MARWIN SOLOMON

M.Solomon2@lse.ac.uk | +44 7436 683661 | LinkedIn

EXECUTIVE SUMMARY

Data Scientist from LSE collaborating with Deutsche Bank to build explainable stochastic models. Have built machine learning projects pertaining to finance, investment and sports analytics to implement acquired mathematical and statistical knowledge.

EDUCATIONAL QUALIFICATIONS AND ACHIEVEMENTS

- **MSc in Data Science - London School of Economics and Political Science** *(Sep '22 - Sep '23)*
  - **Attained Distinction along with Graduate Support Scheme Scholarship awarded to meritorious students**
  - Key Modules: Machine Learning & Data Mining, Distributed Computing for Big Data, Data Analysis & Statistical Methods
- **B.Sc. (Hons) Mathematics - St. Stephen's College, University of Delhi** *(Jul '18 - Jul '21)*
  - Finished with a **CGPA of 9.8/10**, batch **rank of 3** out of 42 students and amongst **Top 1%** across the University of Delhi
  - Key Modules: Probability Theory & Statistics, Mathematical Finance, Differential Equations, Real Analysis, Group Theory
- **Class 12th (A Levels Equivalent) - St. Martin's Diocesan School, Delhi Cantt** *(Apr '16 - May '17)*
  - Scored **95%/100** and secured **1st position** in the outgoing batch of **30+ students**
  - Awarded **Special Award for Best Application**, St. Martin's Diocesan School
- **All India Rank 13th (1000+ students)** in MSc in Quality Management Science Entrance 2021, Indian Statistical Institute
- **All India Rank 39th (2000+ students)** in Master of Operational Research Entrance 2021, University of Delhi

WORK EXPERIENCE

- **Deutsche Bank - Capstone Project (Location: London, UK)** *(Dec '22 - Sep '23)*

Leading German financial institution with operations spanning across 70 countries

  - **Elected by a panel of 5 professors to lead a group of** students to construct stochastic models for the company
  - Identifying **factors responsible for the movement of FX rates across the G10 markets with EUR & USD as base currencies**
  - Integrating Local Interpretable Model-Agnostic Explanations (LIME) algorithm. LIME architecture obtains a linear transformation of complex stochastic models to evaluate trading strategy and provide easier explanations to stakeholders
  - Incorporating Permutation Feature Importance (PFI) to **extract the crucial subset of financial datasets**. Helps to ascertain the important features which characterise the data related to varying asset classes
- **Delhivery - Operations Research Intern (Location: Gurugram, India)** *(Jun - Sep '22)*

India's leading Logistics and Supply Chain company providing services across 2300 cities and towns

  - Leveraged Python for data cleaning to **construct optimization models** for predicting cost-efficient routes pan-India
  - Analysed the vehicle routing dataset of **2500+ routes** for the company with respect to operational costs and delivery time
  - Achieved **reduction in operational costs by 8%** and **delivery time by 5%** for the delivery services of the company
- **Peacock Solar - Marketing Analysis and Digital Marketing Intern (Location: Gurugram, India)** *(Jun - Aug '20)*

Startup in the solar ecosystem rendering services by leveraging data science to help India achieve 50GW solar capacity

  - Created a Business Model Canvas for consulting reporting client segments, key partners and activities
  - Formulated a Lead Generation Plan through data analysis with reports of potential leads in Indore, Jaipur and Kota
- **St. Stephen's College - Nagpaul Fellowship (Location: New Delhi, India)** *(Aug '19 - Mar '20; Jul '20 - Apr '21)*

Auspicious fellowship in the domain of applied and pure mathematics offered by the Department of Mathematics, St. Stephen's College

  - **Selected by a panel of 8 professors to lead a group of** students from a competitive **pool of 50+ students**
  - **Modeled the rate of change of deposits based on percentage of quotas** of four major banks in the Greek Market
  - **Developed various epidemic models and graphs based on COVID-19 Modeling** research paper
  - Presented the findings to a panel of professors from the Department of Mathematics, St. Stephen's College

LSE DATA SCIENCE PROJECTS

- **S&P 500 Classification Using PySpark (Domain: Distributed Computing for Big Data)**
  - Dataset: Data from past 20 years obtained from Yahoo Finance about the 503 stocks that constitute the S&P 500 Index
  - Aim: To understand the efficiency of machine learning algorithms in terms of time (time complexity) over cloud computing
  - Outcome: **Time complexity of the algorithms to produce result decreases with improvement in computational resources**
- **Stock Price Prediction Using Generative Adversarial Networks (Domain: Finance and Investment)**
  - Dataset: Historical prices of technology stocks with financial tweets for incorporating public sentiment for analysis
  - Aim: To analyse the effect of financial news over accurately determining the prices of the stocks
  - Outcome: **Financial news helps in improving the accuracy of the price predictions**. News helps in capturing the public sentiments which influence financial markets. GAN models perform considerably well across machine learning models

PERSONAL DATA SCIENCE PROJECT

- **FIFA Project (Domain: Sports Analytics)**
  - Dataset: FIFA 22 dataset available over Kaggle characterised by attributes about all the players in the gaming title
  - Aim**:** To engineer a model with data visualisation for clustering the players based on attacking abilities to help clubs identify potential rising talents in the scouting and transfer market across summer and winter windows
  - Outcome: **Identified Erling Halaand as the best player** for transfer on the basis of age, growth potential and overall rating

TECHNOLOGY AND INTERESTS

- Programming Languages - Python (Jupyter Notebook), R, C++, JavaScript
- Libraries - NumPy, Pandas, Matplotlib, Seaborn, Plotly, Scikit-learn, TensorFlow, Keras, BeautifulSoup4, dplyr, ggplot2
- Cloud Computing and Data Engineering - Google Cloud Platform (GCP), Apache PySpark, Hadoop
- Database Systems & Version Control Repository Management – SQL, GitHub
- Web Development, Application Software - HTML, CSS, Microsoft Office, Microsoft PowerPoint, Microsoft Excel
- Football, Music, Skipping, LeetCode Coding, Kaggle Competitions