KU LEUVEN LEUVEN.AI

Epistemic AI — A new learning paradigm

SCIENCE ↓ AI

Science of AI – Total Uncertainty in AI

Keivan Shariatmadar

Mecha(tro)nic System Dynamics (LMSD), KU Leuven, Belgium

Leuven.AI scientific workshop 2024

keivan.shariatmadar@kuleuven.be

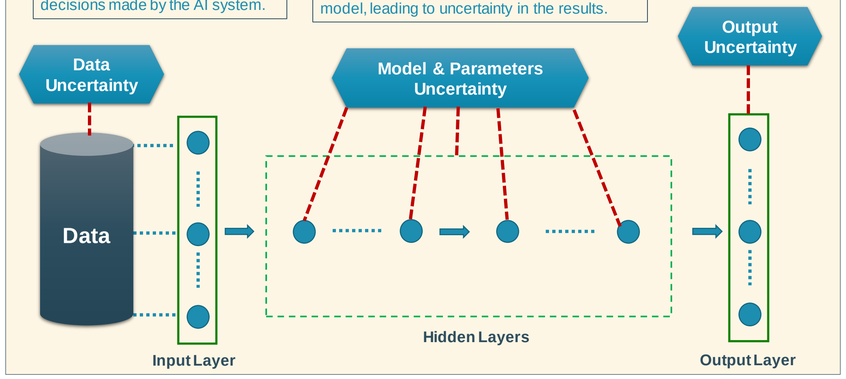
Abstract

AI has been addressing uncertainty to achieve highly accurate results. This challenge becomes more pronounced with smaller data sets or variations within them, significantly impacting decision-making, forecasting, and learning mechanisms. This study explores the nature of uncertainty in AI by integrating ideas from established theories, recent developments, and practical applications, ultimately proposing a new comprehensive definition of total uncertainty in AI. From foundational theories to current methodologies, this work offers a comprehensive perspective on managing total uncertainty and its complexities in AI, helping us understand its significance and value across various domains.

Sources of Uncertainty in AI

AI models are trained on data, and the quality and accuracy of the data can affect the performance of the model. **Noisy** or **incomplete data** can lead to uncertain predictions or decisions made by the AI system.

AI models are complex and can have various **parameters** and **hyperparameters** that need to be tuned. The choice of **model architecture**, **optimisation** algorithm, and hyperparameters can significantly impact the performance of the model, leading to uncertainty in the results.



Standard Neural Network – Cross-Entropy

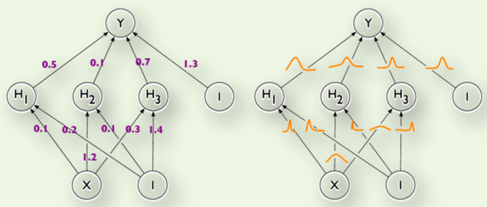
$$H(P,Q) \coloneqq S(P) + KL(P||Q)$$

$$S(P) \coloneqq -\sum_{x \in X} P(x).\log(P(x))$$

$$KL(P||Q) \coloneqq \sum_{x \in X} P(x).\log(Q(x)/P(x))$$

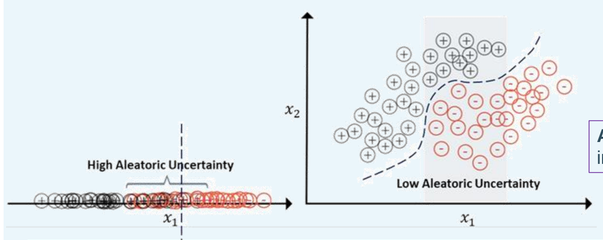
$H(P,Q)$ is the cross-entropy of Q from P, $S(P)$ is the entropy of P and $KL(P||Q)$ is the divergence of Q from P.

Bayesian Neural Network (BNN)



Left: each weight has a fixed value, as provided by classical backpropagation. Right: each weight is assigned a distribution, as provided by Bayes by Backprop.

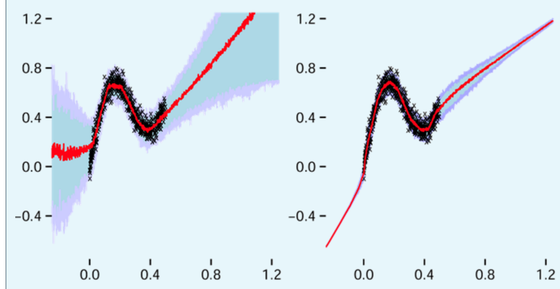
Supervised Learning



Aleatoric Uncertainty (AU) in Supervised Learning

Adding more variables to prediction models can result in decreasing the aleatoric uncertainty, but it modifies the problem set.

Regression – Bayes vs Standard NN



Epistemic Uncertainty (EU) in Regression

Regression of noisy data with interquartile ranges. Black crosses are training samples. The red lines are median predictions. The blue/purple region is the interquartile range. Left: Bayes by Back prop neural network, Right: standard neural network.

Discussion

- Using EnNN, take the highest and lowest prediction as $\boldsymbol{\alpha_2}$?
- In INN, RSNN, and CreNN $\alpha_2 \geq H^* - H_*$?
- Could contamination model be used to define better Total Uncertainty?
- How about using the Expectation rather than Entropy?
- Why do we need Aleatoric Uncertainty, is it necessary?
- Test in regression with the defined true model.



Epistemic AI — A new learning paradigm

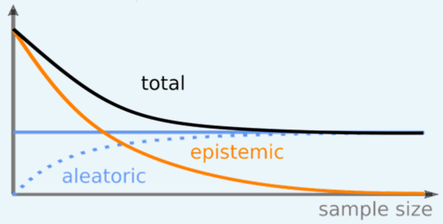
Epistemic & Aleatoric Uncertainties

Aleatoric vs Epistemic

Aleatoric (statistical) uncertainty refers to randomness or variability: What a random sample drawn from a probability distribution will be.

Epistemic (systematic) uncertainty refers to the lack of knowledge: What the relevant probability distribution is.

Total uncertainty $(\boldsymbol{TU}) := \boldsymbol{EU} + \boldsymbol{AU}$?



expected qualitative behaviour of uncertainty components. The dotted line marks the difference between TU and EU (which equals AU when stipulating additivity)

Dependency

The **mean** of two of the same data sets with **different noises** are different:

$$\boldsymbol{E(X_1) \neq E(X_2)} \quad X_1 := f(x) + \varepsilon_1 \sim N(\mu_1, \sigma_1), \quad X_2 := f(x) + \varepsilon_2 \sim N(\mu_2, \sigma_2)$$

Then AU and EU are **dependent**.

$$\boldsymbol{TU \coloneqq \alpha_1 AU + \alpha_2 EU, \qquad \alpha_1 + \alpha_2 > 1}$$

How to define the $\boldsymbol{\alpha_1}$ and $\boldsymbol{\alpha_2}$?
Investigate if the model is robust against increased noise or decreased data size.
Calculate H^* and H_* then $\alpha_2 \geq H^* - H_*$?
$\alpha_2 \geq U_{EU} - L_{EU}$?

References

Hüllermeier and W. Waegeman. Aleatoric and Epistemic Uncertainty in Machine Learning: An Introduction to Concepts and Methods. Machine Learning, 2021.

Peter O. Hristov, Ioanna Ioannou and Scott Ferson. "How to Exploit What We Know About Input and Model: A Trans-probabilistic Approach to the 2022 AIAA UQ Challenge," AIAA 2024-0944. AIAA SCITECH 2024 Forum. January 2024.

Hüllermeier, S. Destercke, and M. H. Shaker. Quantification of Credal Uncertainty in Machine Learning: A Critical Analysis and Empirical Comparison. In UAI, 2022.

L. Wimmer, Y. Sale, P. Hofman, B. Bischl, E. Hüllermeier. Quantifying Aleatoric and Epistemic Uncertainty in Machine Learning: Are Conditional Entropy and Mutual Information Appropriate Measures? 2023 arXiv:2209.03302v2

C. Blundell, J. Cornebise, K. Kavukcuoglu, D. Wierstra. Weight Uncertainty in Neural Networks. 2015 arXiv:1505.05424v2

Dubois and E. Hüllermeier. Comparing probability measures using possibility theory: A notion of relative peakedness. International Journal of Approximate Reasoning, 45(2):364–385, 2007

This work is supported by the European Union's Horizon 2020 research and innovation program under grant agreement No. 964505 (E-pi).