



COLLEGE of ENGINEERING AND PHYSICAL SCIENCES

SCHOOL OF COMPUTER SCIENCE

MSc Seminar

Wednesday September 4, 2019 at 1:00PM in Reynolds, Room
2224

Applying Convolutional Neural Networks to the prediction of
Open and Closed compartments of DNA.

Matthew Kirchhof

Advisor: Dr. Stefan C Kremer

Advisory Committee: Dr. Deborah A Stacey

ABSTRACT:

Previous research has shown that the spatial organization of a genome has significant effects on gene regulation. Understanding how the organization of the genome within a cell affects gene expression has the potential to lead to various breakthroughs surrounding disease, cell differentiation and more. In 2009, a method to capture the conformation of genomes called Hi-C was published, revealing an effective but time consuming means of analysing the interaction of a genome with itself. From this, we can categorize 100kb bins of the genome as open (A) compartments or closed (B) compartments, where A compartments are more expressive and interact with other parts of the genome more often than B compartments.

This seminar presents preliminary research into using convolutional neural networks (CNNs) to predict the A/B compartmentalism of bins across a given genome, using only its DNA encoding. Artificial neural networks have been used to predict these A/B compartments when given a variety of predetermined features such as GC content. However, research has yet to be published on a CNNs ability to predict these compartments without the use of predefined features. If successful, we may be able to derive new undiscovered patterns and features within the genome that correlate to its spatial organization, and therefore its gene regulation. More importantly, we would be able to predict A/B compartments in a faster and less expensive manner across many different cells. Preliminary results using CNNs to predict A/B compartmentalism demonstrate potential for state of the art classification.