



## ANLUN JIANG

Data Engineer

### ABOUT ME

My interest in data and technology was cultivated from both my personal interests as well as my engineering degree. During the summers of 2016 and 2017 I undertook my university's research internship programme, where I worked closely with PhD candidates on their research, data acquisition, processing and analysis. I found contributing insights and recommendations from my processed data very rewarding, which further inspired me to pursue a career in data.

I was also able to develop my interpersonal and leadership skills. I was responsible for supervising and teaching new research members the techniques and theory of the research project.

I enjoy playing piano and rock climbing. I helped run my university climbing club, my interest in technology lead me to help design the website for our club, which involved teaching myself HTML and CSS.

### CONTACT US

Kubrick Group

T: 020 3866 4620

E: [consultants@kubrickgroup.com](mailto:consultants@kubrickgroup.com)

W: [kubrickgroup.com](http://kubrickgroup.com)

### OVERVIEW

Graduating from Imperial College London with a MEng in Materials Science and Engineering has developed my strong analytical and problem-solving skills, as well as given me a strong appreciation of data. Being comfortable working with large data volumes in a variety of formats, I chose my final year project to further establish this key skill. My master's project involved the simulation of graphene semiconductor interactions for solar applications. By utilising MATLAB and OriginLab, I created algorithms and identified trends from the experimental data. This resulted in a working theoretical model that predicted graphene semiconductor potential barrier. My findings revealed new insights into improving the efficiencies of graphene solar cells as well furthered the understanding of the electronic nature of graphene.

Joining Kubrick has proved to be a fantastic opportunity in allowing me to bridge the gap between research related data analysis and the wider data industry.

### WORK

#### DATA ENGINEER - KUBRICK GROUP, LONDON

September 2018 - Present

At Kubrick Group I gained knowledge and real world experience in professional skills, a wide yet in depth array of modern technical and analytic skills. We studied the agile software development lifecycles and gained detailed knowledge of platforms and infrastructure. I specialised in Data Engineering which involved learning and developing advanced skills in Python, Spark, Hadoop, NoSQL and SQL as well as advanced Excel. As a junior data consultant I applied my skills on real client projects including:

- Sourcing, profiling, enrichment, dimensional modelling and analysis of UK Crime data spanning 10 years. I identified trends and correlations using statistical methods visualised in Tableau. I produced both an interactive Dashboard as well as an Infographic telling story behind the analytics identified.
- Forecasting sales and identifying seasonal trends for a leading UK retailer. Working in agile teams, we enriched the client dataset with macroeconomic and weather data as well as scraping of the client website to gather product description data for passing into our NLP algorithms. The analytics were modelled using machine learning regression techniques and visualised in Plotly. The final product was delivered via a Git repository and documented using Markdown. Our issues were all tracked in a JIRA dashboard.

#### IMPERIAL COLLEGE LONDON

July 2017 - Sept 2017

**Research Project:** Metal Organic Framework synthesis on ceramic substrate membrane for gas separation and purification

Sponsor: UK Engineering Physical Sciences Research Council (EPSRC)

- The project aim was to investigate the separation efficiency of gaseous mixtures using different ceramic/MOF membranes and varying experimental parameters. I demonstrated my analytical and innovative skills by developing a membrane structure and combination of parameters for optimum ideal selectivity and presented my findings to my supervisor. My work has been submitted in a scientific journal.
- I gained working knowledge of key approaches to numerical and technical analysis using appropriate software such as OriginPro and Matlab.

July 2016 - Sept 2016

**Research Project:** Regenerative Metal Organic Framework for the removal of contaminants in waste water

- I undertook a research project; "Alkaline treated UiO - 66 - NO2 adsorbent with regenerative abilities for the removal of oxyanions from waste water" in the Chemical Engineering Department under Imperial College's UROP programme.
- I was responsible for more than 50% of data acquisition and analysis. Duties included skilled working in complex machinery such as ICP, SEM and XRD and precise pH and temperature optimisation. This required huge multi-tasking, time management, precise working and critical analysis skills to organise an abundance of raw data into a comprehensible format for analysis.

### EDUCATION

#### IMPERIAL COLLEGE LONDON - MATERIALS SCIENCE AND ENGINEERING (MENG) (2:1)

October 2014 - October 2018

**Modules include:** Mathematics and Modelling, Probability and Statistics, Project Management, Managerial Economics, Optoelectronics, Nanomaterials and Quantum Mechanics

**Final Year Project:** Modelling and Optimising Graphene Solar Cells

### KEY SKILLS

- Python Libraries: Pandas, Numpy, SK Learn, NLTK, Plotly
- Machine Learning Skills: Statistical Learning, Regression, Classification, Clustering
- Cloudera Hadoop version 5.4: HDFS, Yarn, Sqoop, Flume, Avro, Parquet, Impala, Hive and Hue
- SQL & NoSQL: MongoDB, Neo4J, Modelling, SSIS, Alteryx, Advanced SQL
- Apache Spark: RDD, DataFrames, SparkSQL, PySpark
- Source Control and IDE: Git, PyCharm, Visual Studio
- Other Skills: Tableau, Linux CentOS 6, SSH, Putty, Agile, LaTeX, Markdown