

## **CRUNCH Seminars at Brown, Division of Applied Mathematics**

**Friday - April 10, 2020**

### **REAL-VALUED (MEDICAL) TIME SERIES GENERATION WITH RECURRENT CONDITIONAL GANS**

Liu Yang  
Brown University

Generative Adversarial Networks (GANs) have shown remarkable success as a framework for training models to produce realistic-looking data. In this work, we propose a Recurrent GAN (RGAN) and Recurrent Conditional GAN (RCGAN) to produce realistic real-valued multi-dimensional time series, with an emphasis on their application to medical data. RGANs make use of recurrent neural networks (RNNs) in the generator and the discriminator. In the case of RCGANs, both of these RNNs are conditioned on auxiliary information. We demonstrate our models in a set of toy datasets, where we show visually and quantitatively (using sample likelihood and maximum mean discrepancy) that they can successfully generate realistic time-series. We also describe novel evaluation methods for GANs, where we generate a synthetic labelled training dataset, and evaluate on a real test set the performance of a model trained on the synthetic data, and vice-versa. We illustrate with these metrics that RCGANs can generate time-series data useful for supervised training, with only minor degradation in performance on real test data. This is demonstrated on digit classification from 'serialised' MNIST and by training an early warning system on a medical dataset of 17,000 patients from an intensive care unit. We further discuss and analyse the privacy concerns that may arise when using RCGANs to generate realistic synthetic medical time series data, and demonstrate results from differentially private training of the RCGAN.