

# Rivindu Ashinsa

Data Science Intern

Location: Western Province, Sri Lanka Email: ashinsa.rivindu@gmail.com Phone: +94 752 691 645

Website: [rivindu-ashinsa.github.io](https://rivindu-ashinsa.github.io) LinkedIn: rivindu-ashinsa GitHub: rivindu-ashinsa

## Professional Summary

Computer Science undergraduate with hands-on experience in data cleaning, analysis, and visualization. Proficient in Python, SQL, and data reporting with proven ability to support data-driven decision-making. Experienced in processing and analyzing large datasets, creating actionable insights, and building data pipelines. Strong analytical and problem-solving skills with expertise in statistical analysis and data interpretation.

## Technical Skills

**Programming Languages:** Python, SQL (MySQL, MariaDB), R, Java

**Data Analysis & Cleaning:** Pandas, NumPy, SciPy, Data Wrangling, Data Preprocessing, Feature Engineering, Statistical Analysis

**Data Visualization & Reporting:** Matplotlib, Seaborn, Power BI, Data Interpretation, Dashboard Creation

**Machine Learning Tools:** scikit-learn, TensorFlow, Keras, XGBoost

**Development Tools:** Git, GitHub, Jupyter Notebook, VS Code, Streamlit, Linux

**Key Competencies:** Data Cleaning, Data Analysis, Data Visualization, Problem-Solving, Analytical Thinking, Decision Support, Statistical Methods

## Education

### B.Sc. (Hons) Computer Science

University of Westminster, United Kingdom

Sep 2024 – Jun 2027

(Expected)

- Relevant Coursework: Data Science, Machine Learning, Statistics, Database Systems, Algorithms and Data Structures, Artificial Intelligence

### Foundation in Information Technology

Informatics Institute of Technology (IIT), Sri Lanka

Sep 2023 – May 2024

- Key Skills: Python Programming, SQL Database Management, Data Analysis, Statistical Methods

## Relevant Experience

### Data Scientist – Software Development Group Project

University of Westminster (CognivusLabs Project)

2025 – Present

- Performed extensive data cleaning and preprocessing on 10,000+ biomedical signals using Python and SciPy
- Conducted statistical analysis and data interpretation to identify patterns in clinical datasets
- Created data visualizations and comprehensive reports to support research decision-making
- Analyzed time-series data with 3,000+ clinical cases, generating actionable insights for medical teams
- Developed automated data pipelines reducing manual data processing time by 60%

### Data Analyst – Dementia Risk Prediction Project

XPredators Research Team

2025

- Cleaned and analyzed 5,000+ patient records using Python, SQL, and Pandas
- Performed exploratory data analysis (EDA) and statistical testing to identify key health indicators
- Created data visualizations using Matplotlib and Seaborn to communicate findings to stakeholders
- Reduced data dimensionality by 35% through feature selection and data cleaning techniques

- Supported data-driven decision-making by providing analytical insights and recommendations

### **Data Analyst – Heart Disease Prediction**

2025

#### Personal Research Project

- Analyzed and cleaned 1,000+ patient medical records using Python and SQL
- Performed statistical analysis to identify 12 predictive biomarkers for heart disease
- Generated comprehensive data reports and visualizations for medical interpretation
- Conducted data quality assessment and implemented data validation procedures

## **Additional Projects**

---

### **RepoDigest – AI-Powered Code Analysis (LangGraph & GenAI)**

2025

- Built GenAI application using Python, LangGraph, and OpenAI API for automated GitHub analysis
- Extracted and processed data from 500+ repositories, analyzing 50,000+ lines of code per repository
- Implemented AI-driven data extraction pipelines reducing documentation time by 40%
- Applied natural language processing for structured insight generation from code

## **Achievements & Recognition**

---

- **IntelliHack 5.0 (IEEE UCSC):** Top 10 finalist out of 350+ teams for data-driven predictive solutions
- **DataStorm 6.0 (Rotaract):** Top 10 finalist for business analytics and data visualization competition
- **University Group Project:** Developed biomedical data analysis pipeline for CognivusLabs (Westminster SDGP)
- **DataCamp Certification:** Data Scientist Associate – Validated data analysis and Python skills
- **Multiple Data Science Certifications:** IBM, Microsoft, Google, Kaggle (Data Cleaning, Analysis, Visualization)
- **Academic Excellence:** 9 A grades in GCE Ordinary Level Examination

## **Professional Certifications**

---

- DataCamp: Data Scientist Associate Certification
- IBM CognitiveClass.ai: Data Analysis with Python, Python for Data Science
- Kaggle: Data Cleaning, Pandas, Data Visualization, SQL
- FreeCodeCamp: Data Analysis with Python
- Microsoft Learn: Machine Learning Fundamentals, Azure AI Fundamentals
- Google: Machine Learning Crash Course
- LinkedIn Learning: Data Science Fundamentals, Applied Machine Learning

## **References**

---

### **Seyed Ruzaik**

*Assistant Lecturer*

*Informatics Institute of Technology*

Email: ruzaik.s@iit.ac.lk

Phone: +94 76 233 4461

### **Thuan Pakeer**

*Associate AI Research Assistant*

*CurveUp*

Email: naheempakeer@gmail.com

Phone: +94 76 856 0394