

Yovin Ransika Yahathugoda (BEng, MSc) | Data Scientist

Mobile: +44 7932614336 | Email: yovin.yahathugoda@gmail.com

LinkedIn Profile - <https://www.linkedin.com/in/yovin-yahathugoda/> | Digital Portfolio - <https://www.yovinyahathugoda.com/>

Kaggle Profile - <https://www.kaggle.com/yovinyahathugoda/competitions>

GitHub Profile - <https://github.com/yoviny/Deep-learning-Competitions>

Professional Experience

Roche, United Kingdom

06/2022 – Present

Data Scientist Industrial Placement

- Contributed to a team implementing a **federated learning** architecture to unify Multiple Sclerosis segmentation algorithms (**PyTorch**, **TensorFlow**, Matlab, **ONNX**) across different sites for 3D MRI images to improve model performance and enhanced data privacy while increasing access to medical data. Responsibilities included algorithm testing, data curation, and pipeline construction.
- Implemented an **unsupervised algorithm**-based (clustering) solution for Roche studies, enhancing reusability and reducing costs across similar countries.
- Leveraged **NLP** and **Graph models** (Neo4J & NetworkX) to predict clinical study similarity and identify missing NCT IDs, facilitating accurate internal data source matching for missing studies.
- Developed a PDF parsing pipeline for drug protocol documents, utilizing **OCR** and NLP to extract data, which was stored in **Snowflake** in a semi-structured format for indexing and subsequent machine learning tasks.
- Experienced in writing **reproducible code**, adhering to standards via **version control**, and utilizing **GitLab**.
- Utilized Agile methodologies in a team setting to develop Minimum Viable Products (MVPs) and Proof-of-Concepts (PoCs) using **Python** (**NLTK**, **Scikit-learn**) and machine learning techniques (**K-means**, **Node2Vec**, **multi-label classification**), leveraging existing data to deliver business solutions and evaluate ideas from business stakeholders using a software development life cycle.
- Developed and deployed RS Connect-based **dashboards** for data **visualization** and querying, facilitating dataset exploration and narrative creation for non-technical users.
- Organized and hosted data science workshops and hackathons promoting STEM inclusion and social mobility.

John Keells Research (A subsidiary of John Keells Holdings PLC), Sri Lanka

04/2018 – 11/2019

Research Engineer (Computer Vision & Electronics)

- Enhanced lab experiment efficiency by employing ML **classifiers** for microbial colony detection and **segmentation**, facilitating precise surface area calculations.
- Utilized **image processing** for multispectral image analysis to identify latent stress fractures in wind turbine blades, optimizing maintenance and repair scheduling.
- Developed a **computer vision**-based 3D printing monitor that detects print failures, effectively reducing material waste.
- Contributed to team brainstorming for innovative, marketable ideas by conducting prior art searches, validating concepts through proof-of-concept product development, and fostering potential intellectual property creation.

Education

MSc Data Science (Expected Grade – Distinction)

09/2021 – 10/2023

Nottingham Trent University, UK

Received a **25% scholarship** and additional **£3,000 bursary** for previous academic performance.

Key modules: Software Engineering, Statistical Data Analysis and Visualisation, Applied AI and Data Mining, Deriving Business Value from Data Analytics, and Fundamentals of Big Data and its Infrastructure.

MSc thesis

10/2022 - Present

Focusing on developing a **novel** approach to combine knowledge graphs with transformer-based language models specializing in the financial domain.

BEng (Hons) Electronic Engineering - 1st Class Honours (74%)

09/2014 – 01/2018

Sheffield Hallam University, UK

- Thesis title: Navigation System to Assist the Visually Impaired using Stereo Vision and Audio Feedback - This included Computer Vision and Machine Learning using Python, OpenCV and LIBSVM.

Core Competencies

Programming - Python, PyTorch, Linux, Pandas, Jupyter, OpenCV, **Git, TensorFlow, C, C#, C++, Streamlit, R**, NetworkX

Software – Microsoft Office Suite, Visual Studio, Arduino, Eagle PCB, OrCAD, R Studio, Matlab, Neo4J

Languages – English (Fluent) & Sinhala (Native)

Transferable skills

- **Communication** - Explaining technical presentations in a simple and concise manner for allowing non-technical team members (D-suite) to understand. Leading and collaborating in design workshop sessions at Roche and presenting at data science workshops.
- **Time management** - Demonstrated commitment to deadline adherence and weekly goal attainment across multiple projects as a Data Scientist at Roche and Research Engineer at JKR, leveraging prioritization skills and proficiency in sprint methodologies and frameworks such as JIRA and Trello.
- **Leadership** - Spearheaded and executed data science hackathons within the Roche Advanced Analytics Network (RAAN) through collaboration with colleagues. Led postgraduate and undergraduate teams, ensuring timely completion of all project deliverables. Represented the undergraduate cohort in the final academic year.
- **Project management** - Managed procurement of essential equipment and components for research projects at John Keells Research, adhering to a stringent budget and conducting weekly inventory assessments.
- **Problem Solving/Critical Thinking** – Thinking ‘outside the box’ to solve novel problems and improve business processes during the placement at Roche. Providing solutions to problems faced while working on novel research projects at John Keells Research.

Deep Learning Projects Achievements

- Full-time Kaggle participant involved in Data Science and Deep Learning competitions in areas such as medical imaging between December 2019 to August 2021.
- Implemented a custom ResNeXt model with CutMix augmentation for OCR of Bengali handwritten characters and achieved a leader board ranking in the **top 3%**
- Developed a custom U-Net model ranking in the **top 16%** for detecting functional tissue units (FTUs) in Kidney biopsies.

Memberships

Institution of Engineering and Technology (IET) – Member (MIET)

2018 - Present

Institute of Electrical and Electronics Engineers (IEEE) - Member

2016 – Present

Certifications

[Azure Data Fundamentals](#) – Microsoft

2022

Industry Recognised Courses

[Big Data Specialisation](#) - University of California San Diego

2020

[Deep Learning Specialisation](#) - By Andrew Ng (Deep Learning platform / Coursera)

2020

[DL0101EN: Deep Learning Fundamentals with Keras](#) - IBM

2019

[PH125.8x: Data Science: Machine Learning](#) - HarvardX

2019

Interests & Achievements

Extra-Curricular Activities

- Volunteering – John Keells Foundation (CSR), Buddhist Society of Lyceum International School Nugegoda
- President of Aeronautical Club at Lyceum International School Nugegoda.
- Shortlisted in two categories in the [Astronomy Photographer of Year 2021](#) competition hosted by Royal Museums Greenwich, UK. Nominated for the People’s Choice Award in the same competition.

References

References available upon request.