SOOMIN MYUNG

Graduate Visa (UK, until 31 Dec 2025)
Address: 32 Quay Street, Salford, M3 5GD, UK
Nationality: South Korea
LinkedIn | **GitHub**

Email: msm1640@gmail.com | **Phone:** +44 7442646421

EDUCATION

The University of Manchester   **Manchester, UK**
MSc Data Science, Distinction   Sep 2022 - Oct 2023

- **Core Modules:** Programming in Python for Business Analytics, Statistics and Machine Learning, Understanding Databases, Applying Data Science

The University of Nottingham   **Nottingham, UK**
MSc Computer Science (Artificial Intelligence), Distinction   Sep 2020 – Oct 2021

- **Core Modules:** Machine Learning, Systems and Networks, Databases, Data Modelling and Analysis, Games in Unity

Seoul National University   **Seoul, South Korea**
BA Economics, 3.83/4.3 GPA (94.30%)   Mar 2006 – Aug 2014*
* Mandatory military service in South Korea (2007 – 2009), study for Public Administration Exam (2010 – 2012)

- **Core Modules:** Microeconomics, Macroeconomics, Econometrics, Statistics, Mathematics for Economics, Agricultural Firm Management, Market Economy and Food Policy
- **Scholarship:** Superior Academic Performance (2006, 2007), the top 0.5% of the year in the degree course

SKILLS

- Machine Learning
- Deep Learning
- Artificial Neural Network (ANN)
- Convolutional Neural Network (CNN)
- XGBoost
- Data Analysis
- Exploratory Data Analysis (EDA)
- Data Visualisation
- Data Pre-processing
- Data Pipelines
- Python
- TensorFlow
- Scikit-learn
- NumPy
- Pandas
- R
- MySQL
- NoSQL
- MongoDB
- PyMongo
- Microeconomics
- Macroeconomics
- Econometrics
- Statistics
- Natural Language Processing (NLP)
- Transformer
- Principal Component Analysis (PCA)
- Large Language Models (LLMs)
- LangChain
- Tableau

PROJECTS

CNN for CIFAR-10 Image Classification (Python) | GitHub   **2023**

- A project to design a CNN model to see how hyperparameter tuning can improve classification performance.
- Designed a 5-layer CNN model for image classification with hyperparameter tuning (number of layers, number of filters, batch size, epoch, kernel size) and batch normalisation.
- Referenced journal articles to get ideas for initial model settings and hyperparameter tuning.
- Enhanced model accuracy from 69.8% to 81.6%.

Amazone Shopping Mall (MongoDB, PyMongo) | GitHub   **2023**

- Designed a NoSQL DB system for a shopping mall to efficiently manage all the stages of shopping.
- Designed schemas to take advantage of the flexibility of NoSQL and decided whether to do referencing or embedding for each schema, depending on whether the schema was write- or read-intensive.
- Built pipelines in PyMongo to support features such as product browsing, recommendation, shopping basket, order, dispatch of products and drivers, inventory management, and past orders.

Rossman Sales Prediction (Python) | GitHub   2023

- Given a dataset, the project was to train and test a sales forecasting model for a drugstore franchise.
- Performed Exploratory Data Analysis (EDA) to discover the characteristics (data type and distribution of each feature) and any issues (missing values, duplicates, outliers, multicollinearity, and unit inconsistencies) of the dataset.
- Pre-processed the data, such as missing value imputation, one-hot encoding for categorical variables, and introducing new variables to better represent and use the information in the dataset.
- Performed SHAP value analysis to see each feature's contribution to sales prediction for feature selection and further business insights.

Monopoly (MySQL) | GitHub   2023

- Designed a MySQL DB system to make a Monopoly game using triggers and procedures.
- The system gets a dice number, it automatically proceeds the game and controls the game/player status (turn, budget, tile bonus, properties, bankruptcy, etc.) until there is only one player left in the game as the winner.
- Data normalisation for efficiency and to avoid anomalies.

WORK EXPERIENCE

Data Consultant, DATA70202 Project   England
Rail Growth Limited, Consultancy   Feb 2023 - Jun 2023

- Conducted research to estimate the potential carbon emissions reductions by a modal shift from 100% road truck transportation of deep-sea containers to rail transportation between Sohar and Muscat in Oman.
- Collected data from the European Commission, Eurostat, OECD, World Bank, EEA, and Defra to estimate the potential reduction in externalities and further justify the modal shift.
- Developed own methods to properly estimate various kinds of externalities using only reliable data within limited data availability.

Project Manager, Full-time   South Korea
Korea Energy Technology Evaluation and Planning, Government Agency   Nov 2017 - Dec 2018

- Supported government energy technology policy project management.
- Held meetings for project selection, audit, and evaluation.
- Worked on government reports for all relevant tasks.

Assistant Budget Manager, Full-time   South Korea
Korea Rural Community Corporation, Government Agency   Apr 2014 - Dec 2014

- Supported government budget management using Excel and SAP.

OTHER ACTIVITIES

Certificates | LinkedIn   2020 - Present

- LangChain with Python Bootcamp, Udemy
- Data Modelling in Power BI, Microsoft, Coursera
- NLP with Classification and Vector Spaces, DeepLearning.AI, Coursera
- Fundamentals of Visualization with Tableau, UC Davis, Coursera
- C++ For C Programmers, UC Santa Cruz, Coursera
- C for Everyone, UC Santa Cruz, Coursera

Korea Information Technology Academy   2019

- Python, Java, HTML, CSS, MySQL

US Army, Yongsan Garrison, South Korea   Aug 2007 – Jul 2009
Korean Augmentation to the United States Army, Driver

- Served as a driver for the Air and Missile Defence Division (AMDD).