Bayesian Neural Networks

By Hong

Agenda

1. Why would you bother with BNNs?
2. Why do BNNs hate you(r hardware)?
3. How can we approximate BNNs?
4. What are the disadvantages of Approximation?
5. Variational Inference Tutorial

Why bother with Bayes?

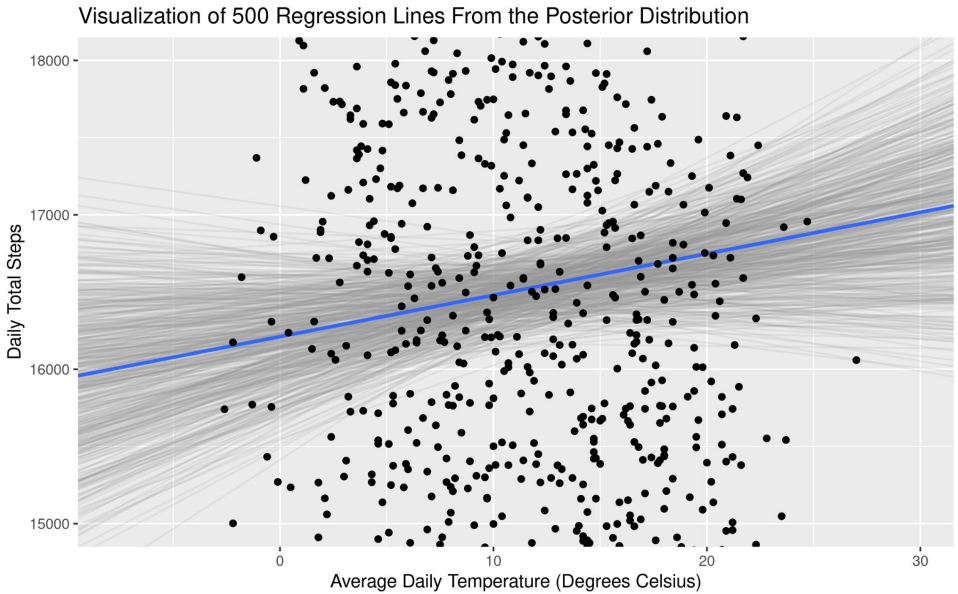
PRIOR PROBABILITY		
Initial estimate of how likely it is that he is cheating on you.	x	4%
A NEW EVENT OCCURS: MYSTERIOUS UNDERWEAR ARE FOUND		
Probability of underwear appearing conditional on his cheating on you.	y	50%
Probability of underwear appearing if he is *not* cheating on you.	z	5%
POSTERIOR PROBABILITY		
Revised estimate of how likely it is that he is cheating on you, given that you've found the underwear.	$\frac{xy}{xy + z(1-x)}$	29%

Why bother with BNNs?

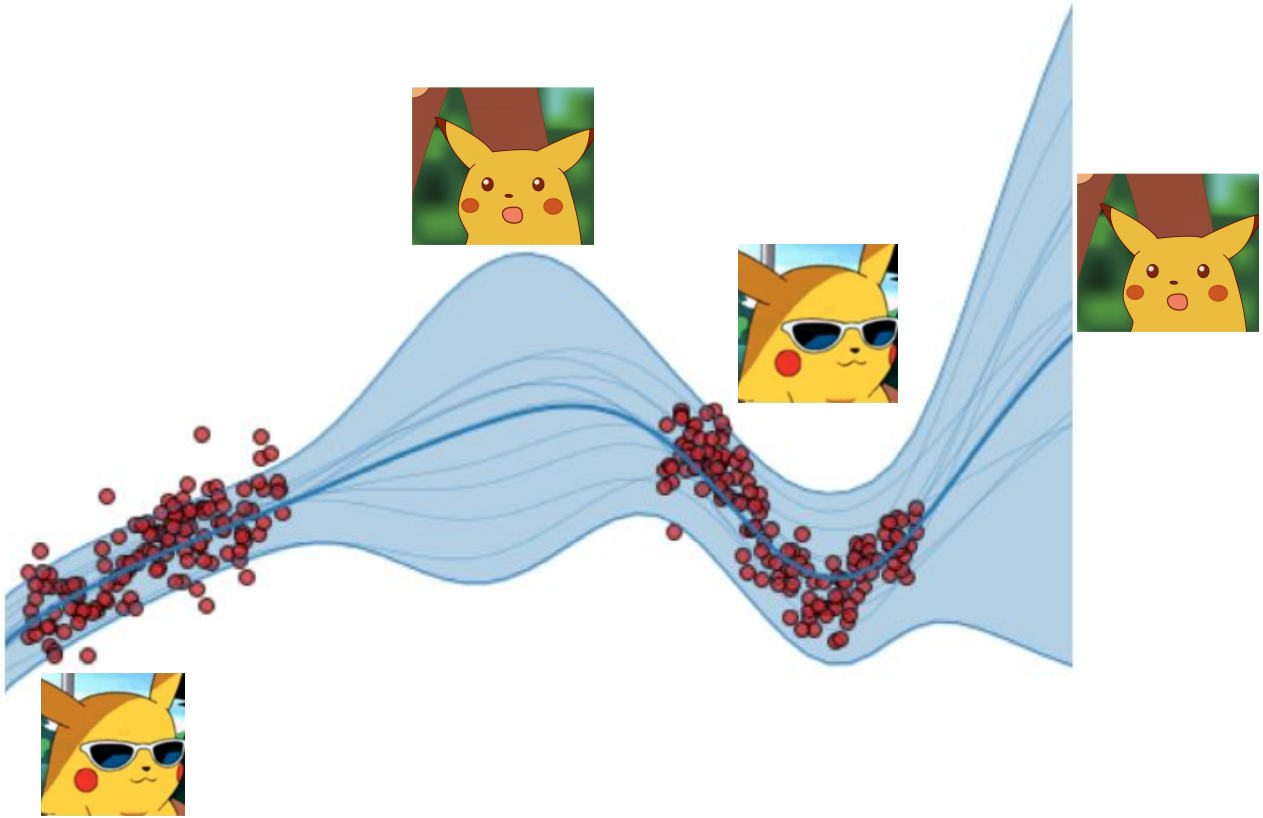
$$P(W|D) = \frac{P(D|W)P(W)}{P(D)}$$

What are the advantages of BNNs?

1. Quantifiable Uncertainty
2. That's all
3. Yep

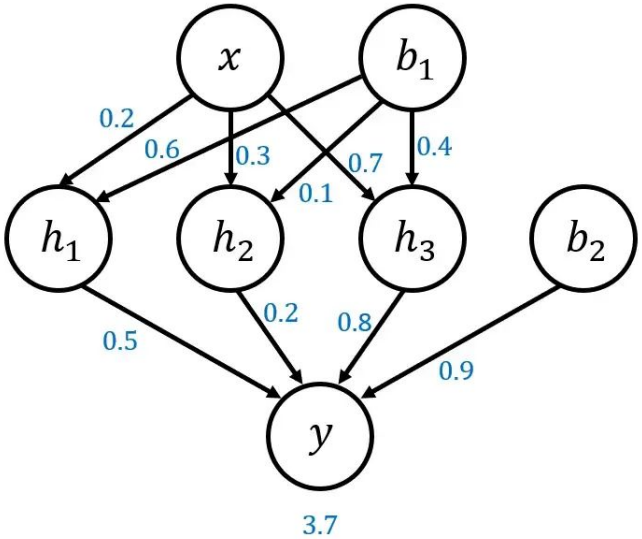


Why is Quantifiable Uncertainty Important?

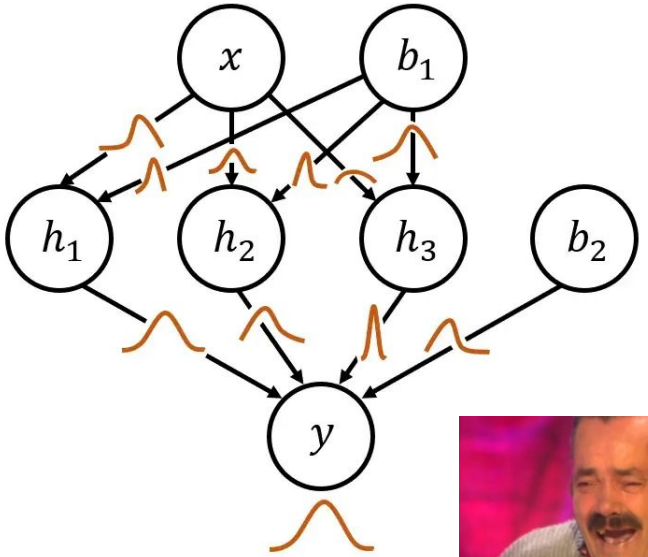


Why do BNNs hate you(r hardware)?

Standard Neural Network



Bayesian Neural Network

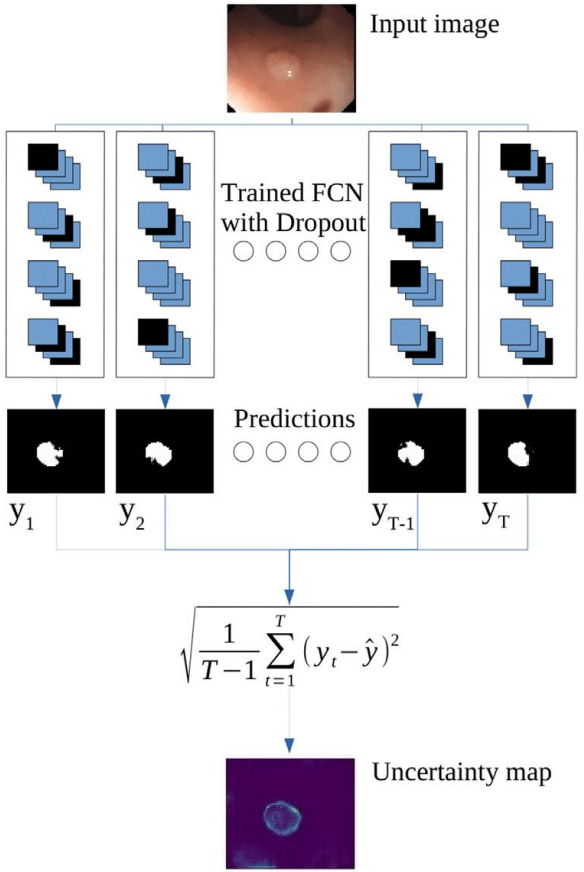


"Backprop over the posterior of the parameter space"

Why are we still here if BNNs are intractable?

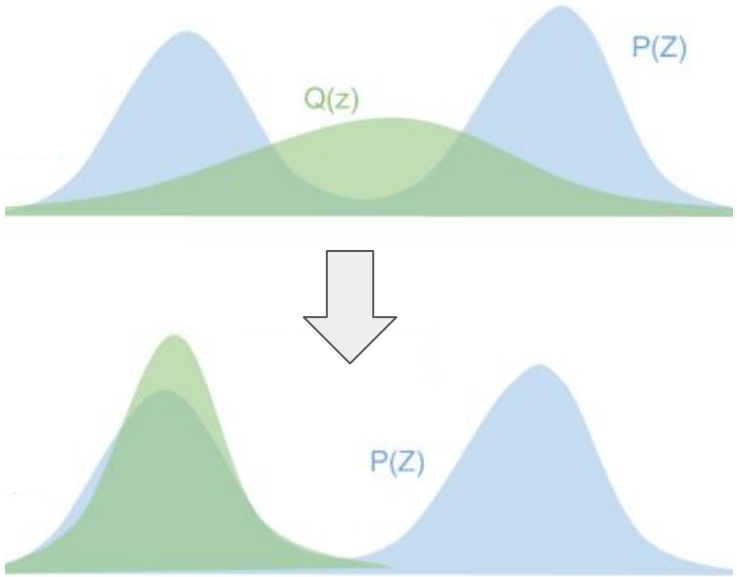
Monte Carlo Dropout

1. Generate more samples from the network using dropout
2. *Claims* to approximate a bayesian deep gaussian process https://arxiv.org/pdf/1506.02142.pdf
3. Very easy to implement
4. Costly inference



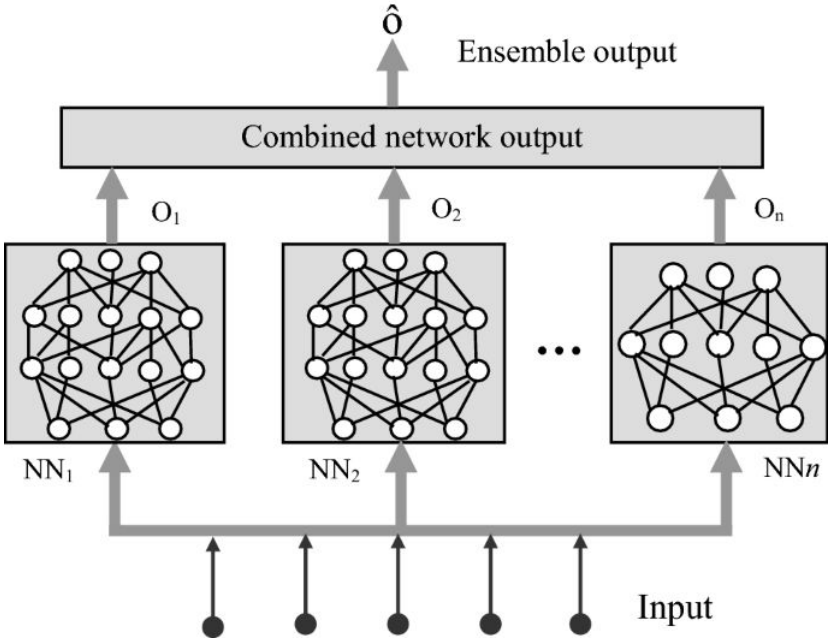
Variational Inference

1. *Assume* that P(W|D) has the shape of a parameterizable distribution (eg. Normal distribution)
2. Backprop the parameters of the distribution, rather than weights
3. In addition to likelihood loss (MSE), minimize the KL Divergence between the variational posterior and the prior
4. Parameterization may not match P(W|D)

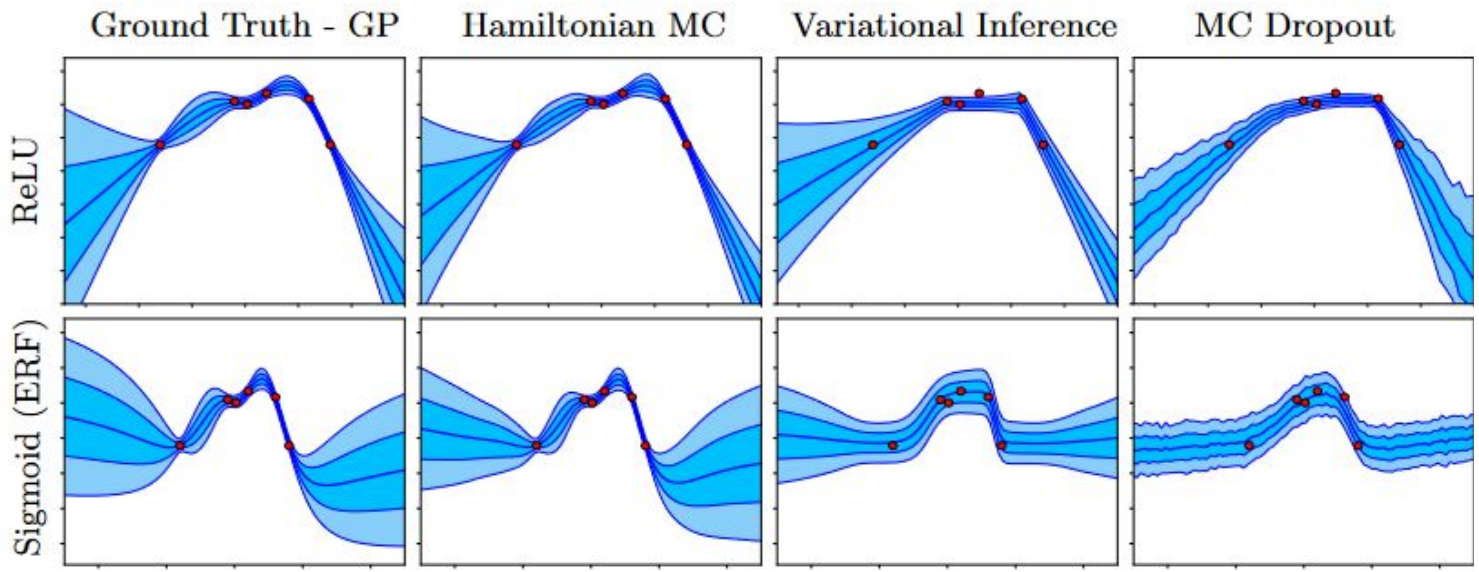


Deep Ensemble

1. Train multiple networks with different initialization
2. Assume that the networks do not converge to the same MAP weights
3. Costly training



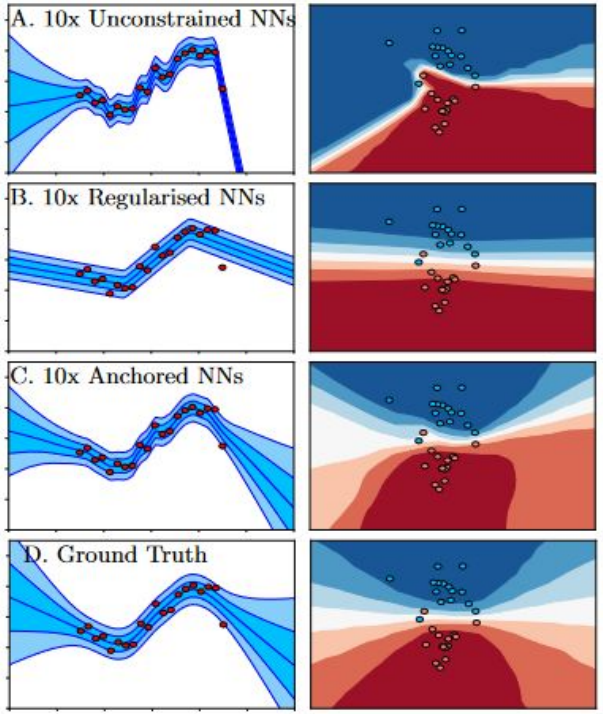
Disadvantages of Approximation



Hamiltonian MC https://arxiv.org/pdf/2010.06772.pdf

Disadvantages of Approximation

1. Approximation methods can introduce their own biases
2. Can be quite expensive to compute
3. Research in better BNN approximations is ongoing



Tutorial on Variational Inference

Go to https://colab.research.google.com/drive/1pK47PsCAoqF6GR-9KR63Ix7Oh-106aSU?usp=sharing