Uncertainty Quantification in Graph Neural Networks

Albert Kjøller Jacobsen (s194253), Aron Djurhuus Jacobsen (s194262),
Felix Borum Burmeister (s194266) and Phillip Chavarria Højbjerg (s184984)

- Human Centered Artificial Intelligence, DTU Compute, Technical University of Denmark

Introduction

Deep Neural Networks (DNNs) have demonstrably shown increased predictive performance when modeling complex regression and classification tasks due to their data-driven nature. Though DNNs can mostly solve the prediction task the naive implementation does not contain appropriate mechanisms for estimating model uncertainty. For several down-stream tasks - such as medical treatment classification or prediction of molecular properties used for medical development - this is a major flaw of DNNs as a highly certain prediction from such a model might have critical costs for the deployment task. In general the total uncertainty of a model can be split into two; i) *aleatoric* uncertainty capturing uncertainty of the data and ii) *epistemic* uncertainty capturing how uncertain a model is about its predictions. While *aleatoric* uncertainty can trivially be obtained from training data, approaches such as Bayesian Neural Networks (BNNs) have been used for modeling epistemic uncertainty as they place probabilistic priors on the network parameters, however BNNs are cumbersome to train. Another approach is known as *Evidential Learning* (EL) where a higher-order, evidential distribution is fitted through an evidence acquisition process. Sampling the evidential distribution yield lower-order likelihood functions. As such EL provides a framework for training uncertainty-aware predictive DNNs.

Key points

- We train a DNN outputting parameters of the **Normal Inverse-Gamma (NIG) distribution** and reproduce use the EL framework for reproducing the 1-dimensional (1D) example of fitting a 3rd order polynomial as done by Amini et al. [1]
  - We evaluate **epistemic uncertainty estimates** using calibration plots and entropy on two test sets; one based on in-domain (ID) test points and one based on out-of-domain (OOD) test points. We compare to a **Gaussian MLE**.
  - We find that the proposed framework can learn to associated high uncertainty to OOD prediction, but that it is **highly sensitive** to the evidential regularization.
- We create a simple **new synthetic graph dataset** simulating 3-dimensional (3D) molecules by sampling. We train and evaluate it for predicting the synthetic potential energy function using a similar approach to the 1D-example.
- We apply the evidential learning approach on the **QM7** dataset of 3D molecular graphs and use it for solving a **molecular property prediction** task estimating the total energy of a molecule. We compare it to a mean-predicting baseline.
- We exploit **invariant message-passing networks** to encode 3D graphs (both synthetic and molecular) into their respective graph representation from the original atomic coordinates.

Normal Inverse-Gamma (NIG) distribution

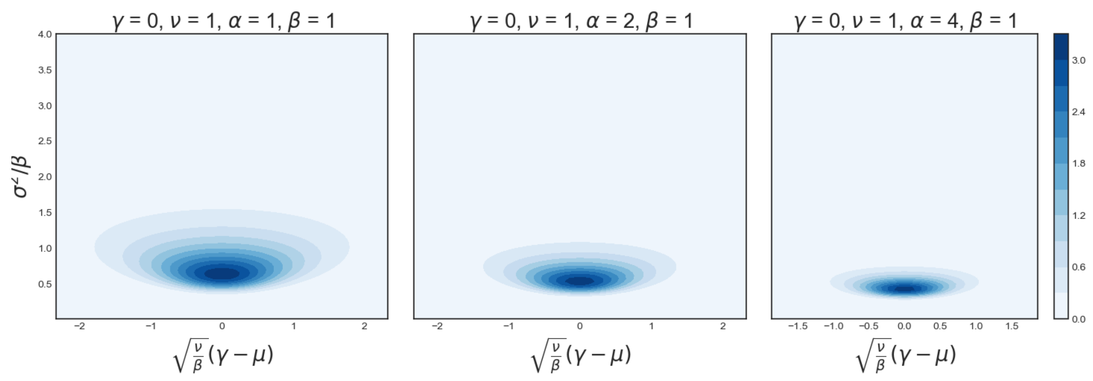

Figure 1: The Normal Inverse-Gamma distribution as the higher-order distribution with varying $\boldsymbol{\alpha}$

$$p(\underbrace{\mu, \sigma^2}_{\boldsymbol{\theta}} | \underbrace{\gamma, \upsilon, \alpha, \beta}_{\boldsymbol{m}}) = \frac{\beta^\alpha \sqrt{\upsilon}}{\Gamma(\alpha)\sqrt{2\pi\sigma^2}} \left(\frac{1}{\sigma^2}\right)^{\alpha+1} \exp\left\{-\frac{2\beta + \upsilon(\gamma-\mu)^2}{2\sigma^2}\right\}$$

$$\underbrace{\mathbb{E}[\mu] = \gamma}_{\text{prediction}}, \quad \underbrace{\mathbb{E}[\sigma^2] = \frac{\beta}{\alpha-1}}_{\text{aleatoric}}, \quad \underbrace{\mathrm{Var}[\mu] = \frac{\beta}{\upsilon(\alpha-1)}}_{\text{epistemic}}.$$

Model performance - 1D

Table 1 and Figure 2 show the results obtained by comparing the Gaussian MLE with the EL-based model. We observe that the the model perform on par in terms of RMSE with higher RMSE for OOD. The EL model learns to associate a high epistemic uncertainty for OOD while the Gaussian MLE does not to a similar extent.

Table 1: Results on the 3rd order polynomial.

Method	RMSE	NLL	σ^2
In-Domain			
Gaussian MLE	3.6339	1.3324	18.6463
Evidential	3.7485	9.5130	8.9698
Out-of-Domain			
Gaussian MLE	56.5099	5.6624	47.9385
Evidential	58.0255	12.2256	2198.7055

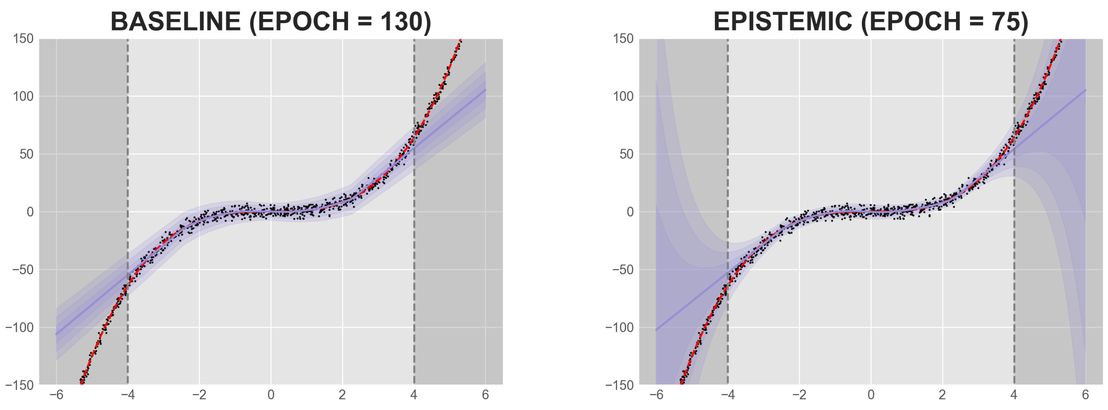

Figure 2: Model uncertainties visualized on the 1D example. Dark grey areas are OOD.

Uncertainty Quantification - 1D

On Figure 3 (left) we observe that both the Gaussian MLE and EL-model learns to be relatively more uncertain when the respective error is high. This is in line with Figure 2, where both models become more uncertain for OOD data. However, Figure 3 (right) reveals that the entropy for the EL-model is factors higher for OOD than ID. While the same trend is observed for the baseline, the trend is not as evident.

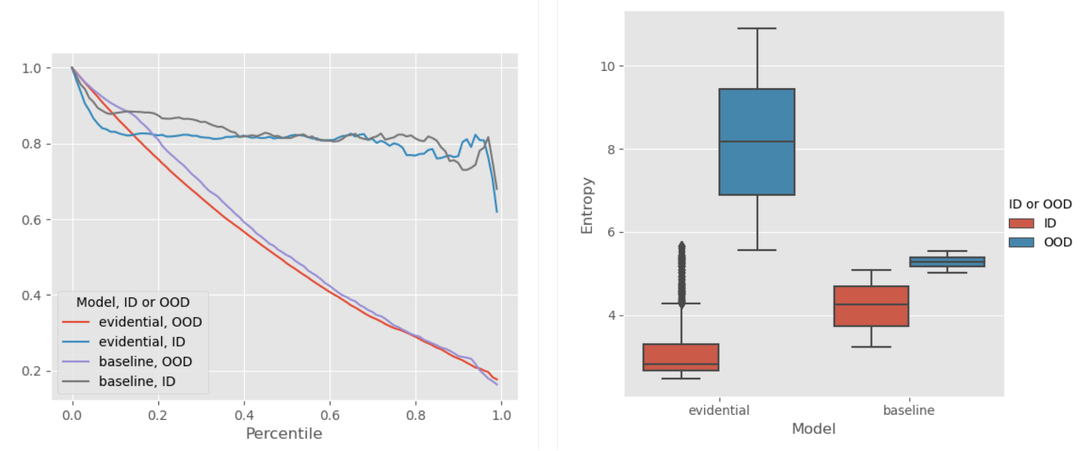

Figure 3: Uncertainty quantification for the Gaussian MLE and EL-model - both for ID and OOD data. *Left:* Calibration plots - normalized RMSE per percentile of $\boldsymbol{\sigma^2}$. *Right:* Entropy measures.

Synthetic data - fitting an invariant GNN

The synthetically generated data set - consisting of an ID and an OOD part - is presented in Figure 4. The average simulated potential energy, *U(x)*, is generally lower for OOD data (blue) than for ID data (red) in order to show similar properties of evidential learning as for the 1D example.

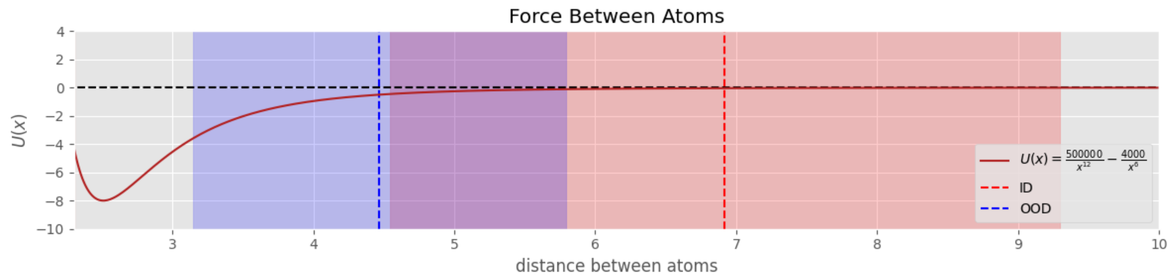

Figure 4: The potential energy function, *U(x)*. A synthetic data set is generated as a combination of ID and OOD samples. The average distance between atoms and the associated standard deviations for ID and OOD data is shown in red and blue, respectively.

Table 2 shows the performance of an Evidential and a Gaussian MLE based GNN model on the synthetic dataset. In general both models have increased values on the Out-of-Domain data compared to In-Domain which is expected considering the task at hand. However, most notably the Evidential model has some unrealistic $\boldsymbol{\sigma^2}$ values.

Table 2: Results on the synthetic data.

Method	RMSE	NLL	σ^2
In-Domain			
Gaussian MLE	7.0681	1.5678	51.9137
Evidential	5.2799	3.1984	$\mathbf{2.3919 \cdot 10^{11}}$
$\mu_{\text{data}=\mathbf{-9.0756}}$	8.4726	-	-
Out-of-Domain			
Gaussian MLE	8.3570	1.6224	53.4266
Evidential	9.9351	3.7365	$\mathbf{2.3266 \cdot 10^{11}}$
$\mu_{\text{data}=\mathbf{-25.8228}}$	14.9330	-	-

QM7 dataset - fitting an invariant GNN

Table 3 shows the results from an evidential and a gaussian based GNN model trained on the QM7 dataset [2, 3] aimed at predicting the atomization energies - i.e. the extra energy needed to break up a molecule. It is observed that the gaussian MLE GNN has a better score than the Evidential GNN on all measurements. This could be due to poor optimization, especially considering the $\boldsymbol{\sigma^2}$-values where the Evidential model was expected to be superior. On the other hand, the RMSE values are relatively high when compared to other papers of e.g. 5.95[4].

Table 3: Results on the QM7 dataset

Method	RMSE	NLL	σ^2
Gaussian MLE	14.2349	1.7824	310.7302
Evidential	19.3000	7.0335	$9.9999 \cdot \mathbf{10^9}$
$\mu_{\text{data}=\mathbf{-1538.0377}}$	233.9189	-	-

Acknowledgements

We wish to thank Mikkel N. Schmidt for his generous attention and supervision on this project. Even on sick days he was able to set time aside and help us on our way.

References

[1] A. Amini, W. Schwarting, A. Soleimany, and D. Rus. Deep evidential regression. 2020.

[2] L. C. Blum and J.-L. Reymond. 970 million druglike small molecules for virtual screening in the chemical universe database GDB-13. *J. Am. Chem. Soc.*, 131:8732, 2009.

[3] M. Rupp, A. Tkatchenko, K.-R. Müller, and O. A. von Lilienfeld. Fast and accurate modeling of molecular atomization energies with machine learning. *Physical Review Letters*, 108:058301, 2012.

[4] A. B. Tchagang and J. J. Valdes. Prediction of the atomization energy of molecules using coulomb matrix and atomic composition in a bayesian regularized neural networks, 2019.

s194253@student.dtu.dk s194262@student.dtu.dk s194266@student.dtu.dk s184984@student.dtu.dk