Dropout as Bayesian approximation: Representing Model Uncertainty in Deep Learning

Yarin Gal, Zoubin Ghahramani
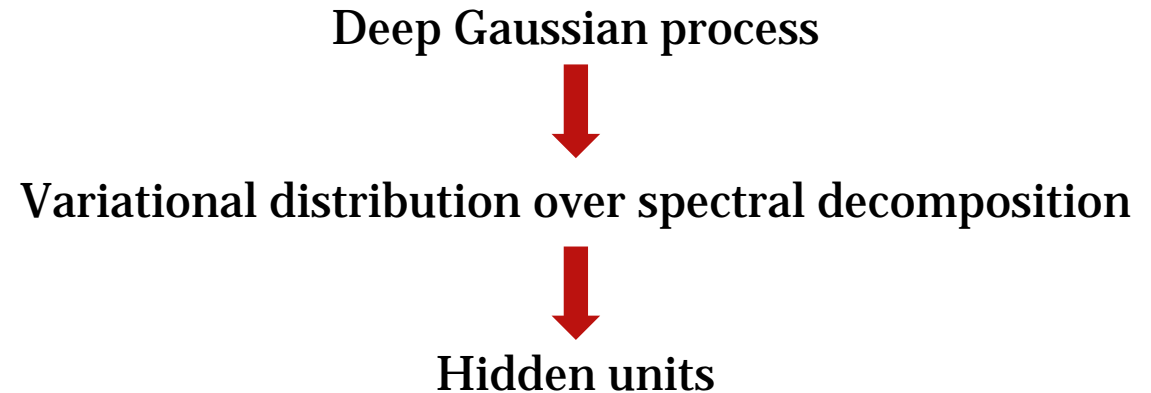
Challenge

❑ **"dropout applied before the every weight layer"**
  – How about using dropout associated with only few layers?
  – How about extend this to GNN layers?

❑ **"dropout with probability 0.1/ 0.5"**
  – Any special reason?

Deep Gaussian process

⬇

Variational distribution over spectral decomposition

⬇

Hidden units

Simple and Scalable Predictive Uncertainty Estimation using Deep Ensembles

Balaji Lakshminarayanan, Alexander Pritzel, Charles Blundell

❑**Deep ensemble provide a useful uncertainty estimate in practice**

– Is there any theoretical guarantee for its consistency in model estimation?
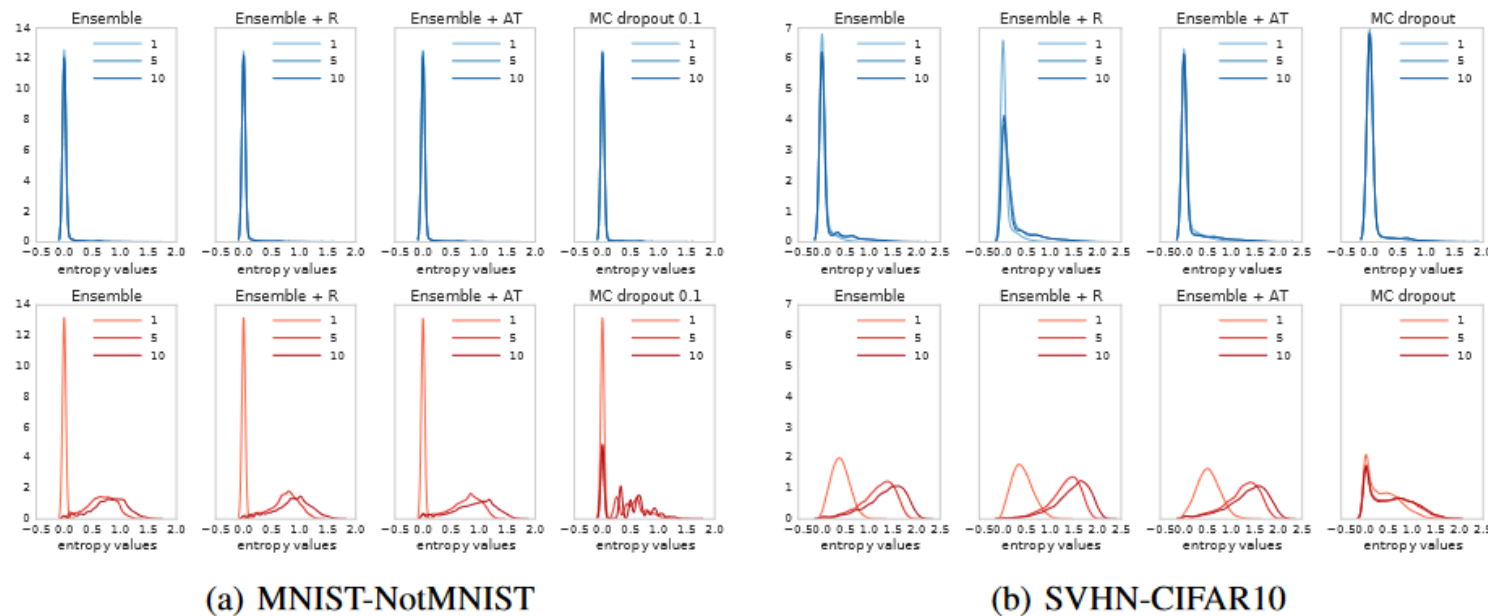

Figure 3: : Histogram of the predictive entropy on test examples from known classes (top row) and unknown classes (bottom row), as we vary ensemble size M.