Bayes Filters and Recurrent Neural Networks

Abstract:
Recurrent Neural Networks (RNNs) have seen a massive surge in popularity in recent years, particularly with the advent of modern architectures such as LSTMs. These sophisticated modern models have resulted in significant performance gains across a number of challenging tasks. Despite their success, we still struggle to provide a rigorous theoretical analysis of these models, or to truly understand the mechanism behind their success. Prior to the success of RNNs, time series data modelling was dominated by Bayes Filters in their many forms. In contrast to RNNs Bayes Filters are grounded in axiomatic probability theory, resulting in a class of models which can be easily analyzed and whose action is well understood. In this work we propose a new class of models called Predictive State Recurrent Neural Networks, which combine the axiomatic probability theory of Bayes Filters with the rich functional forms and practical success of RNNs. We show that PSRNNs can be learned effectively by combining Backpropropogation Through Time (BPTT) with a method-of-moments initialization called Two-Stage Regression. Furthermore we show PSRNNs reveal interesting connections between Kernel Bayes Rule and conventional RNN architectures such as LSTMs and GRUs. Finally we show PSRNNs outperform conventional RNN architectures, including LSTMs, on a range of datasets including both text and robotics data.

Speaker:
Carlton Downey

Thesis Committee:
Geoff Gordon (Chair)
Ruslan Salakhutdinov
Byron Boots (Georgia Tech)
Arthur Gretton (University College London)

Nov. 20, 2017
10:00am
4405 GHC

Link to draft document: