<FIRST> <LAST>

<Phone> / <Email> / [<Github>](https://github.com/HungN0102) / <Location>

EDUCATION

**Cass Business School**

Master of Science (M.Sc.) in Business Analytics **Sep 2020 – Sep 2021**

* Achieved: Distinction.
* Relevant Modules: Data Management Systems, Machine Learning, Natural Language Processing.

**Queen Mary University of London**

Bachelor of Science (B.Sc.) in Mathematics with Actuarial Science. **Sep 2017 – Jun 2020**

* Achieved: Upper Second Class.
* Relevant Modules: Python, R, Statistics, Statistical Modelling, Calculus, Financial Reporting.

SKILLS

* Programming Languages: HTML, CSS, JavaScript, SQL, NoSQL, Python, R, Spark, Linux.
* Tools: Tableau, Power Bi, GCP, AWS, Docker, Git.

PROJECT EXPERIENCE

**Web-scraper** **Jun 2022 – Jul 2022**

* Created a Python-based web-scraping application to extract data and visualized the proportion of technical skills required for Software Engineer across 2253 job descriptions on LinkedIn and Glassdoor.
* Applied Machine Learning theories to design Topic Modelling with TF-IDF implementation to analyze 10 most common topics across different sectors. Discovered that finding solutions and developments are the main requirements within the Technology sector.

WORK EXPERIENCE

**IWSR**

Data Engineer **Mar 2022 - Present**

* Designed and maintained complex T-SQL queries, views and stored procedures in multiple databases hosted on AWS RDS. Resulting in tuned SQL queries that extract and transform faster while producing less errors.
* Built an ETL pipeline that ingested millions of transactional data using S3 buckets, Spark, Python and AWS Glue.
* Engineered multiple automated data-preprocessing tasks that eliminated all manual works using Python scripting.

**4C Associates**

Data Analyst **Nov 2021 – Mar 2022**

* Utilized Natural Language Processing’s techniques to classify the item’s category, then used Power Bi to perform data modeling then visualize monthly spent data provided by clients namely Reckkit, Bank of England, Just Eat.
* Developed statistical algorithms and implemented Python to the daily processes, reducing from 5 to 2 hours per day of analyzing and validating data provided by NHS.

**Rolls-Royce**

Applied Research Project **Apr 2021 – Sep 2021**

* Developed Supervised Machine Learning Algorithms such as tree-based models to produce survival analysis on 1120 biological factors that influence Motor Neuron Disease patients’ life expectancy.
* Utilized Python and Machine Learning models such as Logistic-Lasso Regression and Linear Support Vector Machine to identify toxic language in 200,000 text data. Resulting in 96% accurate predictions.
* Produced Tableau dashboards and prepared presentations to introduce projects to senior data scientists and industry professionals. Recommended possible future works that can meet business objectives.