- Army Cadet – Lance Corporal

References

Mr. Anusha Samarakkodi

Senior Software Engineer - h2o.ai
anusha.samarakkodi@h2oai.com
+94 77 733 4187

Mr. Kenneth Thilakarathna

Lecturer - University of Colombo(UCSC)
kmt@ucsc.cmb.ac.lk
+94 76 310 5987