Decompose uncertainty into :

* Aleatoric uncertainty
* Epistemic Uncertainty :
  + OOD uncertainty
  + Model uncertainty

OOD uncertainty : depends on the expressiveness of the functions in the model class. The higher thecapacity of the model class, the more the models in this class will extrapolate differently in data-scarce regions resulting in higher OOD uncertainty.

Solution: GPs (not scalable with data)🡪 BNN (intractable inference)🡪 NLM (priors on the weight of the feature map).

GPs: fine with OOD but not NLM. (epistemic does not increase in data-poor region)

Ambition: why NLM are not suited for OOD (decisions boundaries of the posterior) + new framework for training NLM that provides good uncertainty (for the 3 types).

Differenciation of this paper: provide a way to specifically track every type of uncertainty.

Process of NLM: very clear: Hamiltonian MC: spot the regions with a priori high mass for the posterior.

Tool to measure confidence: Entropy: why?

Analysis of the OOD uncertainty for NLM: Objective Posterior must include decision boundaries that bound high-likelihood data regions. (we want boundaries that will surround the data)

Objective of NLM: construct a feature map that allows to predict the outcome y. Constructing such map does not allow to efficiently bount the data (for the posterior with the dot product with the weights): the resulting NLM classifier will underestimate OOD uncertainty. Toy experiment in order to show that the features learned to predict y from x are not relevant in order to predict OOD or not. What is working: a feature map that is trained to jointly classify b,y|x. We have shown here that there is a problem inherent to the training of NLM.

From this intuition, we can construct a new paradigm of training: BaCOUn: Bayesian Classifier with OOD Uncertainty.

General training framework for NLMs:

* Generate OOD samples on the boundary of data: use normalizing flows (go into a latent space using invertible functions where the distribution is reasonnable, sample from the boundaries and go back into the original space).
* Train a classifier to distinguish between K+1 (K+OOD) classes: Neural Network with softmax
* Use the feature map and insert it in the NLM: HMC, VI

Evaluate OOD uncertainty: Epistemic Uncertainty on OOD data points vs ID data points.

Will we be able to bound the data ?