

Taissir Boukrouba

Neasden, London, NW10 | taissirboukrouba@outlook.com | +44 7387138628

- **LinkedIn:** <https://www.linkedin.com/in/taissir-boukrouba/>
- **GitHub:** <https://github.com/taissirboukrouba>

PROFILE

Data science master's student at the University Hertfordshire with a BSc in computer systems with shown competence in data modelling, data visualisation, and data wrangling seeking a work experience in data science.

CORE SKILLS & TOOLS

<ul style="list-style-type: none">• Python, R• Machine Learning Modelling• Natural Language Processing• Data Visualisation• Deep Learning• Data Wrangling	<ul style="list-style-type: none">• NumPy, Scikit-learn, Pandas• TensorFlow, Statsmodels, PyTorch• Plotly Express, Seaborn, Matplotlib• Keras, OpenCV• SQL, Scala• Excel
--	---

EDUCATION

Msc Data Science, University of Hertfordshire 2023-2024

- Expected to graduate with Distinction
- Key modules: Machine Learning & Neural Networks (First), Data Handling & Visualisation (First), Data Mining & Discovery (First)

Bsc Computer Systems, University of UHBC 2019 – 2022

- Graduated with a Distinction
- Key modules: Algorithms & Data Structures, Statistics, Graph theory
- Dissertation: Plant classification using Vision Transformers
- Best Undergraduate project (Excellent)

RELEVANT EXPERIENCE

Data Science Intern – MajestEYE Jul 2023 – Sep 2023

- Researched about the company/customer's domain (banking)
- Created data dictionary and listed the data model (dimension & fact tables)
- Enhanced invoice image quality using CNNs
- Removed stamp and watermarks using openCV and deep learning algorithms

- Extracted tables and texts from using different OCR methods
- Used Basic Language models for Named Entity Extraction (NER)

Applied Data Science Lab – WQU May 2023 - Jun 2023

- Created an interactive dashboard for a clustering problem (survey of consumer finances)
- Performed exploratory data analysis on housing prices in Mexico
- Predicted earthquakes damages for households in India using decision trees
- Used Gradient Boosting for the Taiwan bankruptcy data
- Applied data wrangling using SQL

Data Science and Business Analytics Intern – The Sparks Foundation Apr 2023 - May 2023

- Performed EDA on SampleSuperstore dataset (Retail)
- Predicted the number of optimum clusters for the Iris dataset
- Used Linear Regression to predict student score percentage based on study hours

Data Science & Big Data Intern – Djezzy Mar 2023 - May 2023

- Explored data dictionary and researched client's domaine (insurance)
- Implemented the big dataset into the Apache Spark Environment
- Solved different data problems (input typos, redundance ..etc) using Scala
- Introduced Fairness into the data using manual and automated Curations
- Created Multiple KPIs that explains the company's financial status and performance
- Deployed the work into a data product as requested by the client (dashboard)
- Presented the data product to the client

RELEVANT PROJECTS

ACADEMIC

- Plant Classification App Using Flutter and Vision Transformers (ViT)

CORPORATE

- **RiskIntel:** An Interactive dashboard that displays multiple financial KPIs
- **EasyInvoice:** An App that automates PDF invoice reading into text and tables using Deep Learning Models

ONGOING

- Structured information retrieval with LLMs for academic research using RAG
- Global Fuel Statistics Dashboard Before & After Russia-Ukraine War
- Virus Genome Classification Using Biopython & NIH Genome data

HOBBIES AND INTERESTS

- Reading & Writing
- Language Learning
- Travelling