The University of Texas at Austin
Department of Statistics and Data Sciences
College of Natural Sciences

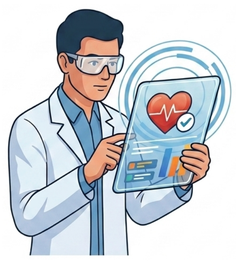
CosmicAI
Funded by the U.S. National Science Foundation and the Simons Foundation

Density-Informed Pseudo-Counts for Calibrated Evidential Deep Learning

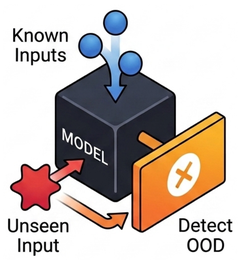
Pietro Carlotti • **Nevena Gligić** • **Arya Farahi**
Contact: pietro.carlotti@utexas.edu

Motivation

Predictions are important, but calibrated **uncertainty** is essential for:



Safety-critical decision-making



Detection of unseen inputs

Yet, modern classifiers are **overconfident** under limited data or covariate shift.

Evidential Deep Learning

EDL models probabilities via a neural-network-parameterized Dirichlet distribution:

$$p \sim \mathrm{Dir}\left(\alpha + \mathrm{NN}_X^{\phi}\right).$$

Training is done by minimizing a loss balancing data fit and regularization:

$$\mathcal{L}_{\mathrm{EDL}}^{\lambda}(\phi) = \mathcal{L}_{\mathrm{data}}(\phi) + \lambda\, \mathcal{L}_{\mathrm{reg}}(\phi).$$

A single forward pass yields both a **prediction** and an **uncertainty** quantification.

Statistical Interpretation

We cast EDL as a Bayesian model with a **tempered** pseudo-likelihood:

$$Y_i \mid p_i \overset{\text{ind.}}{\propto} \prod_{k=1}^{K} p_i(k)^{\nu\, \mathbb{1}(Y_i=k)}, \quad i = 1, \ldots, n,$$

$$p_i \overset{\text{i.i.d.}}{\sim} \mathrm{Dir}(\alpha), \quad i = 1, \ldots, n.$$

Conjugacy yields a closed-form **posterior** given the observed labels:

$$\pi_{Y_i}^{\nu} = \mathrm{Dir}(\alpha + \nu\, e_{Y_i}), \qquad i = 1, \ldots, n.$$

We induce **covariate dependence** by approximating the posterior via mean-field amortized **variational inference** (VI), parameterized by a neural network:

$$q_{X_{1:n}}^{\phi} = \prod_{i=1}^{n} q_{X_i}^{\phi}, \quad \text{where} \quad q_{X_i}^{\phi} = \mathrm{Dir}\left(\alpha + \mathrm{NN}_{X_i}^{\phi}\right), \quad i = 1, \ldots, n.$$

Then, parameters are learned by minimizing the KL divergence to the posterior:

$$\widehat{\phi}_n = \operatorname*{argmin}_{\phi \in \Phi} \mathrm{KL}\left(q_{X_{1:n}}^{\phi} \,\|\, \pi_{Y_{1:n}}^{\nu}\right).$$

Minimizing this amortized VI objective is exactly equivalent to EDL training:

$$\widehat{\phi}_n = \operatorname*{argmin}_{\phi \in \Phi} \mathcal{L}_{\mathrm{EDL}}^{\frac{1}{\nu}}(\phi; X_{1:n}, Y_{1:n}).$$

In EDL, epistemic uncertainty of a Dirichlet distribution is measured by **vacuity**:

$$\mathrm{Vacuity}(\alpha) = \frac{K}{\alpha_0}, \qquad \text{where} \qquad \alpha_0 = \sum_{k=1}^{K} \alpha(k).$$

Even when the oracle distribution is available, uncertainty remains constant:

$$\mathrm{Vacuity}\left(\alpha + \mathrm{NN}_X^{\phi^\star}\right) = \frac{K}{\alpha_0 + \nu}.$$

EDL's uncertainty is therefore **spurious** because it is hyperparameter-driven.

DIP-EDL

We propose replacing ν with a local, density-informed pseudo-count:

$$q_{X_i}^{\psi,\phi} = \mathrm{Dir}\left(\alpha + n\, \mathrm{DE}_{X_i}^{\psi}\, \mathrm{NN}_{X_i}^{\phi}\right), \qquad i = 1, \ldots, n,$$

where

- n governs the degree of concentration.
- DE^{ψ} learns the true covariate density P_X^*.
- NN^{ϕ} learns the true conditional class distribution $P_{Y|X}^*$.

Salient Properties

Our method enjoys some desirable properties:

- **Distributional awareness**: predictions contract to the prior under scarce data:

$$\mathrm{DE}_{X_i}^{\psi} \approx 0 \implies q_{X_i}^{\psi,\phi} \approx \mathrm{Dir}(\alpha).$$

- **Asymptotic concentration**: consistent estimators yield converging predictions:

$$\mathrm{DE}_{X_i}^{\psi} \xrightarrow{p} P_X^*(X_i), \;\; \mathrm{NN}_{X_i}^{\phi} \xrightarrow{p} P_{Y|X}^*(\cdot \mid X_i) \implies q_{X_i}^{\psi,\phi} \xrightarrow{p} P_{Y|X}^*(\cdot \mid X_i).$$

- **Modularity**: the density estimator and classifier can be trained independently:

$$\hat{\psi}_n = \operatorname*{argmin}_{\psi \in \Psi} \mathcal{L}_{\mathrm{DE}}(\psi; X_{1:n}), \qquad \hat{\phi}_n = \operatorname*{argmin}_{\phi \in \Phi} \mathcal{L}_{\mathrm{NN}}(\phi; X_{1:n}, Y_{1:n}).$$

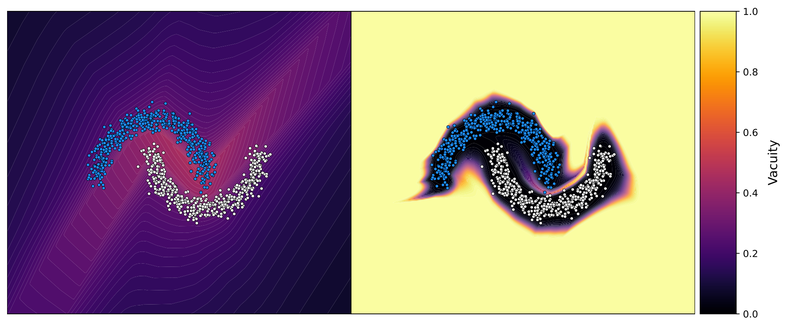
- **Architecture-agnosticism**: it requires no specific estimator or classifier.

Experiments

We evaluate DIP-EDL on simulated, benchmark, and scientific data.

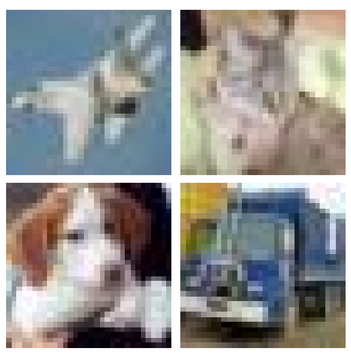
Simulated Data

We visualize the vacuity of EDL and DIP-EDL on synthetic **two-moons** data:



Benchmark Data

We train models on **CIFAR-10** (ID), evaluating on **SVHN** (OOD).

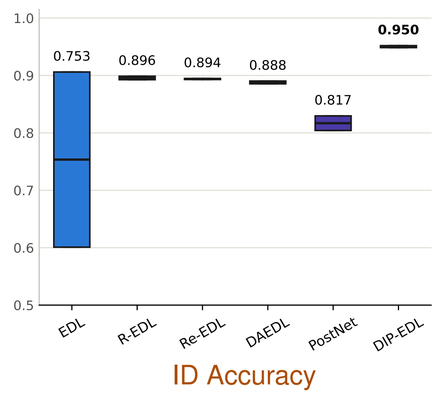


CIFAR-10

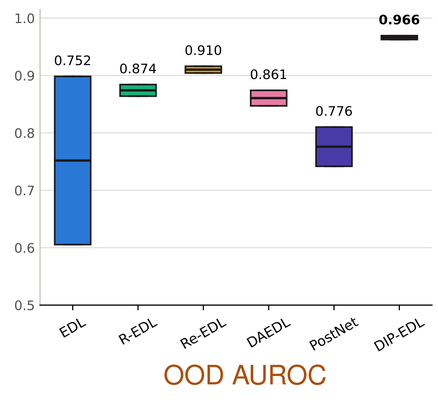


SVHN

We assess prediction with **accuracy** and uncertainty quantification with **AUROC**:



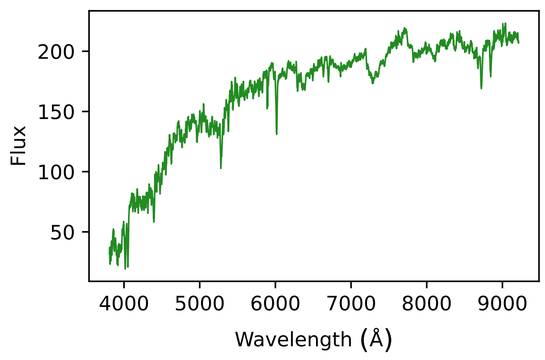
ID Accuracy



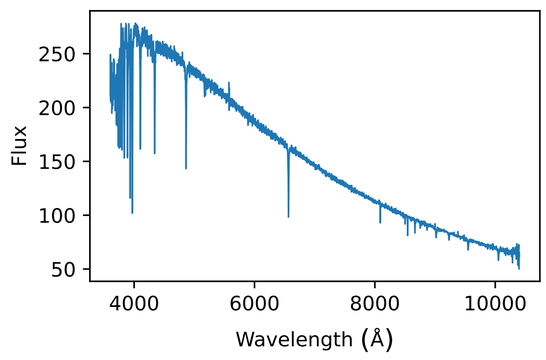
OOD AUROC

Scientific Data

We train models on **LAMOST Galaxy/Quasar** (ID), evaluating on **Star** (OOD).

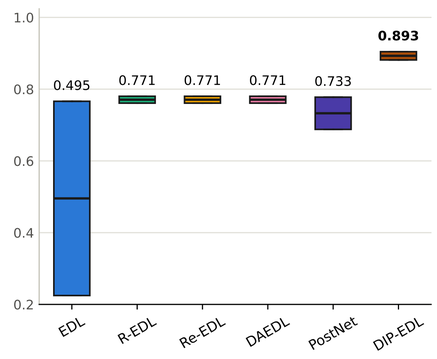


Galaxy

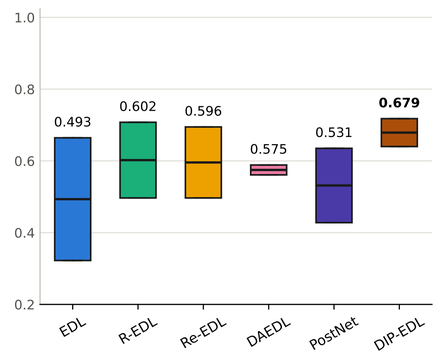


Star

We assess prediction with **accuracy** and uncertainty quantification with **AUROC**:



ID Accuracy



OOD AUROC

Ablation Study

We ablate DIP-EDL's three components to isolate their contributions.

n	DE^{ψ}	NN^{ϕ}	CIFAR-10		
			ID Accuracy	ID BS	OOD AUROC
✓	✓	✓	0.9496	0.0978	0.9686
×	✓	✓	0.9496	0.7653	0.9686
✓	×	✓	0.9496	0.0831	0.5014
✓	✓	×	0.1000	0.9000	0.9686

This reveals three distinct roles:

- Removing n compromises **ID BS**.
- Removing DE^{ψ} compromises **OOD AUROC**.
- Removing NN^{ϕ} compromises **ID Accuracy** and **ID BS**.

References

Sensoy, M., Kaplan, L., and Kandemir, M. Evidential deep learning to quantify classification uncertainty. *Advances in Neural Information Processing Systems*, 31, 2018.