Thematic Research Training 2018-2019
Artificial Intelligence/Machine Learning (AI/ML) group

AI/ML Third meeting
Monday 20th May 2019, 11:00-17:00
Edinburgh, King's Buildings, Hudson Beare Building - Classroom 7
(*https://preview.tinyurl.com/yycwlr75*)

This event will consist of a day of 4 talks by users of AI/ML approaches in a biological or biomedical context. The aim is to demonstrate the variety of approaches available and to showcase their use in research.

10:30: Coffee/tea and pastries

11:00-11:40

Speaker: James Holehouse

Title: Using moment-based maximum likelihood inference to infer parameters from experimental data

Abstract: A current challenge in molecular biology is to find the correct mechanisms through which genes regulate their own expression. Modelling using the chemical master equation (CME) and the stochastic simulation algorithm (SSA) allow us to see the theoretical behaviour of gene specific regulatory mechanisms, such as auto-regulatory feedback or the genetic toggle switch. The recent advent of single molecule data *in vivo* allows one to compare experimental probability distributions, for molecule numbers at specific times, to their theoretical counterparts. This talk concerns the inference of parameters (such as reaction rates) from experimental data onto a given reaction network using the moment-based maximum likelihood method. I will elucidate the applications and challenges behind this approach using examples of real inference from simulation data.

11:45–12:45

Speaker: Professor Neil Carragher

Title: Multiparametric high content imaging: An enabling technology for emerging A.I. applications in drug discovery

Abstract: In this presentation I will introduce how the latest advances in automated microscopic imaging and image analysis has contributed to a new field of cell pharmacology and functional genomics called *High Content Analysis*. I will describe how multiparametric high content imaging technology can provide important functional context and foundations for emerging Artificial Intelligence (A.I.) solutions to identifying novel drug mechanisms-of-action, drug combinations, biomarkers and guiding novel chemical design. I present our

recent examples of the application of machine learning and deep learning approaches for classification of drug mechanism of action across genetically distinct cancer cell lines representing different clinical subtypes and I will provide a specific case study illustrating the application of these methods to a compound library screen performed across oesophageal cancer cell lines. Finally we describe ongoing research projects to build large high content phenotypic screening data sets fully integrated with genomic, proteomic, small molecule target annotation and chemical structure data to support the A.I. Community.

12:45–14:00: Lunch

14:00–14:40

Speaker: **Dr Steve Shave**

Title: Encoding complex data types; AI friendly representations of small molecule

Abstract: Connectivity graphs representing small molecules do not trivially lend themselves to AI and machine learning techniques. Traditionally, this problem was addressed through the use of molecular descriptors, hoping not to lose too much information in the encoding process. In this talk, the use of descriptors for small molecule physicochemical property prediction is demonstrated along with more cutting-edge neural network architectures, enabling the lossless encoding of molecular graphs and exploration of chemical space.

14:50– 15:30

Speaker: **Dr Julian Pietsch**

Title: Discovering cellular dynamics in the age of artificial intelligence

Abstract: Natural biological environments are constantly varying and in recent years it has become clear that cells both monitor such environmental change and generate dynamic signals in order to prepare for future conditions. Investigating the regulatory architectures responsible for such behaviour is challenging, however, as it requires precise control of the cellular microenvironment and the ability to track individual cells through time. We use microfluidics and time-lapse microscopy to study growth regulation dynamics in yeast, and rely heavily on artificial intelligence to automate the annotation and processing of >50,000 images per experiment. In this presentation, I will briefly describe how we use machine learning to identify and outline cells in images and the improvements in accuracy we get through the use of convolutional neural networks. The resulting single-cell time-series data sets pose an additional challenge: there is no general statistical framework for addressing even simple queries, like whether control data is significantly different from experimental. We have developed a machine learning approach, founded on information theory that makes such questions tractable. I will introduce this method and highlight instances where traditional machine learning methods, like support vector machines, still outperform neural networks. Where possible I will also provide practical starting points to begin using such techniques in your own work.

EASTBIO Training pages: http://www.eastscotbiodtp.ac.uk/thematic-research-skills-training#ai/ml
For any queries, please email EASTBIO at enquiries@eastscotbiodtp.ac.uk.