Sparse Function-space Representation of Neural Networks

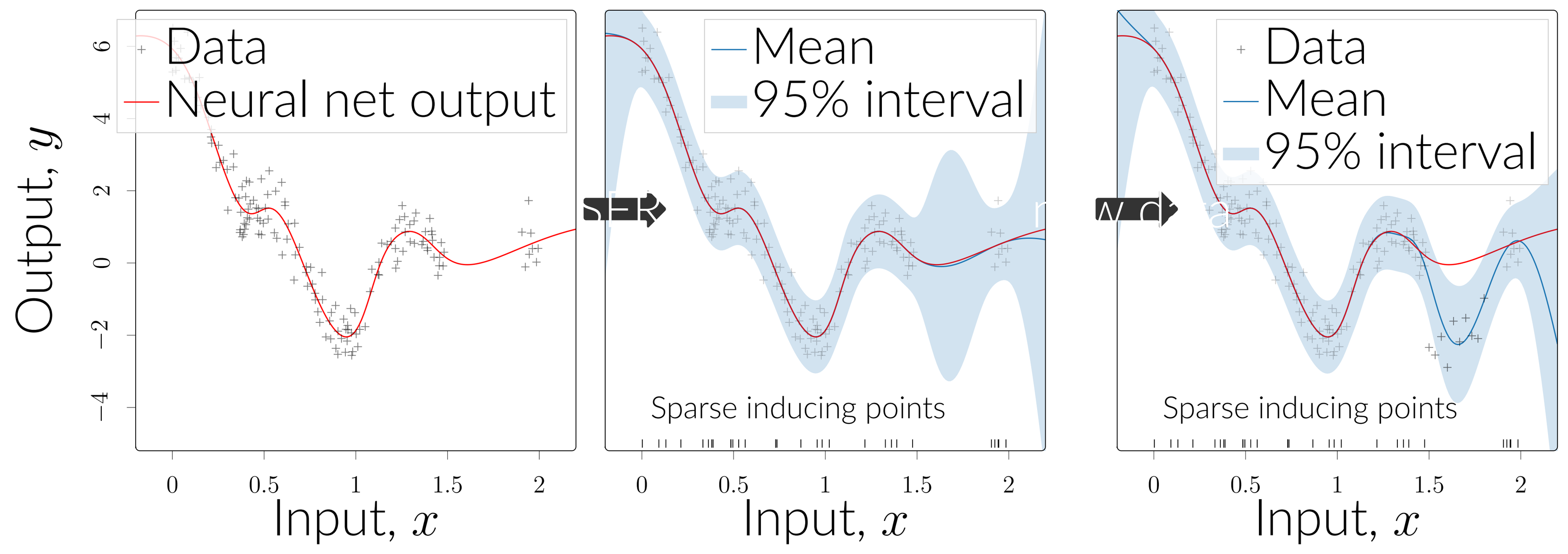
Aidan Scannell*[1,2] Riccardo Mereu*[1] Paul Chang[1]
Ella Tamir[1] Joni Pajarinen[1] Arno Solin[1]

[1]Aalto University [2]Finnish Center for Artificial Intelligence

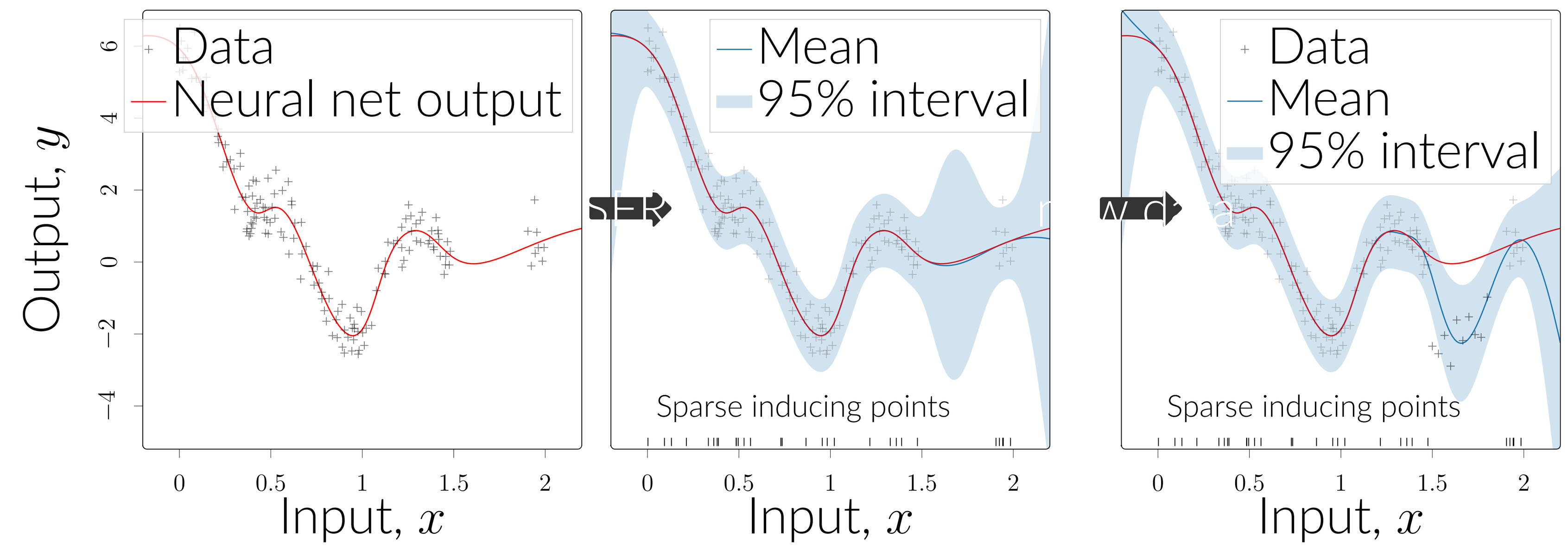
Aalto University

TL;DR

TODO: Rewrite Deep neural networks are known to lack uncertainty estimates, struggle to incorporate new data, and suffer from catastrophic forgetting. We present a method that mitigates these issues by converting neural networks from weight-space to a low-rank function-space representation, via the so-called dual parameters. In contrast to previous work, our sparse representation captures the joint distribution over the entire data set, rather than only over a subset. This offers a compact and principled way of capturing uncertainty and enables us to incorporate new data without retraining whilst retaining predictive performance. We provide proof-of-concept demonstrations with the proposed approach for quantifying uncertainty in supervised learning on UCI benchmark tasks.



Regression w/ 2-layers MLP. Left: Predictions from the trained neural network. Middle: Our approach summarizes all the training data with the help of a set of inducing points. The model captures the predictive mean and uncertainty, and (right) makes it possible to incorporate new data without retraining the model.



Regression w/ 2-layers MLP. Left: Predictions from the trained neural network. Middle: Our approach summarizes all the training data with the help of a set of inducing points. The model captures the predictive mean and uncertainty, and (right) makes it possible to incorporate new data without retraining the model.

Model & Methods

Sub-heading

- Item here
- Another item

Another sub-heading

- Item here
- Another item

Background

- Item here
- Another item

Results

model dataset	BNN	GLM	GP Subset (GP)	GP Subset (NN)	NN MAP	SFR (GP)	SFR (NN)
Glass	0.9562 ± 0.2649	0.8966 ± 0.1424	1.5108 ± 0.0367	1.1420 ± 0.0738	1.8350 ± 0.6147	1.3709 ± 0.0375	1.0922 ± 0.1363
Waveform	0.2693 ± 0.0310	0.2676 ± 0.0227	0.2702 ± 0.0192	0.2670 ± 0.0162	0.2670 ± 0.0162	0.2657 ± 0.0163	0.2670 ± 0.0162

UCI Results. Comparisons and ablations on UCI data with negative log predictive density

- Item here
- Another item

Discussion

- Item here
- Another item

References

[1] V. Adam, P. Chang, M. E. E. Khan, and A. Solin, "Dual parameterization of sparse variational Gaussian processes," in *Advances in Neural Information Processing Systems 34 (NeurIPS)*, pp. 11474–11486, Curran Associates, Inc., 2021.

[2]

[3] A. Immer, M. Korzepa, and M. Bauer, "Improving predictions of Bayesian neural nets via local linearization," in *Proceedings of The 24th International Conference on Artificial Intelligence and Statistics (AISTATS)*, vol. 130 of *Proceedings of Machine Learning Research*, pp. 703–711, PMLR, 2021.

aidan.scannell@aalto.fi