

Abdhurrahman Ashfarul

Data Scientist

Kandy, Sri Lanka

+94760510714

abdhurrahman360@gmail.com



LinkedIn: <https://www.linkedin.com/in/abdhurrahman-ashfarul-9969851b5/>

Github: <https://github.com/abdhurrahmanashfar>

SUMMARY

Passionate and adaptable individual with a strong interest in data science. Completed comprehensive coursework and obtained certifications. Seeking hands-on professional internship experience to further enhance skills and knowledge in the dynamic field of data science. Committed to continuous growth and development.

CORE COMPETENCIES

- ▶ Strong Analytical Skills
- ▶ Problem-Solving Skills
- ▶ Adaptability
- ▶ Attention to Detail
- ▶ Time Management
- ▶ Emotional Intelligence
- ▶ Creativity

SKILLS

- ▶ **Programming Languages:** Python, SQL
- ▶ **Tools:** Git, PgAdmin4
- ▶ **Libraries/Frameworks:** Numpy, Pandas, Scikit-learn, TensorFlow
- ▶ **Data Visualization:** Matplotlib, Seaborn
- ▶ **Version Control:** Git, Github
- ▶ **Databases:** PostgreSQL
- ▶ **Data Formats:** JSON, CSV, XLSX, YAML
- ▶ **Machine Learning:** Linear Regression, Classification, Neural Networks

EDUCATION

- ▶ **Esoft Uni** (September 2025 – Present)
 - Bachelor of Science (B.Sc.) in Data Science
 - Expected Graduation: 2026 (Final Year Student)
- ▶ **Coursera** (December 2023 – Present)
 - Mathematics for Machine Learning (Linear Algebra)-Imperial College London
 - Deep Learning Specialization by DeepLearning.ai institute
 - IBM Data Science Methodology
 - IBM Generative AI: Introduction and Applications
- ▶ **AiCore** (November 2022 - November 2023)
 - Graduated as a Data Scientist at AiCore.
- ▶ **Udemy** (May 2021 - May 2022)
 - Completed 20 business case projects using Machine learning.
- ▶ **365 Data Science** (April 2020 - April 2021)
 - Completed 26 Data Science related courses.
- ▶ **G.C.E (O/L) -Srilanka English Medium** (2019)

EXPERIENCE

▶ Data Science Intern – Codveda Technologies (April 2025 – May 2025)

- Completed 9+ structured tasks across web scraping, preprocessing, visualization, machine learning, NLP, and deep learning.
- Built and evaluated models using Linear/Logistic Regression, Random Forest, SVM, K-Means, and ARIMA.
- Created an end-to-end NLP pipeline using TF-IDF and Word2Vec, achieving high classification accuracy.
- Developed a CNN in TensorFlow/Keras for handwritten digit recognition (MNIST), reaching 98%+ accuracy.
- Visualized insights using Matplotlib, Seaborn, and trained models with hyperparameter tuning.

Tools: Python, pandas, scikit-learn, TensorFlow, matplotlib, NLTK, Word2Vec, BeautifulSoup

▶ Data Scientist trainee at analyticsguard.ai (Apr 2024 - present)

- Company Overview: analyticsguard.ai is a dynamic startup leveraging cutting-edge technologies to enhance cybersecurity. Tools like Databricks, Fivetran, Google Drive, and Power BI are integral to our operations.
- Predictive Modelling and Anomaly Detection: Developing and testing machine learning models on datasets from CIDDS and University of Brunswick to detect real-time cybersecurity threats and anomalies using Databricks, significantly reducing risk exposure.
- Data Analysis and Strategic Insights: Analysing and extracting strategic insights from datasets to develop new cybersecurity products, employing Databricks, Fivetran, and Power BI.
- Visualization and Reporting: Designing and delivering data visualizations and reports using advanced analytics tools and planning to communicate complex cybersecurity trends to stakeholders.
- Cross-Functional Team Collaboration: Collaborating with other team members using slack, outlook and using Google Drive for effective team coordination.
- Research and Development in Cybersecurity Trends: Engaged in ongoing research, referencing studies like Flow-based Benchmark Data Sets for Intrusion Detection and Attack Vectors in Healthcare Devices - A Multi-Protocol Dataset for Assessing IoMT Device Security, to stay ahead of emerging cybersecurity techniques where the data of these studies are being used for testing our ML Models

▶ BCG GenAI Job Simulation on Forage (April 2024)

- Completed a job simulation involving AI-powered financial chatbot development for BCG's GenAI Consulting team.
- Gained experience in Python programming, including the use of libraries such as pandas for data manipulation.
- Integrated and interpreted complex financial data from 10-K and 10-Q reports, employing rule-based logic to create a chatbot that provides user-friendly financial insights and analysis.

▶ BCG Data Science Job Simulation on Forage (January 2024)

- Completed a customer churn analysis simulation for XYZ Analytics, demonstrating advanced data analytics skills, identifying essential client data and outlining a strategic investigation approach.
- Conducted efficient data analysis using Python, including Pandas and NumPy. Employed data visualization techniques for insightful trend interpretation.
- Completed the engineering and optimization of a random forest model, achieving an 85% accuracy rate in predicting customer churn.
- Completed a concise executive summary for the Associate Director, delivering actionable insights for informed decision-making based on the analysis.

▶ British Airways Data Science Job Simulation on Forage (January 2024)

- Completed a simulation focusing on how data science is a critical component of British Airways success.
- Scraped and analyzed customer review data to uncover findings.
- Built a predictive model to understand factors that influence buying behaviour

▶ Cognizant Artificial Intelligence Job Simulation on Forage (January 2024)

- Completed a job simulation focused on AI for Cognizant's Data Science team.
- Conducted exploratory data analysis using Python and Google Colab for one of Cognizant's technology-led clients, Gala Groceries.
- Prepare a Python module that contains code to train a model and output the performance metrics for the Machine Learning engineering team.

- Communicated findings and analysis in the form of a PowerPoint slide to present the results back to the business.
-

PROJECTS

▶ Machine Learning & NLP Projects (Codveda Internship)

- **NLP Spam Detection:** Built end-to-end pipeline using TF-IDF, Word2Vec, Logistic Regression, achieving high F1-score.
- **MNIST Digit Classifier (CNN):** Designed and trained convolutional neural network in Keras, reaching 98%+ test accuracy.
- **Web Scraper & Data Cleaning:** Scraped tabular data using BeautifulSoup and cleaned using pandas for model-ready format.
- **ARIMA Time Series Forecast:** Decomposed and forecasted real-world time series with RMSE evaluation.

▶ Modelling Airbnb Listings

- Processed and cleaned dataset using Pandas to improve data quality.
- Visually analysed the dataset to understand it better.
- Performed feature selection to understand different features affecting the target columns.
- Trained, compared and evaluated machine learning models (Random Forest, Linear/Logistic Regression, XGBoost etc) for classification (determining different Airbnb categories) & regression (predicting tariff) use cases, reducing errors by 20%.
- Performed hyperparameter tuning and cross validation to optimise the results for particular metrics, such as precision in classification.

▶ Multinational Data Centralisation

- Developed a system that extracts retail sales data from 5 different data sources; PDF documents; an AWS RDS database; RESTful API, JSON and CSV files.
- Created a Python class which cleans and transforms over 120k rows of data before being loaded into a Postgres database.
- Developed a star-schema database, joining 5-dimension tables to make the data easily queryable allowing for sub-millisecond data analysis
- Used complex SQL queries to derive insights and to help reduce costs by 15%
- Queried the data using SQL to extract insights from the data; such as velocity of sales; yearly revenue and regions with the most sales.

▶ Hangman

- Built an implementation of the Hangman game in Python.
- Handled errors in user input and built the conditional logic to run the game.
- Used OOP principles including encapsulation and abstraction.

REFEREES

Ahmed Eldiwany

AI Principal Engineer

AI Core

236 Portobello Road, London W11 1LJ

0204 5364180

ahmed@theaicore.com