Weekly Study Report

Wang Ma

2025-02-25

Electrical, Computer, and Systems Engineering Department
Rensselaer Polytechnic Institute

Outline

(NeurIPS 2024) Are Uncertainty Quantification Capabilities of Evidential Deep Learning a Mirage?

1. (NeurIPS 2024) Analysis of UQ Abilities of Evidential NN

Core Idea. Model the output of a NN to be a distribution, then learn and infer the distribution in with a single model.

1.1.1 The strength of EDL

- Efficient Computation: single forward process produces both prediction and uncertainty
- Good Performance on downstream tasks like OOD detection
- The ability to disentangle uncertainty

1.1.2 Limitations and pitfalls of EDL

- Non-vanishing epistemic uncertainty, even with infinite training data, the estimated Epistemic uncertainty of EDL still exists, meaning the estimation of EU is not reliable
- There is non-existence of a proper loss function for training a distribution on the output space.

For Classification problem, we usually model the output distribution to follow:

$$p_\psi(\pi|x) = \mathrm{Dir}\big(\pi; \alpha_\psi(x)\big).$$

And we usually set a prior, $p_\psi(\pi|x) = \mathrm{Dir}\ (\pi, \mathbb{1})$, which means a uniform distribution on the output.

The author found different kinds of EDL methods have an unified loss function form:

$$L(\psi) = \mathbb{E}_{(x,y)\sim p(x,y)}\Big[D\Big(p^{(\nu)}(\pi|y), p_{\psi(\pi|x)}\Big)\Big] + \lambda \cdot \mathrm{Regularization},$$

where $p^{(\nu)}(\pi|y)$ is the target distribution, and $p_{\psi(\pi|x)}$ is the output distribution of NN.

A Theoretical Conclusion:

If EDL methods do not introduce model diversity, the learned distribution will converge to a "fixed" distribution rather than a accurate or sharp distribution.

1.2 An Unified View of EDL

For example, if we use Dirichlet prior, the model tends to have an "optimal" solution:

$$p_\psi(\pi|x) \longrightarrow \mathrm{Dir}\ (\pi, \alpha_0 + \nu\eta(x)),$$

where $\eta(x)$ is the true label distribution $[p(y = 1|x), ..., p(y = C|x)]$, and α_0 is the hyperparameter of the prior $\mathrm{Dir}(\pi; \alpha_0)$.

This means, when the training data is infinite, EDL's estimation of epistemic uncertainty will not vanish, but be forced to approximate to a Dirichlet distribution with some hyperparametrrs.

- Epistemic uncertainty cannot not vanish as the training data being infinite
- The estimated Aleatoric uncertainty will also be influenced by the hyperparametres, which does not reflect the objective randomness in the data.

The author call the EDL-estimated uncertainties spurious.

1.3 Why EDL succeeds on downstream tasks like OOD?

1. The equivalence to Energy-based Model

   For learned $\alpha_\psi(x)$ of EDL, the sum $\sum_i \alpha_{\psi,i}(x)$ can be seen as the negative log term in energy function, while OOD data make the sum large, and ID data make the sum small.

2. EDL prefers smaller hyperparameters.

   When the hyperparater is proper, the EDL can get sharp prediction on ID data and flat prediction on OOD data, which matches OOD detection.

1.4 What is missing?

The missing components in EDL is **uncertainty/diversity in the model/parameter space**.

Without diversity in model-space, the EDL methods only depend on the the distribution predicted by single parameters, which was trained with some targets.

Conclusion.

This paper theoreticall shows that the uncertainties learned by single-model method EDL are meaningless.

1.4.1 Proposed methods:

Distillation-based EDL (what Hanjing did in his Bayesian-Evidential paper).

2. Unified Uncertainties: Combining Input, Data and Model Uncertainty into a Single Formulation

2.1 Input Uncertainty Propagation

For every input data x, it is determined by a mean and variance: $x_i = \{\mu_i, \sigma_i\}$.

2.1.1 First-order Taylor Expansion Method

Using Taylor Expansion to propogate the input uncertainty to output:

$$f_\theta(\mu_i \pm \sigma_i^2) \approx f_\theta(\mu_i) \pm J\sigma_i^2 J^T,$$

2.1.2 MC Samling

Sample multiple inputs and do multiple times forward process:

$$f_\theta(\mu_i \pm \sigma_i^2) \approx \mathbb{E}[f_\theta(x_i)] \pm \mathrm{Var}[f_\theta(x_i)] \quad x_i \sim N(\mu_i, \sigma_i^2)$$

2.2 Uncertainty Estimation Formulation

An ensemble model: $F(x) = \left[f_{\theta_1}(x), ..., f_{\theta_5}(x)\right]$, where x is determined by (μ, σ),suppose the outpt is logits-form. Then we let

- $\mu^s = \mathbb{E}\left[f_{\theta_i}(x)\right]$ and $\sigma^{2^s} = \mathrm{Var}\left(f_{\theta_i}(x)\right)$

Then the paper present three values:

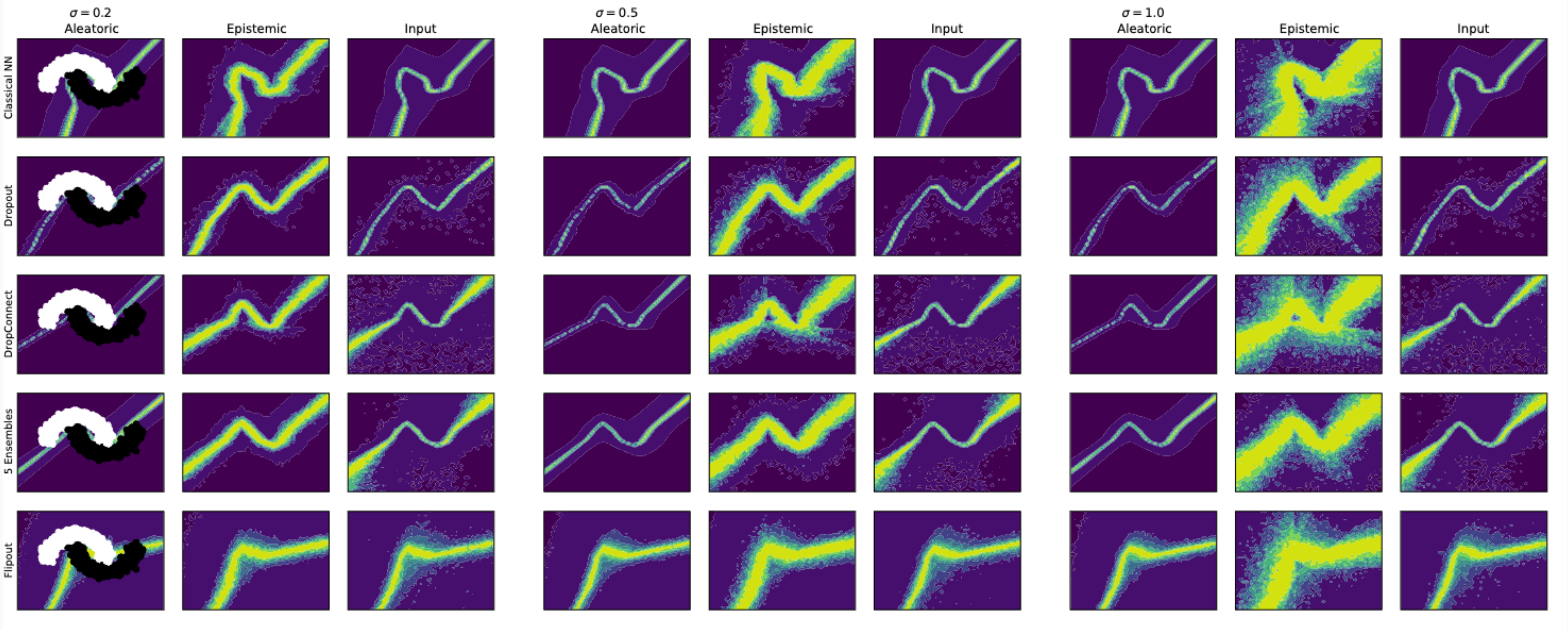
- The prediction: $\mu^o = \mathbb{E}[\mu^s]$
- Variance came from input variance, $\sigma^o_{\text{inp}} = \mathrm{Var}[\mu^s]$
- Variance contributed to epistemic: $\sigma^o_{\text{epi}} = \mathbb{E}\left[\sigma^{s^2}\right]$

Three kinds of predictions considering uncertainties:

- $p_{\text{ale}}(y|x) = \text{softmax }(\mu^o)$
- $p_{\text{inp}}(y|x) = \text{sampling\_softmax }(\mu^o, \sigma_{\text{inp}})$
- $p_{\text{epi}}(y|x) = \text{sampling\_softmax }(\mu^o, \sigma_{\text{epi}})$

where $\text{sampling\_softmax }(\mu(x), \sigma^2(x)) = \frac{1}{N} \sum_j \text{softmax}(\hat{z}_j),\ \hat{z}_j \sim N(\mu(x), \sigma^2(x))$

2.2 Uncertainty Estimation Formulation



2.2 Uncertainty Estimation Formulation

1. A key finding. From the view of source, input uncertainty is similar to aleatoric uncertainty; but the explicit noise in the input propagates, the influence will be shown on Epistemic uncertainty. This means the noise in the input will influence the decision boundary.
2. In the experiments, when considering the input uncertainty, the prediction of the model will be robust from the noise, while the prediction based on epistemic uncertainty loses clear decision boundary.

3. OOD Detection Results on Credal Ensemble:

3. OOD Detection Results on Credal Ensemble:

Instead of using their existing code (their code was written in Tensorflow and I manually convert the code to Pytorch version), I rewrite the logic and algorithms of the Credal Net locally with Resnet18 on Cifar10. And I got the following results.

- Trained Resnet18 on Cifar-10.
- OOD data is SVHN.
- Using Epistemic Uncertainty to do OOD detection

base model: Res18	Accuracy ⬆	AUROC ⬆	AUPRC ⬆
Deep Ensemble	95.88%	0.9406	0.9613
Credal Ensemble	**96.27%**	**0.9774**	**0.9903**
Single Credal Nets	94.79%	0.8693	0.9556