Intelligent
Embedded Systems

Separation of Aleatoric and Epistemic Uncertainty in Deterministic Deep Neural Networks

Denis Huseljic, Bernhard Sick, Marek Herde, Daniel Kottke
{dhuseljic | bsick | marek.herde | daniel.kottke}@uni-kassel.de

University of Kassel, Faculty 16
Electrical Engineering / Computer Science

December 9, 2020

Motivation

Intelligent Embedded Systems

Situation: Standard Deep Neural Networks (DNN) lack in the estimation of uncertainty associated with predictions.

Problem: In autonomous driving, for example, predictions must be reliable to avoid life-threatening situations.

Task: Provide models that measure epistemic uncertainty that describes the reliability of a prediction, and aleatoric uncertainty which describes the risk of a predicted class while maintaining proper generalization capabilities.

Approach: **AE-DNN**, a method for separating Aleatoric and Epistemic uncertainty in DNN.

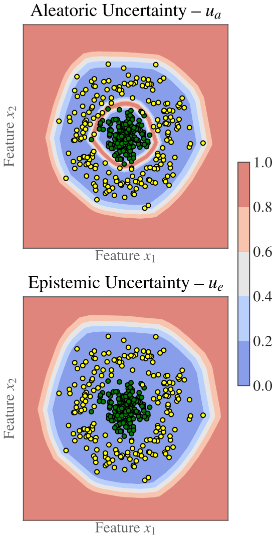
Motivation

AE-DNN can identify the following situations:

(a) *Low epistemic, low aleatoric uncertainty:*
$\rightarrow$ Ideal as predictions are reliable.

(b) *Low epistemic, high aleatoric uncertainty:*
$\rightarrow$ Predictions are reliable, but not distinct (e.g., due to sensor noise).

(c) *High epistemic uncertainty:*
$\rightarrow$ Predictions are unreliable (aleatoric uncertainty can be disregarded).



Characteristics:

- At run-time, AE-DNN allows for a detection of samples that were never seen during training (referred to as OOD detection) or an estimate of the risk coming with a decision.
- For computational efficiency, the inference avoids multiple forward passes through DNN as needed in ensembles, for instance.
- The inference is deterministic in the sense that the same input always leads to the same output (in contrast to Bayesian NN or Monte Carlo dropout).
- OOD samples are generated by means of Generative Adversarial Networks (GAN). As a result, AE-DNN does not require explicitly provided OOD data sets.
- The hyperparameter within the convex combination (OOD vs. ID) allows for a control of the degree of desired certainty in a concrete application.

Comparison

Intelligent Embedded Systems

A comparison to related techniques for uncertainty modeling in DNN is given below. We denote optimization criteria and mark benefits and flaws by $+$ and $-$.

	Epistemic Uncer.	Aleatoric Uncer.	Inference Time	Training Time	ID Optim. Criterion	OOD Optim. Criterion
Ordinary	$-$	$+$	$++$	$++$	MLE	N/A
Ensembles	$+$	$++$	$-$	$--$	MLE	N/A
Dropout	$+$	$+$	$--$	$+$	MLE	N/A
EDL	$--$	$--$	$++$	$++$	Bayes-risk + KL	N/A
PN	$++$	$--$	$++$	$+$	KL	KL
AE-DNN	$++$	$+$	$++$	$+$	MLE	KL

The most similar approach to AE-DNN is Prior Networks (PN) which differs regarding the ID optimization criterion (Kullback-Leibler (KL) in PN and Maximum Likelihood (MLE) in AE-DNN).

Method

We can summarize our method AE-DNN as follows:

A. For ID samples, we optimize the parameters of the DNN such that its output defines a Dirichlet-Categorical distribution over the classes.

B. For OOD samples, we optimize the parameters of the DNN to enforce the Dirichlet distribution over the class probabilities, which is part of the above Dirichlet-Categorical distribution, to be a uniform distribution over the simplex of possible values.

C. Since we obtain a Dirichlet distribution for every sample, we can derive measures describing the heteroscedastic aleatoric and epistemic uncertainty.

The model parameters' optimization is based on a convex-combination of *A* and *B*.

Proposed Uncertainty Measures

The aleatoric uncertainty $u_a \in [0, 1]$ of a sample $\boldsymbol{x}^*$ is given by

$$u_a = \frac{\mathbb{H}\left[\boldsymbol{y}^* | \boldsymbol{\alpha} = \mathrm{f}^{\boldsymbol{\omega}}(\boldsymbol{x}^*) + 1\right]}{\ln K}.$$

The epistemic uncertainty $u_e \in [0, 1]$ of a sample $\boldsymbol{x}^*$ is given by

$$u_e = \frac{K}{\|\mathrm{f}^{\boldsymbol{\omega}}(\boldsymbol{x}^*) + 1\|_1}.$$

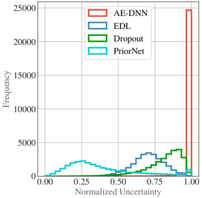
Notation:

- $\mathbb{H}[\cdot]$ denotes the entropy,
- $\mathrm{f}^{\boldsymbol{\omega}}(\boldsymbol{x}^*)$ is the model output for sample $\boldsymbol{x}^*$,
- K is the number of classes, and
- the label $\boldsymbol{y}^*$ is distributed according a Dirichlet-Categorical distribution.

Evaluation

The table summarizes results for SVHN (as ID) vs CIFAR10 (as OOD).

Data Sets	Methods	Generalization Accuracy (↑)	Aleatoric Uncertainty ECE (↓)	NLL (↓)	BS (↓)	Epistemic Uncertainty AUROC (↑)	UH
SVHN vs. CIFAR10	Ordinary	0.875±0.009	**0.012**±0.010	0.440±0.031	0.018±0.001	0.850±0.028	
	Ensembles	**0.900**±0.004	0.046±0.004	**0.361**±0.015	**0.015**±0.001	0.913±0.004	
	Dropout	0.881±0.010	0.015±0.008	0.400±0.026	0.017±0.001	0.921±0.009	
	EDL	0.196±0.000	0.089±0.007	2.291±0.011	0.090±0.000	0.615±0.017	
	PN (OOD gen.)	0.840±0.038	0.107±0.036	0.589±0.134	0.025±0.006	0.933±0.046	
	AE-DNN (OOD gen.)	0.859±0.014	0.014±0.009	0.485±0.038	0.021±0.002	0.970±0.017	
	PN (OOD av.)	0.882±0.009	0.101±0.031	0.468±0.032	0.019±0.001	0.993±0.002	
	AE-DNN (OOD av.)	0.879±0.011	0.019±0.005	0.427±0.033	0.018±0.001	**0.997**±0.001	



For further details and additional results, we refer to our implementation which is available at `https://github.com/hsljc/ae-dnn`.

Questions?