**Syed Haris Ali Wasti**

1 Queens Terrace, E13 9AL

07554787649

[Haris.wasti@hotmail.com](mailto:Haris.wasti@hotmail.com)

**PERSONAL PROFILE:**

I'm a Data science and analytics student at the University of Westminster going into year 3. With getting a 70% average grade for year one and year 2. I'm hoping to take a placement opportunity to leverage myself even further in the field of data science. Through multiple projects tackling real-world problems, I have realised how much impact and importance data science can bring to the world. These experiences have only increased my passion to actively work as a data scientist and make a positive impact on the world.

**---------------------------------------------------------------------------------------------------------------------------------------------**

**EDUCATION:**

**BSc (Hons) Data science and Analytics | September 2021 – June 2025 |** University of Westminster

**1st-year modules(done):**

* + Applied Mathematics
  + Software Development 1(Python)
  + Software Development 2(Java)
  + Requirements Modelling
  + Statistical Modelling and Analysis
  + Database Design and Implementation

**2nd-year modules(done):**

* + Data Engineering
  + Data Visualisation and Communication
  + Machine learning and Data mining
  + Data Science Project Lifestyle
  + Business Analytics
  + Algorithms: Theory, Design, and Implementation

**A - Levels | September 2019 - June 2021 |** London Design and Engineering UTC, London

* A Levels**: Maths, Physics, Chemistry**

**GCSEs | September 2014 - June 2019 |** Lister Community School, London

* Maths (7), Science (7,6), English (5,4)

**---------------------------------------------------------------------------------------------------------------------------------------------**

**Projects:**

**Machine learning and data mining, R Studio**

[Github](https://github.com/HarisWasti/Machine-learning)

**Assignments included:**

* Suitably prepare a realistic data set for data mining / machine learning  
  and discuss issues affecting the scalability and usefulness of learning  
  models from that set
* Evaluate, validate and optimise learned models
* Effectively communicate models and output analysis in a variety of  
  forms to specialist and non-specialist audiences

**Data analysis, visualisation narrative and presentation, Python (Pandas), R (GGPlot2, dplyr):**

[Github](https://github.com/HarisWasti/Netflix-Imdb-analysis)

**Assignments included:**

* Using appropriate tools for data visualisation and dashboarding.
* Applying visualisation techniques to discover and investigate patterns and relationships in a Netflix IMDB dataset.
* Making a creative dashboard to answer and find interesting insights in the Netflix dataset.
* Cleaning a dataset to make it more approachable for analysis and visualisation.
* Finding relationships between different variables to be able to find trends such as length of a movie compared to its rating or votes? Maybe if the movie is too long it will be enjoyed less? What about if a movie is too short? What is the ideal length where a movie is most enjoyed?
* Through visualisations and analysis, I was able to find which movies/ shows people tend to enjoy the most on Netflix.

**Implementing a data pipeline, Matlab:**

[**Github**](https://github.com/HarisWasti/bicycle-store)

**Assignments included:**

* Data collection I found a Cycle store dataset for the pipeline
* Data Exploration where I examined the dataset to get a sense of its structure and content.
* Data transformation where I cleaned the data by filling missing values or removing unwanted columns or rows, I also handled outliers and finally making new columns such as age column by using the date of birth column to make the data more readable and easier.
* Finally creating metadata files for mean, median and mode for numerical and categorical variables.

**Create 2 Databases containing Unstructured Data, Matlab:**

[Github](https://github.com/HarisWasti/Unstructured-Data) (For both)

**Content-based Image Retrieval (CBIR):**

**Assignments included:**

* Image collection where I collected dozens of images all clothing to use for CBIR.
* Image preprocessing where I rescaled, renamed and denoised the images to make it easier to work with.
* I then annotated each image with its keywords and descriptions and stored them as metadata.
* Finally, I performed Image feature extraction where I extracted relevant image features such as mean, norm of the channel pixel intensities, grayscale, texture and shape features. I stored them in a JSON file.

**Sentiment Analysis model:**

**Assignments included:**

* Data collection I found a database where employees review their company.
* I cleaned the data by text normalisation, and tokenization, removed punctuation, and full stops and sampled 50 random reviews.
* I then performed text vectorization where I made the text suitable for machine-learning models, I used methods such as bag of words and word embeddings.
* Finally, using the clean reviews I labelled each of them with a sentiment (Positive, neutral, Negative) and saved the results into a metadata document.

**---------------------------------------------------------------------------------------------------------------------------------------------**

**WORK EXPERIENCES:**

**2023 SIA Training (Ilford (Cranbrook House – IG1 4PG)), Ace Education Academy, London**

* Enhanced security protocol understanding, conflict resolution, medical responses and communication skills
* Training done with a team which boosted my communication and teamwork skills
* Techniques for resolving conflicts and fostering positive interactions.
* CPR, chest compressions and defibrillator usage.

**2023 Volunteering Community Helper, London – Mehmood Mirza – 07916344575**

* Delivered food packages to those unable to attend the distribution centre.
* Met and greeted people on arrival at the food bank to create a welcoming atmosphere.
* Comfortably interacted with people with limited use of English.

**---------------------------------------------------------------------------------------------------------------------------------------------**

**ACHIEVEMENTS AND AWARDS:**

* Obtained an average 70% for first and second year of my Data Science and Analytics degree
* ACT Awareness Certificate
* ACT Security Certificate

**INTERESTS / OTHER SKILLS:**

**Programming languages:** Python, R, Matlab, PowerBI, Java, SQL, Excel

**Languages:** English, Urdu, Punjabi

**Interests:** Cycling, Walking, Games

**REFERENCES:**

Available on request