**DLab Internship Essay: Summer 2019**

Nicola Lawford

This summer, I had the privilege of interning at the Data Science and Engineering Lab (DLab) at King Mongkut's University of Technology Thonburi (KMUTT) in Bangkok, Thailand, where I worked on a research project and a startup, and grew through many new experiences.

My research project was in bioinformatics, the study of biological data and its applications. My aim was to enter the Malaria DREAM Challenge, an international collaborative bioinformatics research project asking participants to develop machine learning models to classify malaria parasite drug resistance based on gene expression data from blood samples. Gene expression measures a relative level of how much a gene is being transcribed to produce proteins, which determines properties of cells in an organism such as how they are structured and how they adapt to changes in the environment--in this case, drug treatment. This was an interesting problem for machine learning because we have relatively few blood data samples with resistance labels to study--a small training set--with many genes which we know little about and may or may not be relevant--high-dimensionality features. In my work, I applied many machine learning algorithms with several feature transformation methods developed at DLab, which transform the expression of many genes into the activity of a small number of pathways performing specific functions in the cell.

I worked on a secondary project for FinQuanti, a KMUTT startup competing in the U!reka startup contest. Their aim is to generate and sell synthetic backtesting data for existing stocks, and work with customers to generate data for new assets in hypothetical market

scenarios. This was another interesting problem: how do we measure the properties of stock price time series, and how do we emulate them while also giving our data the unpredictability of real-world stocks? For FinQuanti, I replicated a paper that answered some of these questions, producing a customizable time series generator in MATLAB.

As one of my first research experiences, my internship empowered me with a self-directed project, and sparked my interest in machine learning and its applications across fields. Given many opportunities to hear presentations and talk to graduate students working on data science research in bioinformatics, chatbots, and smart-home technology, I learned a lot about how engineering researchers apply the latest math and science theory to solve problems and create new things, in any field.

Outside academia, I took every opportunity I could to learn and grow on this international experience. DLab professors generously took us to see many places Thailand: we went to see many temples and markets in Bangkok and Ayutthaya, Thailand's old capital, and hiked a series of waterfalls in Erawan National Park. The picture to the right is of me and other interns at the Temple of Dawn. I tried many new dishes and fruits (my favourites were Tom Yam soup and mangosteen), and I ran in two

races: 8km and 5km. Although it was at times difficult to live and work in a new place far from home, I learned the importance of reaching out to those around me, and made friends by talking to international students in my residents and playing badminton after work at the university gym.

At the beginning of this summer, I had no experience in bioinformatics, machine learning, or finance. I had never been to Asia. In Thailand, I contributed to a bioinformatics and machine learning research project and worked on a financial tech startup prototype. I ran the farthest I have ever run and ate the sweetest and spiciest foods I have ever tasted. I am forever grateful to the professors, graduate students, friends and family who made it possible for me.