SAND2022-4472C

Sandia National Laboratories

# Assessing the Limits of Predictive Uncertainty in Seismic Event Discrimination Using Bayesian Neural Networks

Jorge A. Garcia, Lisa Linville, Thomas A. Catanach

SSA 2022

Sandia National Laboratories is a multimission laboratory managed and operated by National Technology & Engineering Solutions of Sandia, LLC, a wholly owned subsidiary of Honeywell International Inc., for the U.S. Department of Energy?s National Nuclear Security Administration under contract DE-NA0003525

Sandia National Laboratories is a multimission laboratory managed and operated by National Technology & Engineering Solutions of Sandia, LLC, a wholly owned subsidiary of Honeywell International Inc., for the U.S. Department of Energy's National Nuclear Security Administration under contract DE-NA0003525.

# Traditional NNs are great at predicting

## But not so much at **not** predicting

- Overconfident predictions makes it difficult to trust models
- Classification: Tuning of label threshold may help, but there is no other criteria to distinguish False Positives from True Positives.
- Regression: How can we establish if a prediction is wrong?
- **The goal of this is to determine how we can prevent irresponsible predictions?**

# Bayesian NNs might provide an answer

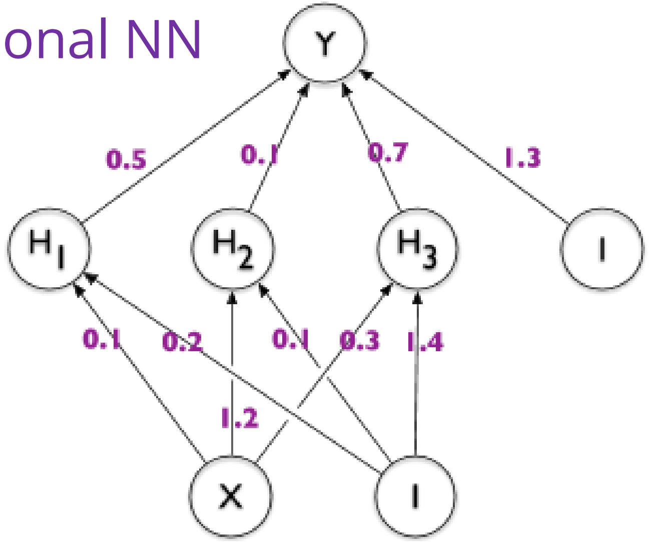
Obtain a distribution of weights

- Allows for sampling different models, resulting in varying predictions

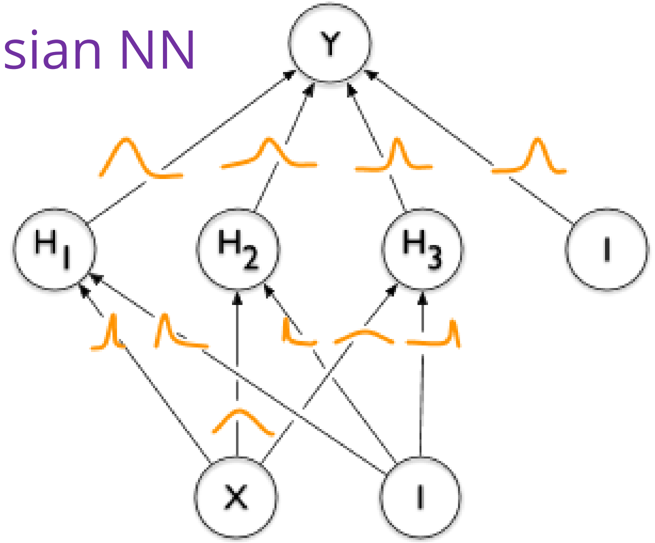
Therefore, obtain a distribution of predictions

- Standard deviation of prediction → confidence on prediction

Traditional NN




Bayesian NN



# Suppose we want to find a model:

$$f(x_0, w) = x_1$$

Want to know what the parameters w are:

Likelihood

Prior

Posterior

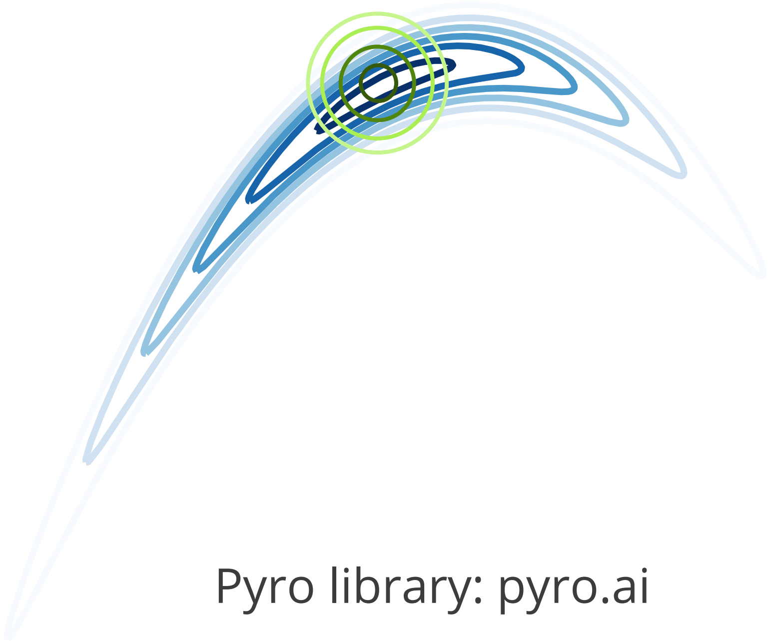
$$p(w|x) = \frac{p(x|w)p(w)}{p(x)} = \frac{p(x|w)p(w)}{\int p(x,w)dw}$$

Evidence

Need posterior, but Evidence is hard to calculate in high dimensions

# Variational Inference (VI)

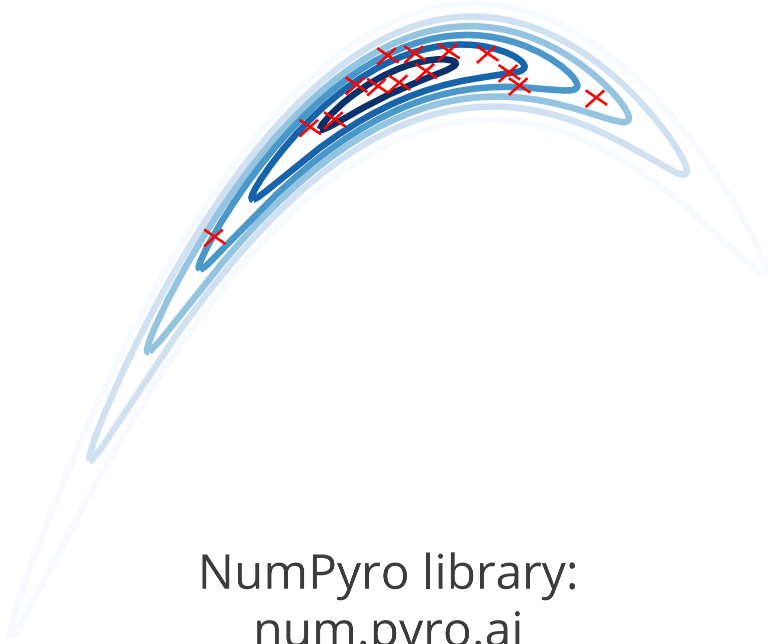
- Use a family of distributions $q(w)$ to find the one that best approximates the posterior $p(w|x)$



Pyro library: pyro.ai

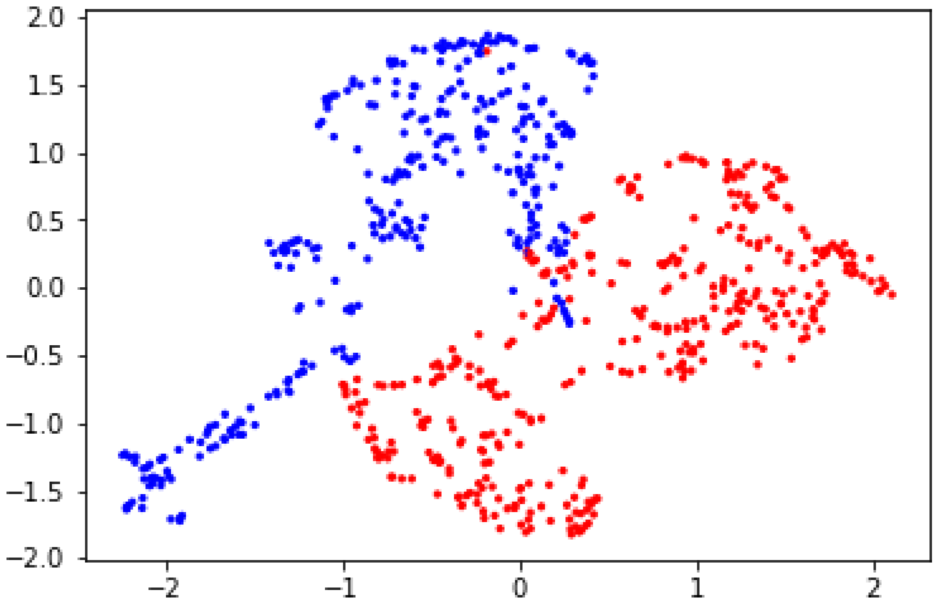
# Markov Chain Monte Carlo (MCMC)

- Find the posterior $p(w|x)$ by drawing parameter samples and testing them against the data

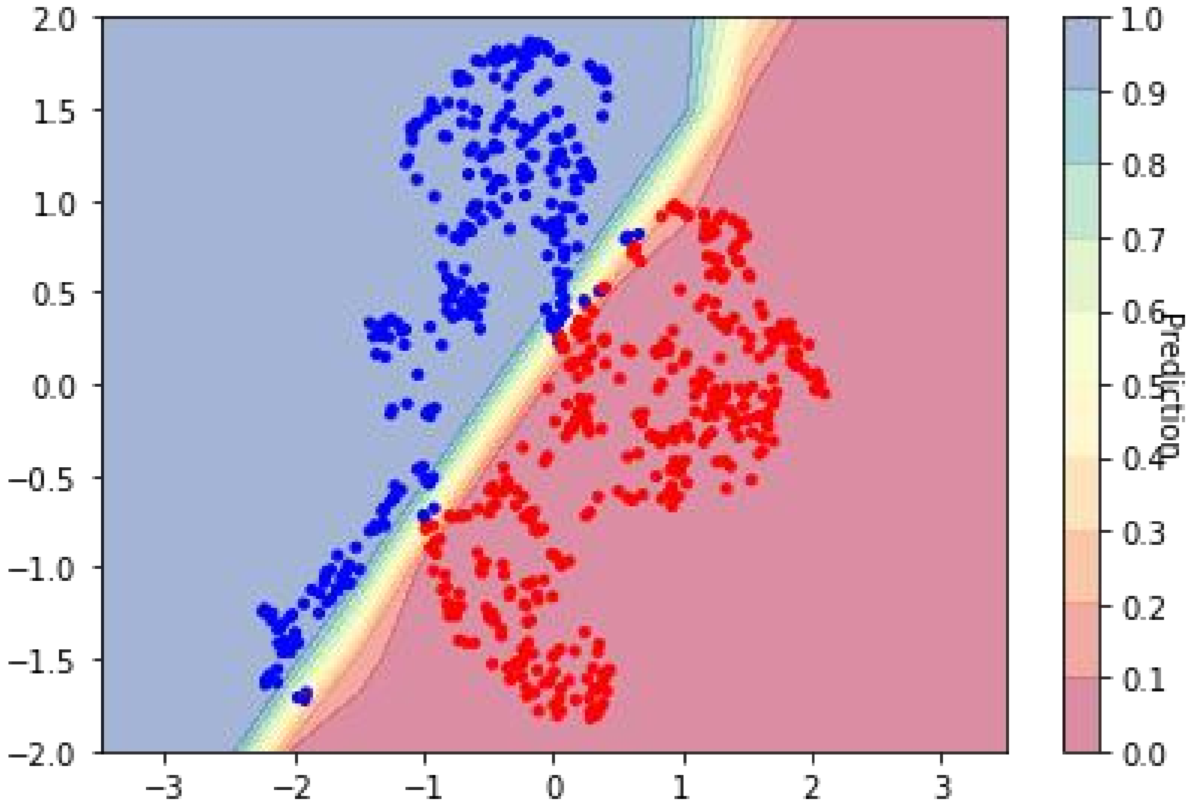


NumPyro library: num.pyro.ai

# Dataset I: Exploratory Dataset




2D mapping of seismic data
(quarry blasts vs earthquakes)




95% acc

Same architecture as used in previous SSL work we've done:

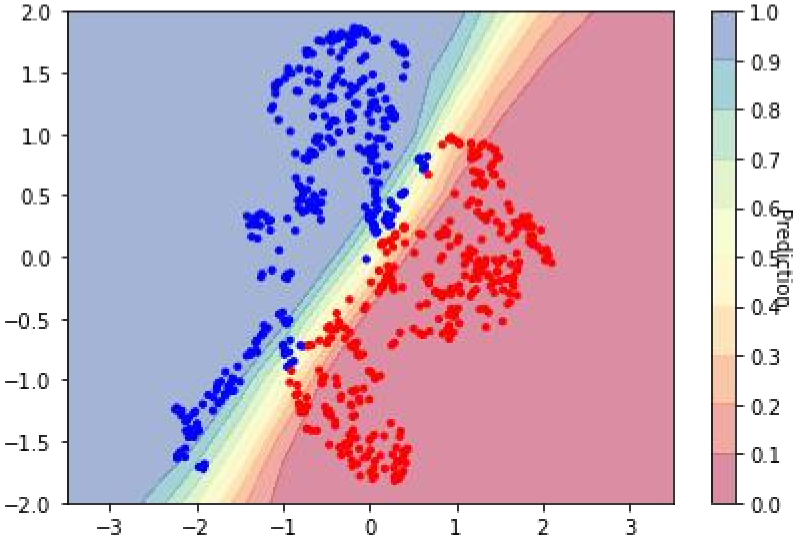
- Feedforward Network with 3 layers
- Tanh activations

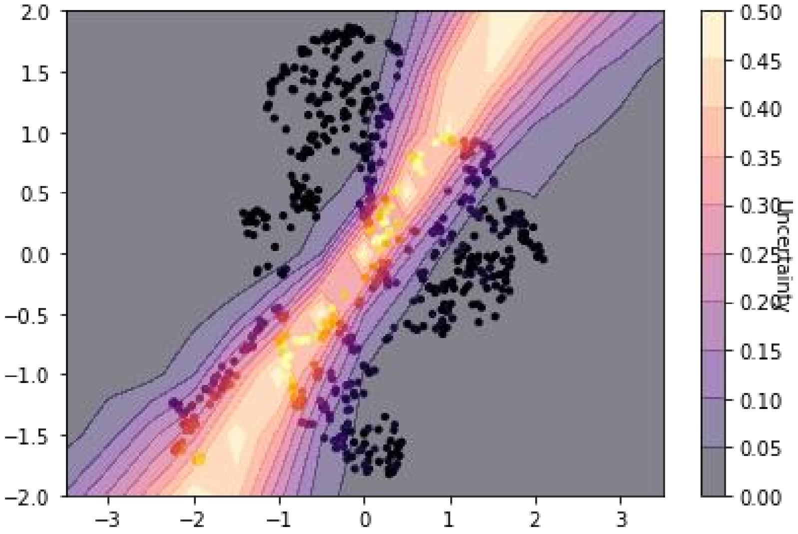
Fit a standard NN using gradient descent

- Naturally, easiest way to separate the data set is a line

# VI: Quick to train, similar results, more info

$\mathcal{N}(0,1)$



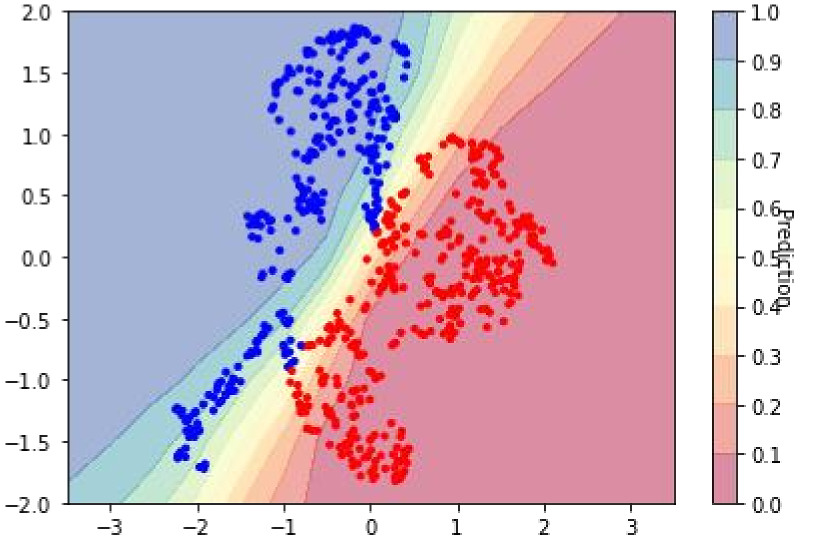



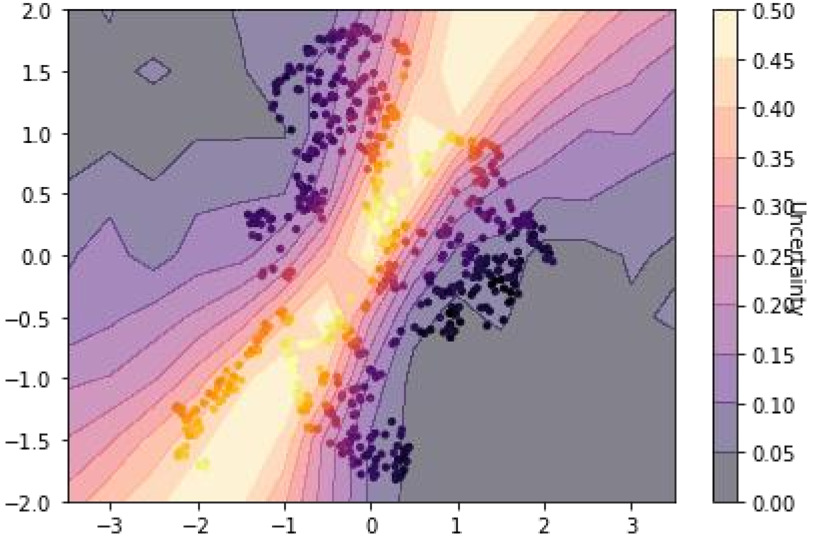

92% acc

We can obtain a distribution of predictions for an input, allowing to obtain a mean prediction (upper) and a standard deviation (lower)

**Prior choice matters**: Larger variance explores a larger model space, resulting in a wider/less "steep" decision boundary

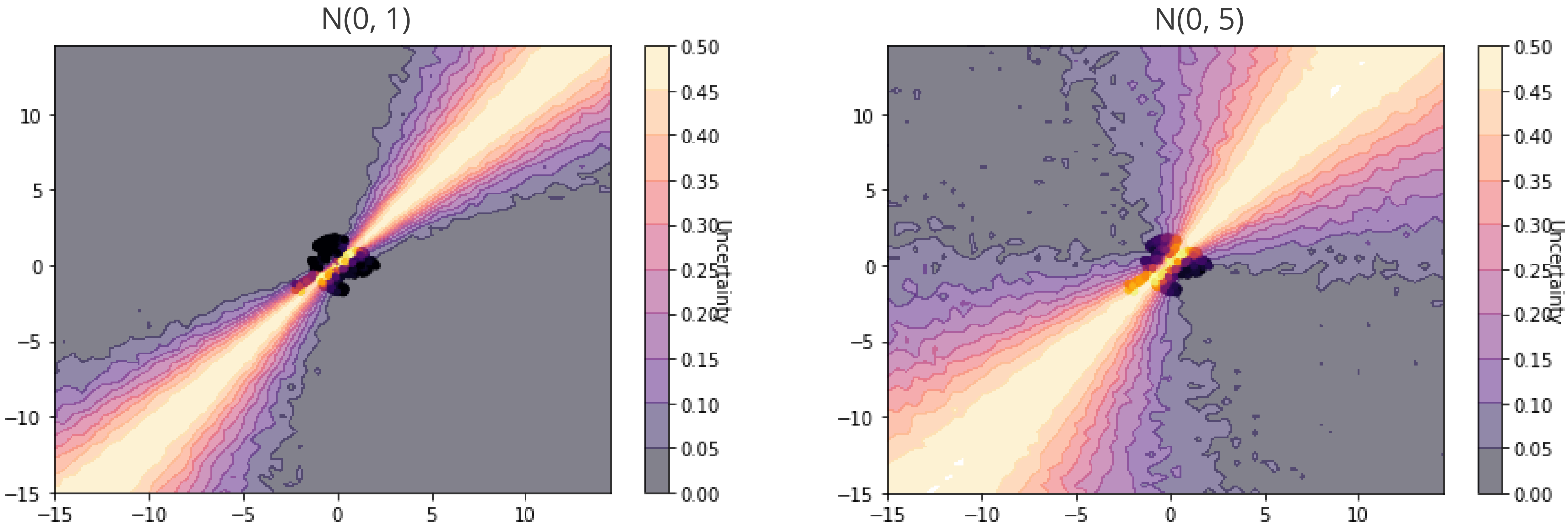
$\mathcal{N}(0,5)$




88% acc

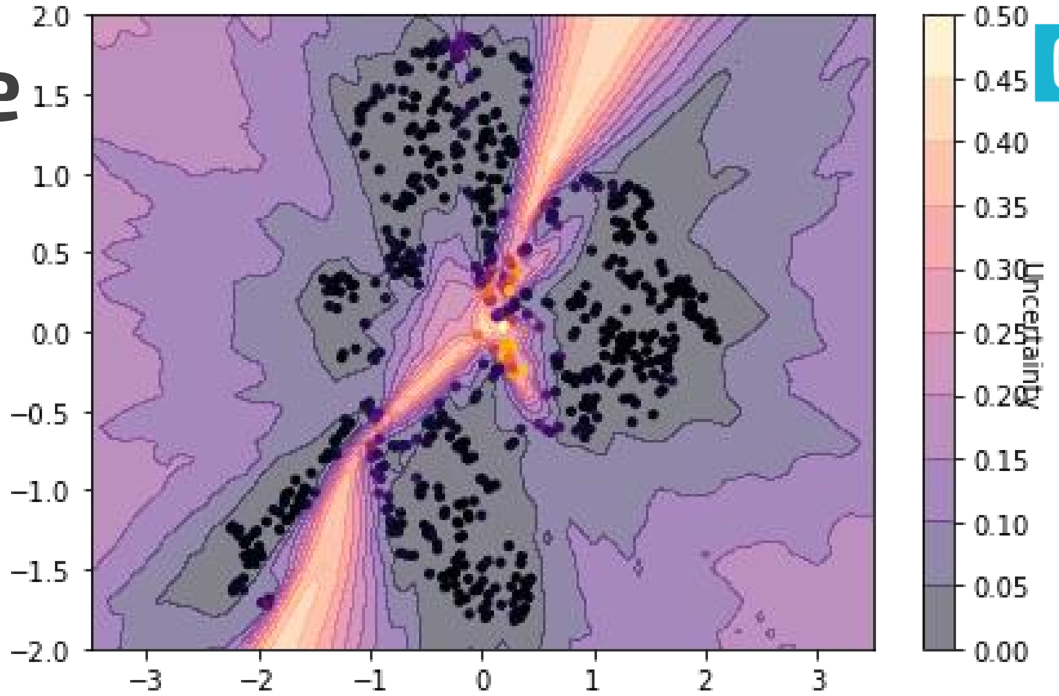
# Models still seem overconfident though

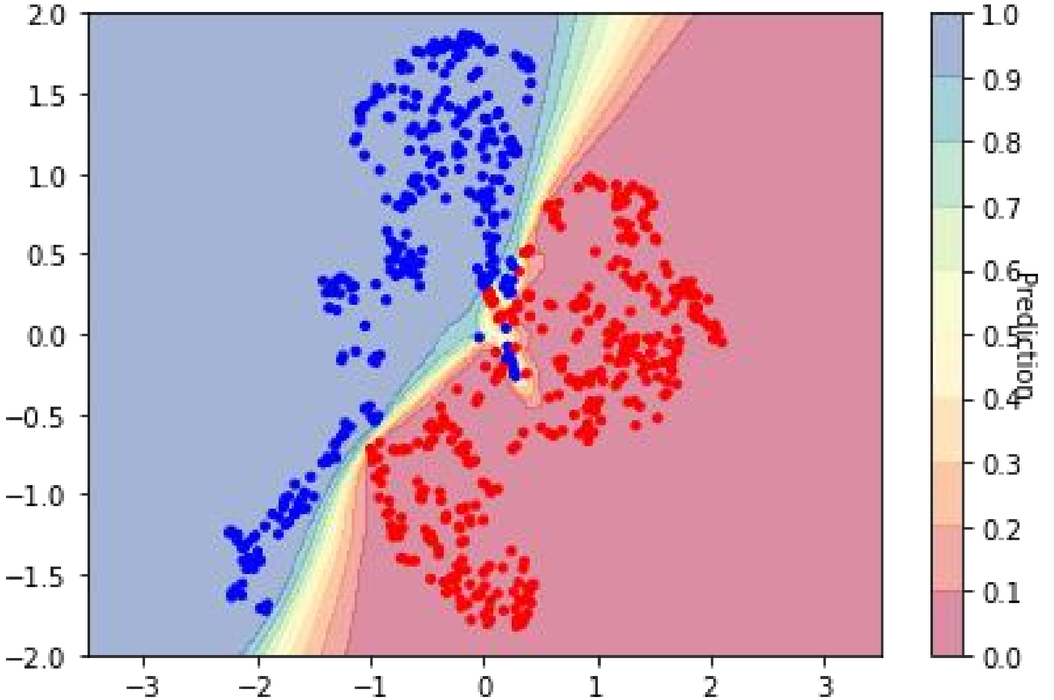



Notice the lack of uncertainty in data-absent regimes

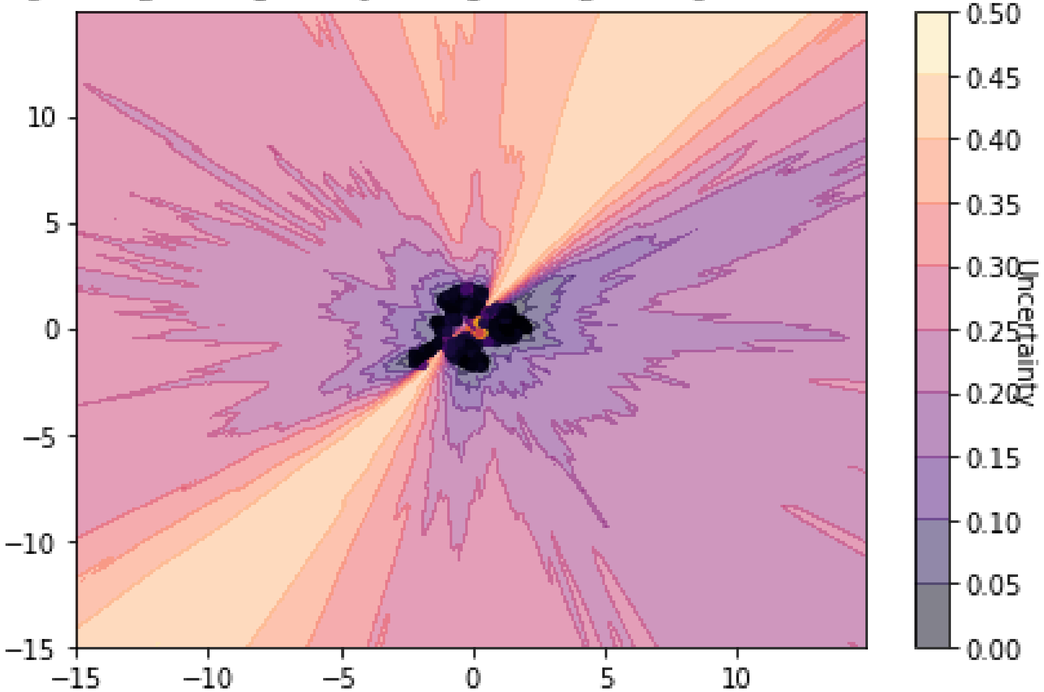
# MCMC: Great and sensible results, but costly

- Captures finer details of data distribution
- Only confident where data is existent
- Long training time: 6 hrs to train vs 1 hr for VI

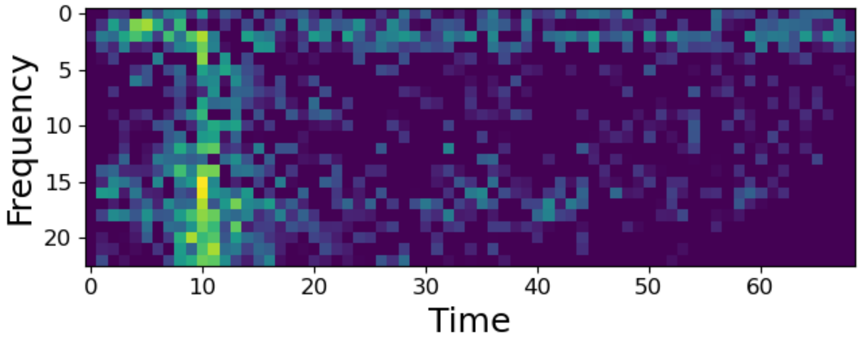


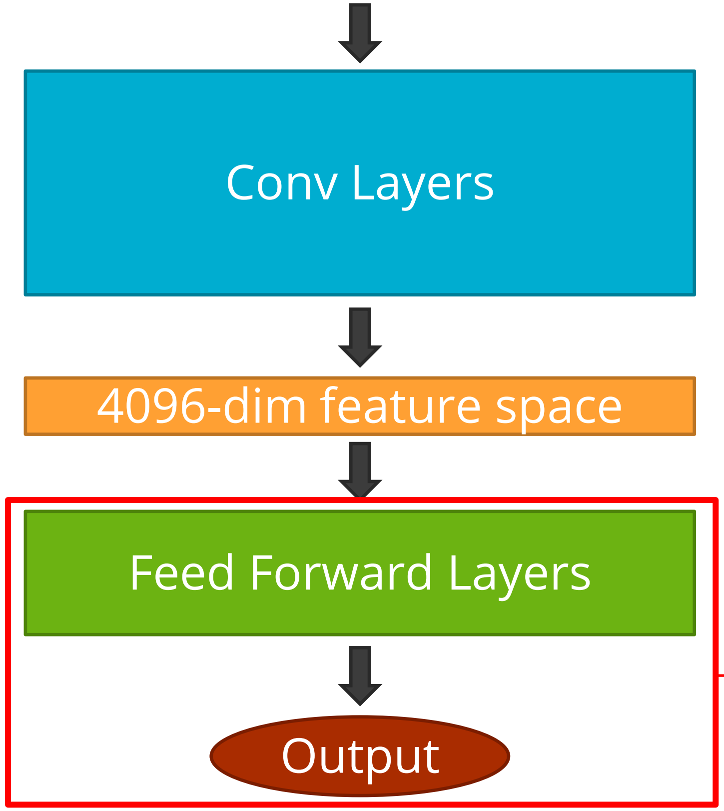




99% acc



# Dataset II: Seismic Waveform Spectrogram Features

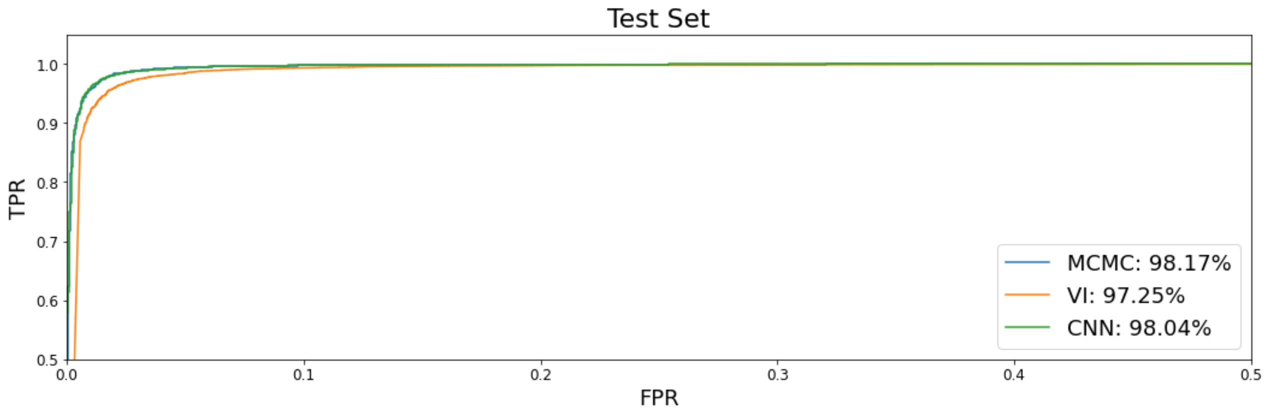



Convert these set weights into a distribution

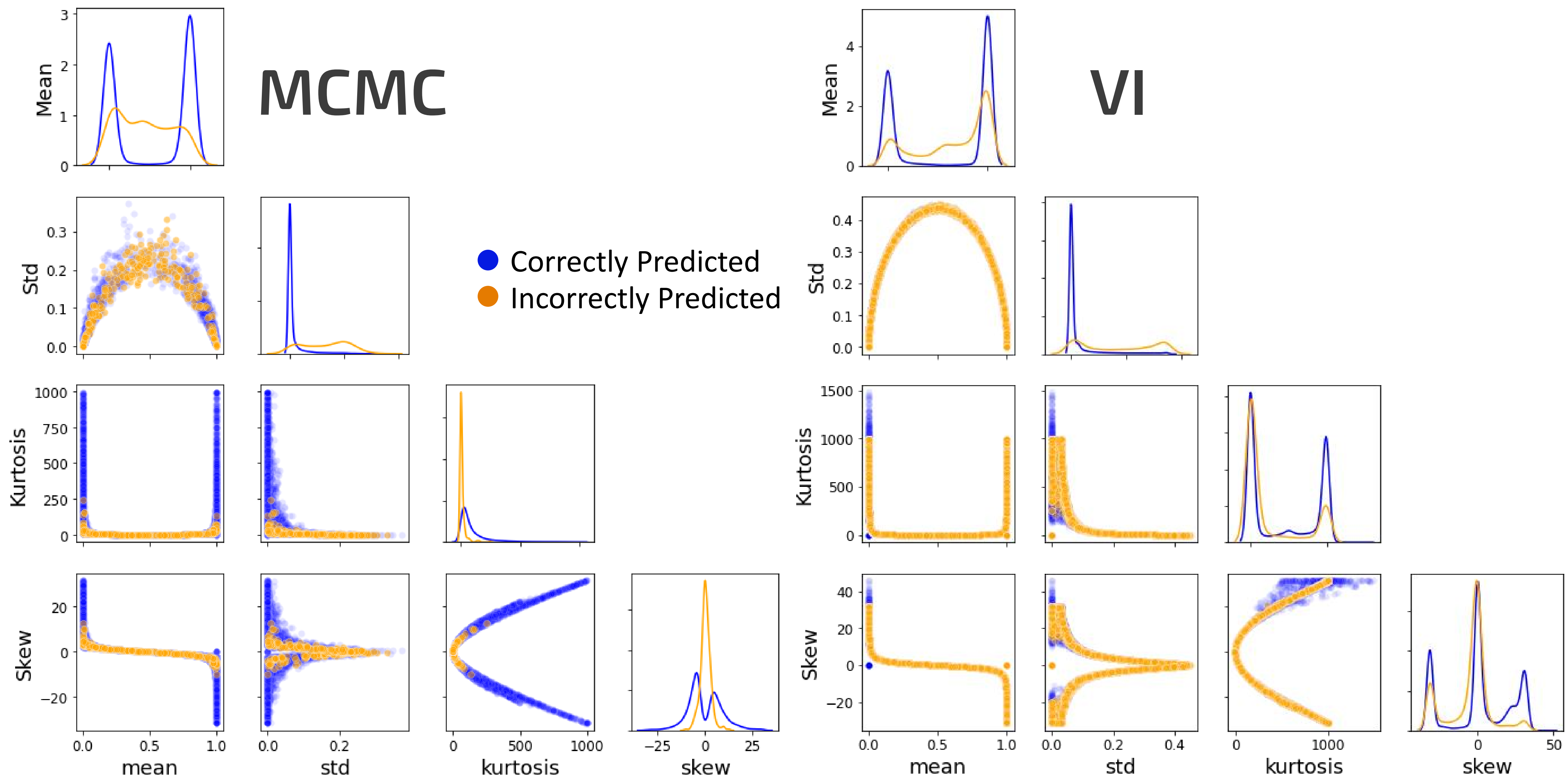
Architecture: VGG Conv NN

- Quarry blasts vs local earthquakes
- Training the whole network in a Bayesian framework would present computational and engineering complexities
- Use Bayesian Inference on only part of the network




- Model performance is still on-par with original baseline despite transferred feature space

# Possibility for exploiting other distribution descriptors



# Conclusion and Next Steps

- So far MCMC has generated more powerful models with more nuanced characterizations of uncertainty
  - The cost is in training time to generate the model (can make or break depending on model or data sample size)
- Currently developing a testbed to compare other methods which generate distributions of predictions
  - MCMC, VI, Bayesian Dropout and Ensemble Methods
- Evaluate performance under a variety of OOD cases
  - Simplest case: Artificial sinusoidal wave
  - More complexity: Seismic events of different nature (e.g. mining induced events, noise)

This research was funded by the National Nuclear Security Administration, Defense Nuclear Nonproliferation Research and Development (NNSA DNN R&D). The authors acknowledge important interdisciplinary collaboration with scientists and engineers from LANL, LLNL, MSTS, PNNL, and SNL.