



NG Ming Hin

☎ (+852) 9349 1019
✉ ngmhmatthew@gmail.com
🌐 LinkedIn / GitHub
🏠 m2ng.github.io

> Hi, I am *Matthew*.

> I have a very strong interest in data, particularly in *data science, machine learning and big data*. My willingness to learn allows me to familiarize myself with new things quickly and be open to others' opinions.

SKILLS

Data-related

TensorFlow

Jupyter stack

Big data

Tableau

SQL

Programming languages

Python

C++

Java

R

Excel VBA

Web-related

CSS

Javascript

React.js

EXPERIENCE

July 2019 — Present (Anticipated completion: June 20, 2021)

The Chinese University of Hong Kong

MPhil in Statistics (CGPA: 3.910 / 4.000)

- Equipped myself with a set of transferable problem-solving skills during research
- Developed a scalable machine learning method to enable association detection over a million of genetic loci for a million of individuals
- Tags: C++ & OpenMP Parallel computing Big data analysis

June 2018 — July 2019

Hong Kong Telecom

Data Science Analyst

- Developed efficient web crawling techniques to gather important information of *hundreds of thousands of products* from a popular e-commerce platform
- Designed and automated the whole pipeline in the recommender system for an e-commerce platform
- Utilized various machine learning methods to provide similar products in the recommender system
- Derived useful insights from browsing behaviour of millions of users using Python, SQL and big data technologies such as Spark and Hadoop
- Implemented an interface to stream web analytics data to a MongoDB database
- Agile project management using Jira
- Familiar with using AWS EC2 instances

Sep 2014 — July 2018

The Chinese University of Hong Kong

BSc in Statistics (CGPA: 3.363 / 4.000)

- Programming coursework: Python Excel & VBA Java R
- Department of Statistics Scholarship 2015/16, 2017/18; Academic Merit in 07/2017; Advantage Trust Statistics Scholarship 2015/2016

PROJECTS

- Machine learning for business**
 - *E-commerce recommender system*
 - *Transform product name and product description into comparable numeric vectors*
- Machine Learning**
 - *Passenger Screening Algorithm Challenge*
 - Predicting winning probabilities in HK horse racing
- Data acquisition**
 - *Stream web analytics data to MongoDB*
 - Create profiles for websites