**ACADEMIC PROJECT SUBMISSION DETAILS:**

<table>
<thead>
<tr>
<th><strong>PROJECT #:</strong></th>
<th>19</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Supervisor/s:</strong></td>
<td>Eibe Frank and Rory Mitchell</td>
</tr>
<tr>
<td><strong>Project Title:</strong></td>
<td>Comparison and evaluation of GPU shuffling algorithms</td>
</tr>
<tr>
<td><strong>Field:</strong></td>
<td>Computer Science</td>
</tr>
<tr>
<td><strong>Division/School:</strong></td>
<td>HECS - Division of Health, Science, Computing &amp; Engineering, School of Computer Sciences, Maths &amp; Statistics</td>
</tr>
</tbody>
</table>

**EXPECTED OUTCOMES:**

1. Research poster summarising the results
2. Open-source implementations of ported algorithms
3. Test and benchmark suite for evaluated algorithms

**STUDENT TASKS:**

1. Research and identify existing algorithms
2. Investigate the feasibility of applying these algorithms to the GPU
3. Port the applicable algorithms to the GPU
4. Conduct an evaluation of the performance of the different algorithms
5. Produce a final research poster demonstrating your findings

**REQUIRED SKILLS:**

1. GPU Programming Experience

**PROJECT ABSTRACT:**

There are a range of algorithms aimed at efficiently generating random permutations on multi-core CPUs and the goal of this project will be to investigate the feasibility of using these algorithms on the GPU, which has a different set of architectural challenges and performance requirements. The project will also compare these approaches to existing algorithms targeted at the GPU.
**ACADEMIC PROJECT SUBMISSION DETAILS:**

<table>
<thead>
<tr>
<th>Supervisor/s:</th>
<th>Vimal Kumar</th>
</tr>
</thead>
<tbody>
<tr>
<td>Project Title:</td>
<td>Malware detection using Deep Learning</td>
</tr>
<tr>
<td>Field:</td>
<td>Computer Science, Cyber Security</td>
</tr>
<tr>
<td>Division/School:</td>
<td>HECS - Division of Health, Science,</td>
</tr>
<tr>
<td></td>
<td>Computing &amp; Engineering</td>
</tr>
<tr>
<td></td>
<td>School of Computer Sciences, Maths &amp;</td>
</tr>
<tr>
<td></td>
<td>Statistics</td>
</tr>
</tbody>
</table>

**EXPECTED OUTCOMES:**

1. Progger output dataset of malicious and benign applications
2. Design of a deep learning model for malware detection

**STUDENT TASKS:**

1. Review of existing DL techniques
2. Progger Deployment and data collection
3. Design of DL Technique for malware detection
4. Conversion of Progger Data to DL input
5. Implementation, testing and tuning of the model

**REQUIRED SKILLS:**

1. Knowledge of Operating Systems
2. Knowledge of Cyber Security
4. Good Programming Skills

**PROJECT ABSTRACT:**

In this project, we will explore the application of deep learning for malware detection. The CROW lab at the University of Waikato has developed a provenance tool known as Progger. Progger runs in the background on a machine and records provenance information of the processes running on the machine. The information recorded includes information about the process itself, the system calls made by the process, the timestamps, etc. In this project, we intend to apply deep learning techniques on the data collected through Progger to detect malicious activity. While this research will complement other research being done in the CROW lab and can be extended for detection on streaming data, the scope of this project will be limited to developing the technique for offline Progger data.
ACADEMIC PROJECT SUBMISSION DETAILS:

<table>
<thead>
<tr>
<th>Supervisor/s:</th>
<th>Michael Mayo</th>
</tr>
</thead>
<tbody>
<tr>
<td>Project Title:</td>
<td>Prediction of Surgically Induced Astigmatism using AI</td>
</tr>
<tr>
<td>Field:</td>
<td>Computer Science</td>
</tr>
<tr>
<td>Division/School:</td>
<td>HECS - Division of Health, Science, Computing &amp; Engineering</td>
</tr>
<tr>
<td></td>
<td>School of Computer Sciences, Maths &amp; Statistics</td>
</tr>
</tbody>
</table>

EXPECTED OUTCOMES:

1. Review of background literature and available tools concerning using of machine/deep learning for regression from medical images
2. Detailed preliminary analysis of our dataset written up in the form of a short paper
3. Poster summarising findings

STUDENT TASKS:

1. Analyse existing patient data (OCT images and tabular numeric data)
2. Investigate regression techniques for predicting future visual acuity scores
3. Investigate medical deep learning techniques that could be fine tuned to deal with OCT images
4. Perform a number of experiments concerning the above techniques with our data
5. Summarise findings in a report and at meetings

REQUIRED SKILLS:

1. Python programming
2. Data analysis/machine learning
3. Image analysis
4. Good verbal/written reporting

PROJECT ABSTRACT:

Astigmatism is a commonly occurring problem causing blurred vision. There are many causes of this condition, for example it can be hereditary. Sometimes the blurring increases over time, while other times it decreases. Amongst several possible treatments, one option is laser surgery. However, laser surgery itself can actually make the astigmatism worse if something goes wrong we call this ‘surgery induced astigmatism’. This project will aim to predict the likelihood of surgery induced astigmatism.
PROJECT ABSTRACT:

We have a dataset of approx. 120 patients with their associated visual acuity measurements before surgery and up to 3 months after surgery. We also have high resolution images of the cornea taken using optical coherence tomography (OCT). The aim of this project will be to come up with an approach for predicting post-operative visual acuity (i.e. whether vision improves or gets worse) from pre-operative OCT scans and pre-operative visual acuity data. This project will be most suitable for a computer science student with prior experience in python and machine/deep learning with images, and someone interested in further study since this is likely to be a topic that goes beyond just the 10 weeks of summer. It is a collaborative project with Waikato DHB.
### ACADEMIC PROJECT SUBMISSION DETAILS:

<table>
<thead>
<tr>
<th>Supervisor/s:</th>
<th>Panos Patros</th>
</tr>
</thead>
<tbody>
<tr>
<td>Project Title:</td>
<td>Reliability of Green Edge Computing with Clusters of Nodes of Raspberry Pis</td>
</tr>
<tr>
<td>Field:</td>
<td>Software/Computer Engineering</td>
</tr>
<tr>
<td>Division/School:</td>
<td>HECS - Division of Health, Science, Computing &amp; Engineering</td>
</tr>
<tr>
<td></td>
<td>School of Computer Sciences, Maths &amp; Statistics</td>
</tr>
</tbody>
</table>

### EXPECTED OUTCOMES:

1. Set up cluster of raspberry pi nodes with kubernetes
2. Deploy and test sample distributed application
3. Evaluate performance with/without fault injection
4. Short paper and poster discussing results
5. The student continues on to do honors/masters/PhD with the ORCA lab

### STUDENT TASKS:

1. Put together nodes of raspberry Pis and connect them as cluster
2. Install micro kubernetes
3. Deploy and test example distributed application
4. Load-test without injecting failures
5. Load-test while injecting failures
6. Analyse and present data and findings in the form of a report and poster

### REQUIRED SKILLS:

1. Software Engineering
2. Computer Engineering
3. Electrical and Electronic Engineering
4. Embedded Systems
### PROJECT ABSTRACT:

People expect their Internet of Things (IoT) devices to just work. However, this requires reliable Internet and continuous power supply, a hurdle for remote and off-grid Aotearoa New Zealand (NZ). In conjunction with a lack of edge standardization and high maintenance costs, the automation of agritech and scientific operations is subsequently hindered.

Edge computing, which is a distributed computing paradigm inspired by the success of cloud computing, can unlock the full potential of IoT devices in NZ. It can expand their range and allow them to operate autonomously should the connection goes down.

We are working on a collaborative project with international partners to effectively enable green edge computing. We are developing a standardised edge architecture that leverages edge nodes comprising stacks of Raspberry Pis; these nodes are then connected with each other and form a cluster on the target location.

For this project, we are looking for an enthusiastic student to set up a proof-of-concept cluster in the lab, manage it with Micro Kubernetes, deploy a sample distributed application on it and take performance measurements. Critically, we want to see how the system behaves if nodes, Pis or deployed containers fail at a configurable rates, which will be the main research deliverable of this work.

The student will work at the ORCA lab (Oceania Researchers in Cloud and Adaptive-systems) and will be expected to continue on as an honors/master/PhD student after working on this project.
ACADEMIC PROJECT SUBMISSION DETAILS:

<table>
<thead>
<tr>
<th>Supervisor/s:</th>
<th>Jessica Turner</th>
</tr>
</thead>
<tbody>
<tr>
<td>Project Title:</td>
<td>Simulating Traffic Data in the Bay of Plenty</td>
</tr>
<tr>
<td>Field:</td>
<td>Traffic Modelling and Simulation</td>
</tr>
<tr>
<td>Division/School:</td>
<td>HECS - Division of Health, Science, Computing &amp; Engineering</td>
</tr>
</tbody>
</table>

EXPECTED OUTCOMES:

1. Identification of gaps in existing functionality of open source tools.
2. Gather traffic related data for road and/or intersections to be modelled.
3. A model which can be used to explore traffic flow for specific roads and intersections based on gathered data.

STUDENT TASKS:

1. Explore and understand how to use open source tools openstreetmap, SUMO and MATSim to create a traffic model of a selected subset of the Bay of Plenty traffic network.
2. Determine data to gather based on existing MATSim examples and carry out traffic counts.
3. Enter data in a workable format for MATSim.
4. Use data to build MATSim Model.
5. Build a SUMO traffic network using open street map.
6. Combine SUMO network with MATSim to model collected data.
7. Explore different solutions to reduce congestion in the model.

REQUIRED SKILLS:

1. Experience in data collection.
2. Understanding of some form of modelling software.
3. Experience using Python.
4. Would be useful to have used openstreetmap, SUMO or MATSim before but not required.
5. Ability to work independently and adapt.

PROJECT ABSTRACT:

Traffic congestion in the Bay of Plenty area is a well-known issue for commuters. Data from Priority One has identified that traffic flows have increased by 5.7% in Tauranga alone for the year between September 2017-2018. This is partially matched with a collective increase in both personal and commercial vehicle registrations in the area of approximately 5%. Traffic congestion adds to the financial and environmental cost of commuting with increased fuel consumption leading to increased air pollution.
This summer research project builds on an initial investigation carried out in the Bay of Plenty area. Various traffic modelling open source tools were identified as useful to simulate potential solutions to reduce traffic congestion with minimal infrastructure investment. However, these solutions need to be data-driven to best motivate potential changes to the existing network and a combination of these tools is necessary in order to closely simulate “real world” scenarios. In this project you will explore using SUMO, openstreetmap and MATSim to model and simulate a subset of the traffic network to explore this idea further and trial potential solutions to reducing traffic congestion.