

# MINJAE KANG

(+44) 7947 485975 | [jerrykang135@gmail.com](mailto:jerrykang135@gmail.com) | [LinkedIn](#) | [Portfolio](#)

## EDUCATION

**University College London (UCL)**, London, UK Sep 2017 – June 2021  
MEng Computer Science; Result: *First Class Honours*  
**Modules include:** Intro to Deep Learning (1<sup>st</sup>), Statistical NLP(1<sup>st</sup>), Information Retrieval and Data Mining(1<sup>st</sup>), Machine Learning for Domain Specialist(1<sup>st</sup>), Maths and Statistics(1<sup>st</sup>), Internet of Things(2:1)  
**Abbey College Cambridge**, Cambridge, UK Sep 2014 – June 2016  
A-Levels: Mathematics (A\*), Further Mathematics (A), Physics (A)

## PROJECTS

**Information Retrieval and Ranking with Deep Learning**, London, UK Jan 2021 – Mar 2021

- Developed information retrieval models that return a ranked list of documents relevant to a query.
- Applied text preprocessing and tf-idf weighting to queries and documents and implemented BM25 and a query likelihood language model.
- With pre-trained dense word embeddings, GloVe, implemented ranking models with Logistic Regression, LambdaMart, and Ranknet.
- Used Python pandas, NumPy, nltk, and Tensorflow/Keras on Jupyter Notebook.

**Comparative Analysis of News Topic Classification Models**, London, UK Jan 2021 – Mar 2021

- Carried out a comparative analysis of deep learning models: RNN, GRU, LSTM, BiGRU, and fine-tuned BERT for news topic classification task.
- Explored the state-of-the-art pre-trained word embedding, BERT, on news topic classification.
- Trained and tested the models with BBC News dataset using Python Tensorflow/Keras.
- Compared the performance of the models with Accuracy, F1-score, Precision, and Recall metrics.

## RELEVANT EXPERIENCE

**Data Engineer Intern, Balaan; an E-commerce Start-up**, Seoul, South Korea June 2020 – Oct 2020

- Built a data pipeline that collects business data using AWS Cloud services S3 and Athena as a Business Intelligence platform to serve across the firm.
- Automated the ETL process for Google Analytics data.
- Created BI dashboards with AWS Quicksight to show weekly KPIs and achievements for employees and to use for the company decision-making.
- Automated the data crawling, cleaning, and mapping with the relevant crawling tools and MongoDB.
- Streamlined data management and mapping process by developing utilisation tools for the sales team.
- Adapted to the company's OKR culture and learned about goal setting and priority.

**Cross-department Smart Classroom IoT Project**, London, UK Jan 2020 – April 2020

- Delivered a smart classroom prototype that autonomously controls a classroom environment such as temperature, brightness, and humidity.
- Designed and built an Internet of Thing (IoT) system with a CC3200 launchpad and environment sensors.
- Developed a data pipeline with IBM cloud and created a UI dashboard with IBM Node-Red.
- Performed EDA and data cleaning on London Air Quality Dataset with Python pandas and NumPy.
- Led a team of four; allocated tasks according to one's strengths and capabilities.

**Software Development for Arthur Murray Dance Studio**, London, UK Oct 2018 – April 2019

- Client-based team project; collaborated with Arthur Murray dance studio and tailored technological and practical services to the client's needs.
- Developed a web application using Node.js(Express framework) with client required functionalities.
- Hosted MongoDB database and the application on Microsoft Azure.
- Liaised with clients in different time zones and coped with changing requirements under time pressure.

**Private A-levels Physics Tutor**, Seoul, Korea July 2016 – August 2016

- Ran one-to-one teaching sessions 12 hours per week.
- Prepared all the teaching materials and handled admin tasks, for example, renting a classroom.
- Adapted different approaches to improve students' results according to their strengths and weaknesses.

## SKILLS

- **Languages:** Python, Scala, SQL, NoSQL, Node.js, C/C++, JavaScript, HTML, CSS
- **Analytics:** Pandas, NumPy, Apache Spark, ETL, Scikit-learn, TensorFlow/Keras
- **Software Tools:** Git, Linux(CRON), Jupyter, Anaconda
- **Cloud:** AWS (S3, Athena, Quicksight), IBM Node-RED