SAND2019-9599D

Estimating Predictive Uncertainty in Machine Learning Models

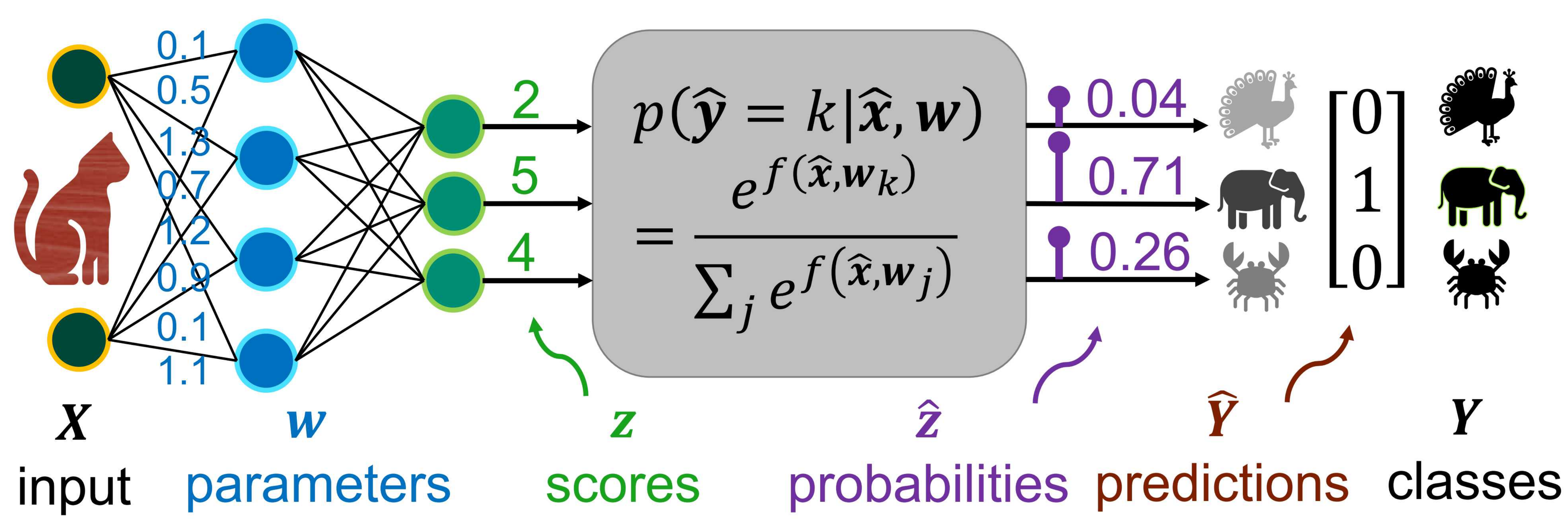
Ahmad A. Rushdi (1463), Optimization and Uncertainty Quantification
Collaborators: Laura P. Swiler (1463), Aubrey C. Eckert (1544), Gabriel Huerta (9136), Brian A. Freno (1544)

Goal: Estimate the predictive uncertainty in machine learning models
from point-estimates to approximate probability distributions

"Predictions without UQ are neither predictions nor actionable."
Begoli E, bhattacharya T, kusnezov D., Nature Machine Intelligence 2019

Problem

In ML/DL classification, an e^x **softmax** layer assigns single-point probabilities to each class



Point Estimates of Neural Networks

tend to be **overconfident** causing unintended and harmful behavior, e.g., when training and test distributions differ, class imbalance, etc.

Input and Model to Output Uncertainties

are not properly characterized/decomposed into:

1) **Aleatoric** (irreducible, e.g., measurement noise)
2) **Epistemic** (reducible, e.g., model parameters)

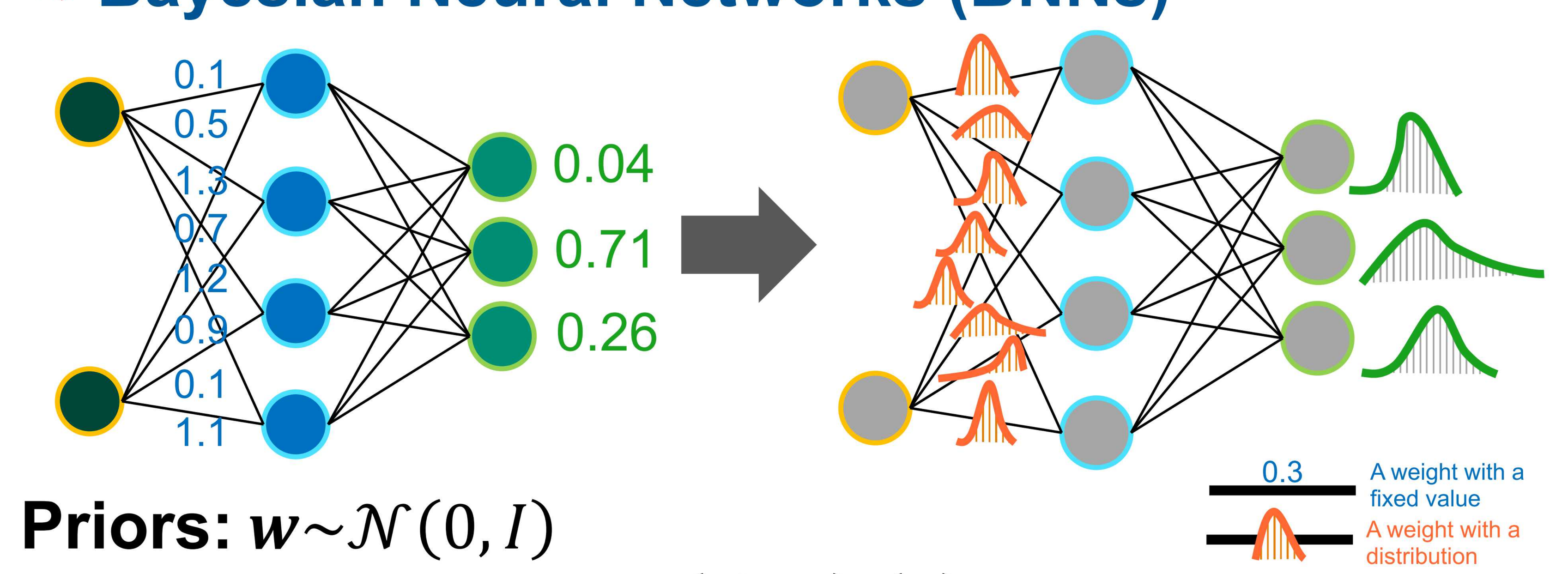
$$w^{\text{MLE}} = \arg\max_w \log p(\mathcal{D}|w) = \arg\max_w \sum_i \log p(Y_i|X_i, w)$$

Maximum Likelihood Estimation

achieved by gradient descent, e.g., back-propagation

Approach

Bayesian Neural Networks (BNNs)



Priors: $w \sim \mathcal{N}(0, I)$

Posterior: $p(w|\mathcal{D}) = \frac{p(Y|X,w)p(w)}{p(Y|X)}$

$$w^{\text{MAP}} = \arg\max_w \log p(w|\mathcal{D}) = \arg\max_w \log p(\mathcal{D}|w) + \log p(w)$$

Maximum a posteriori (MAP)

Gaussian prior: **L2 regularization** (weight decay)
Laplace prior: **L1 regularization**

$$p(\hat{y}|\hat{x}, X, Y) = \int p(\hat{y}|\hat{x}, w) p(w|X, Y) dw$$

Exact inference on w is intractable, **approximation needed**

Results

Variational Inference (VI) Approximation

using a variational distribution $q_\theta(w)$, minimizing the Kullback-Leibler divergence $\mathcal{KL}_{VI}\left(q_\theta(w)||p(w|X,Y)\right)$

$$\theta^* = \arg\min KL\left(q_\theta(w)||p(w|X,Y)\right)$$

integration to optimization

Decomposed predictive uncertainty

Mean $\mathbb{E}_q\{p(\hat{y}|\hat{x})\} \approx \frac{1}{T}\sum_{t=1}^{T} p(\hat{y}|\hat{x}, \theta_t^*)$

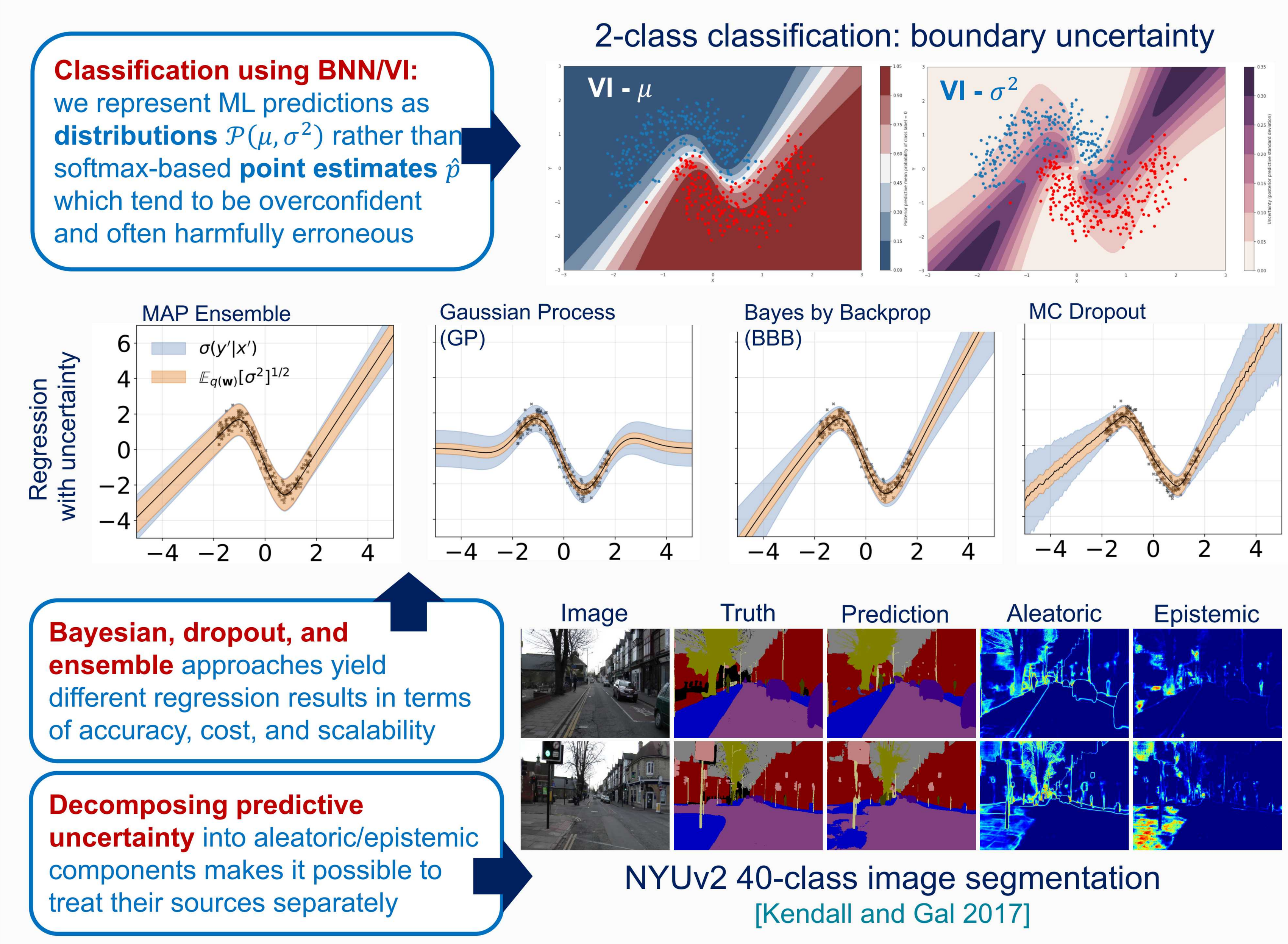
T: MC samples

$\bar{p} = \frac{1}{T}\sum_{t=1}^{T} \hat{p}_t$

Variance $\sigma_q^2\{p(\hat{y}|\hat{x})\} \approx$

$$\frac{1}{T}\sum_{t=1}^{T} \text{diag}(\hat{p}_t) - \hat{p}_t\hat{p}_t^T + \frac{1}{T}\sum_{t=1}^{T} (\hat{p}_t - \bar{p})(\hat{p}_t - \bar{p})^T$$

Aleatoric (irreducible) — **Epistemic** (reducible)

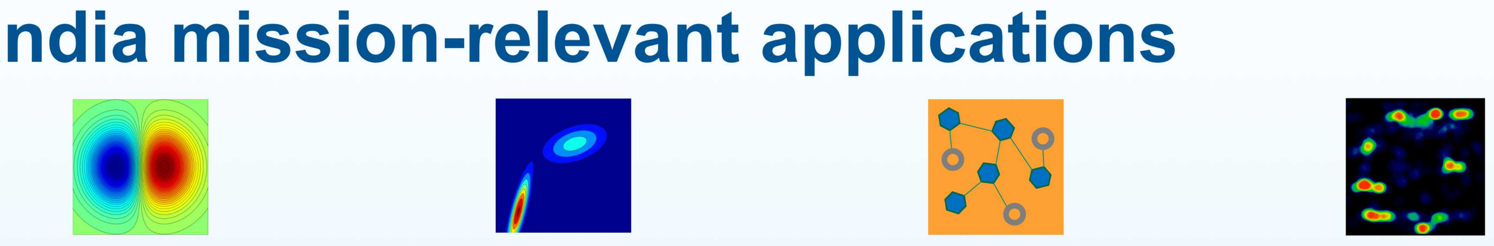


- Presented at Sandia ML/DL Workshop, 08/2019
- **Next**: Multi-fidelity deep ensembles, Sandia applications, and scalability/quality comparisons

Significance

- Enable probabilistic decision making in **high-stakes** and **cost-sensitive** national security, when mistakes cannot be afforded or tolerated
- Evaluate model **trust, reliability,** and **risk** factors, guide **adaptive sampling** for model improvement

Sandia mission-relevant applications



NUMERICAL SIMULATIONS	IMAGE PROCESSING	SENSOR/BIO NETWORKS	HUMAN DIMENSION
Turbulence Modeling	Additive Manufacturing	Graph Neural Networks	Performance Evaluation
Subsurface Modeling	Radar/SAR/X-ray Imagery	Rare Event Prediction	Eye Tracking
Hypersonics	Image Segmentation	Genomic Analysis	Cognitive Load Estimation

Funding
ASC Advanced Machine Learning, FY19-20, 0.5FTE starting 5/2019

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.
SAND No. ____________