Gurpreet Singh Deol

London | 07379806602 | Gurpreet.S.Deol@outlook.com | LinkedIn | GitHub | Website

PROFESSIONAL SUMMARY

Computer Science teacher with an academic background in MSc Data Science and BSc Physics, and a solid foundation in data analytics, machine learning, and ETL development. Skilled in Python, SQL and Power BI with experience in creating predictive models, building automated data pipelines, designing interactive dashboards for analysis and comprehensive report writing. Professional experience in multiple customer facing roles with projects completed in various industries including finance, public health, logistics, sports and environmental conservation.

EDUCATION

Royal Holloway University of London - Egham, Surrey

MSc Data Science and Analytics: Pass with Merit **Oct 2021 - Oct 2022**

Key modules: Data Analysis, Machine Learning, Artificial Intelligence, Databases and Business Intelligence.

- First-class dissertation: Collected, cleaned and analysed current and historical football data with unsupervised machine learning techniques to identify a tactical shift from direct football to possession based over 13 years.
- Gained expertise in developing machine learning models and experience in Python, SQL, R and MATLAB. Completed several projects using machine learning techniques to build predictive models and uncover insights in complex datasets.

BSc Physics: Second Class Honours (Upper Division), 2:1 **Sep 2018 - Jul 2021**

Key modules: Mathematical Methods, Statistics, Quantum Mechanics, Particle Physics and Energy and Climate Science.

- First-class dissertation: Simulated electrical resistance, current and potential fields in 2D geometries using Python.
- Developed strong analytical and data handling skills by working with datasets ranging from astronomical observations to experimental physics measurements, producing clear and comprehensive analytical reports.

WORK EXPERIENCE

Computer Science Teacher - North Brent School – Wembley Multi Academy Trust - London, NW10 2UF **Sep 2025 - Present**

- Currently a part of one of the highest achieving trusts in the country, teaching Computer Science to over 400 students and using data analysis to uncover strengths and areas of improvement to help support their learning.
- Day to day structure includes planning and delivering lessons to help students across various levels of abilities and ensuring lessons are structured to maximise retention and learning with a strong focus on exam practice.

SKILLS

Programming Languages: *Python, PostgreSQL, R and MATLAB.*
Cloud Tools: *AWS (Lambda, EC2, S3) and Azure.*
Data Libraries: *Pandas, NumPy, Scikit-Learn, Keras, TensorFlow, Geopandas and Psycopg2.*
Reporting & Data Visualisation: *Microsoft PowerBI, Streamlit, Excel, Matplotlib and Seaborn.*
Other Tools: *Git/GitHub, PySpark, Docker, Visual Studio Code, PostGIS, Jupyter Notebook and LaTex.*

PROJECTS

Public Health and Air Quality Dashboard | GitHub | Live Dashboard

- Designed and built an end-to-end data pipeline to track air pollution, weather, and health data across 10 major cities.
- Used Python, Docker, and PostgreSQL to automate data collection from APIs and store results locally.
- Processed WHO health data to explore the impact of pollution on mortality rates.
- Created a Streamlit app to display interactive plots, summary cards, and a map using real-time and historical data.
- Also developed Power BI dashboards to visualise long-term pollution and health trends.

Predicting Illegal Fishing with Random Forest | GitHub | Web Application

- Built a Random Forest model (94% accuracy) to classify fishing activity from vessel data (AIS) based on speed, location and distance from shore and ports.
- Collected vessel data using the Global Fishing Watch API and filtered it with ocean and Marine Protected Zones (MPZ) boundaries using GeoPandas and PostGIS.
- Identified possible illegal fishing by detecting overlaps between predicted fishing activity and protected zones.
- Deployed the Random Forest model via a Streamlit app to automate data fetching, apply predictions in real time, and display vessel tracks on an interactive map with visual indicators for suspected illegal fishing.

Sports Data Analysis: Football and Machine learning | GitHub

- Scraped football data over a period of 13 years, creating a dataset with over 2,000 records and 100+ attributes, cleaned the data and applied feature engineering using Pandas to prepare the dataset for analysis.
- Applied Principal Component Analysis and K-means clustering for a time series analysis to identify tactical shifts and trends in football across the decade, visualising results using Matplotlib.
- Presented findings through a comprehensive report in LaTeX, highlighting potential factors influencing these trends.

BOOTCAMPS AND CERTIFICATES

AI Core Data Analytics Intensive Bootcamp **Dec 2023 - Feb 2024**

- Completed courses focused on EDA, SQL, object-oriented programming, Power BI, Azure and AWS.