



## Seminar/Talk

# Bayesian inference and deep learning

**Laurence Aitchison**

Janelia Research Campus

Host: Peter Jonas

Here, I present three projects using principled, Bayesian treatments of uncertainty to address pressing problems in modern deep learning.1.) Bayesian inference in modern convolutional neural networks. Bayesian inference over the parameters of a neural networks is critical to improve generalisation performance, and to reason about what the network does not know due to limited training data. However, Bayesian inference in typical neural networks is impossible (at least without severe approximations) due to the sheer number of parameters in these networks. Here, we show exact inference is possible in state-of-the-art convolutional networks, if we take the limit of infinitely many convolutional filters (at which point the outputs follow a Gaussian process). The network obtains 0.84 % classification error on MNIST, a new record for a Gaussian Process method.2.) Neural network optimization as Bayesian inference. Neural network optimization methods fall into two broad classes: adaptive methods such as RMSprop and non-adaptive methods such as stochastic gradient descent (SGD). This presents a problem for practitioners: which method should use on a particular problem? Or even should you use an adaptive method on some parameters, and a non-adaptive method on others? Here, we resolve this issue, by deriving a Bayesian gradient descent rule that adaptively transitions between adaptive and non-adaptive behaviour. This method provides insight into when we might expect adaptive and non-adaptive methods to be most useful, and is superior to standard neural network optimization methods in practice.3.) Bayesian inference in deep graphical models. Graphical models are a powerful language to encode our knowledge of the dependency (or even causal) structure in data. However, graphical modelling has fallen out of favour recently, due to the success of often unstructured deep-learning. Here, we show that it is possible to combine deep learning and graphical models to form "deep graphical models". To perform inference, we combine strategies from deep learning (variational autoencoder recognition models) with strategies from graphical modelling (message passing). The resulting inference schemes give considerably improved performance over a vanilla deep-learning inference strategy.

**Monday, January 28, 2019 10:00am - 11:00am**

Mondi Seminar Room 2, Central Building

---



This invitation is valid as a ticket for the ISTA Shuttle from and to Heiligenstadt Station.  
Please find a schedule of the ISTA Shuttle on our webpage:  
<https://ista.ac.at/en/campus/how-to-get-here/> The ISTA Shuttle bus is marked ISTA Shuttle (#142) and has the Institute Logo printed on the side.