Cletus Xavier S.
Data Analyst

📱 +94 75 496 0730 ✉ cletusxavier96@gmail.com 🏠 Batticaloa, Sri Lanka

🔗 in/cletus-xavier ⭘ github.com/xcletus3

Profile

Data Analyst with over 2 years of experience in both data analysis and machine learning. Skilled in Python, MySQL, and AWS, I have built and deployed machine learning models, automated data pipelines, and created dashboards to support business decisions. While my role focused on data analysis, I also worked on tasks like developing ML models and using advanced data science techniques. Currently pursuing an MSc in Data Science and AI to further enhance my skills in machine learning and AI. I enjoy turning complex data into clear, actionable insights and solving real-world problems.

Areas of Expertise

- Machine Learning & AI (Deep Learning, NLP, Computer Vision)

- Data Analysis & Visualization

- Statistical Modeling

- Automation & Web Scraping

- GNU/Linux

Professional Experience

Data Analyst, *(LankaPropertyWeb)*   **Colombo 03, Sri Lanka** *06/2022 - present*
- Automated data pipelines using MySQL, Python, and AWS (EC2, Lambda), eliminating manual works and reducing processing time.

- Developed, deployed and maintained machine learning models for property price prediction and image recognition, improving business insights and decision-making.

- Built and maintained interactive dashboards in Looker Studio, integrating MySQL and Google Analytics data, aiding data-driven decisions across departments.

- Created web scraping tools using Python (Selenium, BeautifulSoup) to gather and analyze real estate data from public sources.

- Contributed as a Data Analyst to the Sri Lanka Real Estate Market Outlook Reports (**2023** & **2024**) with data analysis and visualizations to support key market insights.

Education

- **MSc in Data Science and AI (Reading)** *University of SriJayewardenepura*   **Sri Lanka** *2023-2025*

  Relevant Courses: Mathematics & Programming for Data Science, Machine Learning (I & II), Artificial Intelligence, Data Modeling, Spatio-Temporal Data Analysis, Business Intelligence, Bayesian Computation, Computational Linguistics.

- **BSc in Industrial Statistics and Mathematical Finance (General)** *University of Colombo* **Sri Lanka** *2019-2022*

  Relevant Courses: Mathematics (Linear Algebra, Calculus I–III), Statistics, Probability & Distributions, Regression Analysis, Applied Time Series, Statistical Modeling, DataBase Systems, Python Programming and Operational Research.

Online Courses & Certifications

- AutoML Automated Machine Learning BootCamp (No Code ML) - Udemy

- Deep Learning with TensorFlow 2 - 365DataScience

- Convolutional Neural Networks with TensorFlow in Python 365DataScience

- 20 Machine Learning and Deep Learning Projects in Python - Udemy

Projects

Price Meter & House /Land Price Index

- Developed a Python script to summarize local averages of house and land prices for cities, districts, and provinces for 'Price Meter' and accumulate averages for 'House and Land Price Index'

- The python script extracted data from a MySQL database, cleaned it, and calculated averages, automating the entire process.

- Implemented a custom outlier detection method and applied sampling distribution techniques to address the skewed nature of real-world price distributions, resulting in more reliable and accurate price estimations.

Similar Images

- Built a Python-based image processing tool to identify visually similar images, using OpenCV and feature extraction techniques.

- Deployed the tool on AWS EC2 with Lambda triggers, automating image classification for improved efficiency in processing large image datasets.

Image Recognition and Tagging

- Developed a machine learning model using TensorFlow to classify room types (Bedroom, Bathroom, Living room, etc.) from images, achieving 90% classification accuracy.

- Automated image tagging via AWS Lambda and API integration, enhancing the speed and scalability of the tagging process.

Skills

- **Languages:** Python, SQL (MySQL, NoSQL), R

- **Data Analysis & Visualization:** Pandas, NumPy, Matplotlib, Seaborn, Tableau, Looker Studio

- **Machine Learning & AI:** TensorFlow, Keras, Scikit-learn, PyTorch, Hugging Face, YOLO (Ultralytics), PaddleOCR, OpenCV, PIL (Pillow), Threading, Concurrent Futures

- **Cloud Technologies:** AWS (EC2, SageMaker, Lambda, S3, EventBridge)

- **Version Control:** Git, GitHub

- **Other Tools:** Google-AppsScript, Google Spreadsheets/Excel, C++, MatLab, Shell Scripting, Docker, BeautifulSoup, Selenium

- **Soft Skills:** Critical Thinking, Problem-Solving, Collaboration, Adaptability, Analytical Thinking

Languages

- **English** [Full Professional Proficiency]

- **Tamil** [Native Proficiency]

- **Sinhala** [Limited Working Proficiency]

References

Dr. A.A. Sunethra
Senior Lecturer,
Department of Satistics,
University of Colombo.
+94 11-259 0111
sunethra@stat.cmb.ac.lk

Mr. T. Jayarathne
Doctoral Candidate, MBA (UK), BA (Hon) (UK), Chartered MCSI (UK), MIPA (Aus), AFA (UK), CFIP (USA), IIA (USA), Adv.Dip.MA (UK) SAT (SL))
Head of Development Consultancy and Research,
Lanka Property Web (Pvt) Ltd.
+94 (0) 77 777 3051 (WhatsApp)
tharindujayarathne@yahoo.com