Robust uncertainty estimates with out-of-distribution pseudo-inputs training

Anonymous Author
Anonymous Institution

Abstract

Probabilistic models often use neural networks to control their predictive uncertainty. However, when making *out-of-distribution (OOD)* predictions, the often-uncontrollable extrapolation properties of neural networks yield poor uncertainty predictions. Such models then don't *know what they don't know*, which directly limits their robustness w.r.t unexpected inputs. To counter this, we propose to explicitly train the uncertainty predictor where we are not given data to make it reliable. As one cannot train without data, we provide mechanisms for generating *pseudo-inputs* in informative low-density regions of the input space, and show how to leverage these in a practical Bayesian framework that casts a prior distribution over the model uncertainty. With a holistic evaluation, we demonstrate that this yields robust and interpretable predictions of uncertainty while retaining state-of-the-art performance on diverse tasks such as regression and generative modelling.

1 Lipsum

B**'s Pure Strategy**	A**'s expected payoff**
I	$(2-10)x_1+1 = x_1+1$
II	$(1-0)x_1 = 0+x_1$
III	$(0-3)x_1+3 = -3x_1+3$
IV	$(-2-2)x_1+2 = -4x_1+2$

Lorem ipsum dolor sit amet, consectetuer adipiscing elit. Ut purus elit, vestibulum ut, placerat ac, adipiscing vitae, felis. Curabitur dictum gravida mauris. Nam arcu libero, nonummy eget, consectetuer id, vulputate a, magna. Donec vehicula augue eu neque. Pellentesque habitant morbi tristique senectus et netus et malesuada fames ac turpis egestas. Mauris ut leo. Cras viverra metus rhoncus sem. Nulla et lectus vestibulum urna fringilla ultrices. Phasellus eu tellus sit amet tortor gravida placerat. Integer sapien est, iaculis in, pretium quis, viverra ac, nunc. Praesent eget sem vel leo ultrices bibendum. Aenean faucibus. Morbi dolor nulla, malesuada eu, pulvinar at, mollis ac, nulla. Curabitur auctor semper nulla. Donec varius orci eget risus. Duis nibh mi, congue eu, accumsan eleifend, sagittis quis, diam. Duis eget orci sit amet orci dignissim rutrum.

Nam dui ligula, fringilla a, euismod sodales, sollicitudin vel, wisi. Morbi auctor lorem non justo. Nam lacus libero, pretium at, lobortis vitae, ultricies et, tellus. Donec aliquet, tortor sed accumsan bibendum, erat ligula aliquet magna, vitae ornare odio metus a mi. Morbi ac orci et nisl hendrerit mollis. Suspendisse ut massa. Cras nec ante. Pellentesque a nulla. Cum sociis natoque penatibus et magnis dis parturient montes, nascetur ridiculus mus. Aliquam tincidunt urna. Nulla ullamcorper vestibulum turpis. Pellentesque cursus luctus mauris.

Nulla malesuada porttitor diam. Donec felis erat, congue non, volutpat at, tincidunt tristique, libero. Vivamus viverra fermentum felis. Donec nonummy pellentesque ante. Phasellus adipiscing semper elit. Proin fermentum massa ac quam. Sed diam turpis, molestie vitae, placerat a, molestie nec, leo. Maecenas lacinia. Nam ipsum ligula, eleifend at, accumsan nec, suscipit a, ipsum. Morbi blandit ligula feugiat magna. Nunc eleifend consequat lorem. Sed lacinia nulla vitae enim. Pellentesque tincidunt purus vel magna. Integer non enim. Praesent euismod nunc eu purus. Donec bibendum quam in tellus. Nullam cursus pulvinar lectus. Donec et mi. Nam vulputate metus eu enim. Vestibulum pellentesque felis eu massa.

Quisque ullamcorper placerat ipsum. Cras nibh. Morbi vel justo vitae lacus tincidunt ultrices. Lorem ipsum dolor sit amet, consectetuer adipiscing elit. In hac habitasse platea dictumst. Integer tempus convallis

Preliminary work. Under review by AISTATS 2022. Do not distribute.

augue. Etiam facilisis. Nunc elementum fermentum wisi. Aenean placerat. Ut imperdiet, enim sed gravida sollicitudin, felis odio placerat quam, ac pulvinar elit purus eget enim. Nunc vitae tortor. Proin tempus nibh sit amet nisl. Vivamus quis tortor vitae risus porta vehicula.

Fusce mauris. Vestibulum luctus nibh at lectus. Sed bibendum, nulla a faucibus semper, leo velit ultricies tellus, ac venenatis arcu wisi vel nisl. Vestibulum diam. Aliquam pellentesque, augue quis sagittis posuere, turpis lacus congue quam, in hendrerit risus eros eget felis. Maecenas eget erat in sapien mattis porttitor. Vestibulum porttitor. Nulla facilisi. Sed a turpis eu lacus commodo facilisis. Morbi fringilla, wisi in dignissim interdum, justo lectus sagittis dui, et vehicula libero dui cursus dui. Mauris tempor ligula sed lacus. Duis cursus enim ut augue. Cras ac magna. Cras nulla. Nulla egestas. Curabitur a leo. Quisque egestas wisi eget nunc. Nam feugiat lacus vel est. Curabitur consectetuer.

Suspendisse vel felis. Ut lorem lorem, interdum eu, tincidunt sit amet, laoreet vitae, arcu. Aenean faucibus pede eu ante. Praesent enim elit, rutrum at, molestie non, nonummy vel, nisl. Ut lectus eros, malesuada sit amet, fermentum eu, sodales cursus, magna. Donec eu purus. Quisque vehicula, urna sed ultricies auctor, pede lorem egestas dui, et convallis elit erat sed nulla. Donec luctus. Curabitur et nunc. Aliquam dolor odio, commodo pretium, ultricies non, pharetra in, velit. Integer arcu est, nonummy in, fermentum faucibus, egestas vel, odio.

Sed commodo posuere pede. Mauris ut est. Ut quis purus. Sed ac odio. Sed vehicula hendrerit sem. Duis non odio. Morbi ut dui. Sed accumsan risus eget odio. In hac habitasse platea dictumst. Pellentesque non elit. Fusce sed justo eu urna porta tincidunt. Mauris felis odio, sollicitudin sed, volutpat a, ornare ac, erat. Morbi quis dolor. Donec pellentesque, erat ac sagittis semper, nunc dui lobortis purus, quis congue purus metus ultricies tellus. Proin et quam. Class aptent taciti sociosqu ad litora torquent per conubia nostra, per inceptos hymenaeos. Praesent sapien turpis, fermentum vel, eleifend faucibus, vehicula eu, lacus.

UCI benchmarks shifts included	d-VV	VV	VV (no prior)	Mean variance network	Skafte et al	Deep ensembles	Monte Carlo dropout	Noise contrastive priors	Bayes by backprop
$\mathcal{L}$	**21** / 18	3 / 3	0 / 0	n / a	n / a	n / a	n / a	n / a	n / a
$\log p(y\|x)$	2 / 7	5 / 9	0 / 0	0 / 0	4 / 9	0 / 0	0 / 0	2 / 8	**11** / 8
$\text{RMSE}[y, \mu(x)]$	2 / 2	1 / 4	2 / 5	1 / 5	2 / 3	5 / 12	3 / 16	0 / 0	**8** / 0
$\text{RMSE}[\text{Var}[y\|x], (y-\mu(x))^2]$	2 / 4	5 / 9	2 / 4	0 / 0	n / a	2 / 5	5 / 17	0 / 0	**8** / 7
$\text{RMSE}[y, \tilde{y}]$	3 / 4	5 / 7	4 / 2	n / a	n / a	n / a	n / a	1 / 2	**11** / 7
$\mathbb{E}[\text{KL}]$	**23** / 14	1 / 0	0 / 0	n / a	n / a	n / a	n / a	n / a	n / a

Table 1: UCI benchmarks - $\mathcal{L}$

	d-VV	VV (no PIG)	VV (Gaussian Noise)	VV (no prior)	f_mvn
uci_boston	-0.61±0.33	-0.72±0.38	-0.82±0.07	-64.28±29.34	na
uci_carbon	1.19±0.11	1.17±0.12	-0.47±0.01	-3913.5±580.03	na
uci_ccpp	-0.14±0.01	-0.16±0.04	-0.57±0.01	-10.9±1.06	na
uci_concrete	-0.44±0.13	-0.46±0.08	-0.65±0.03	-44.06±14.23	na
uci_energy	0.67±0.03	0.65±0.03	-0.51±0.01	-118.78±81.29	na
uci_kin8nm	-0.25±0.02	-0.28±0.04	-0.59±0.01	-16.76±1.53	na
uci_naval	0.52±0.16	0.12±0.4	-0.54±0.02	-9.57±4.57	na
uci_protein	-1.32±0.01	-1.34±0.01	-1.24±0.02	-14.31±2.59	na
uci_superconduct	-0.54±0.03	-0.56±0.01	-0.72±0.02	-269.08±58.88	na
uci_wine_red	-2.0±0.08	-2.36±0.19	-1.75±0.18	-37.86±31.48	na
uci_wine_white	-1.82±0.06	-2.01±0.1	-1.55±0.17	-869.26±1718.81	na
uci_yacht	-17.6±0.43	1.04±0.1	-0.51±0.01	-551.12±1060.39	na

	f_ens	f_mcd	john	BBB+NCP	BBB	Det
uci_boston	na	na	na	na	na	na
uci_carbon	na	na	na	na	na	na
uci_ccpp	na	na	na	na	na	na
uci_concrete	na	na	na	na	na	na
uci_energy	na	na	na	na	na	na
uci_kin8nm	na	na	na	na	na	na
uci_naval	na	na	na	na	na	na
uci_protein	na	na	na	na	na	na
uci_superconduct	na	na	na	na	na	na
uci_wine_red	na	na	na	na	na	na
uci_wine_white	na	na	na	na	na	na
uci_yacht	na	na	na	na	na	na

Table 2: UCI benchmarks - $\log p(y|x)$

	d-VV	VV (no PIG)	VV (Gaussian Noise)	VV (no prior)	f_mvn
uci_boston	-0.43±0.35	-0.42±0.39	-0.58±0.07	-3.34±1.39	-0.76±0.07
uci_carbon	1.45±0.13	1.45±0.11	-0.29±0.01	0.98±3.08	-3.78±0.05
uci_ccpp	-0.07±0.01	-0.03±0.03	-0.45±0.01	0.05±0.06	-0.58±0.14
uci_concrete	-0.29±0.15	-0.25±0.1	-0.44±0.03	-0.83±0.5	-0.68±0.09
uci_energy	0.87±0.04	0.89±0.03	-0.25±0.01	0.47±0.2	-1.22±0.11
uci_kin8nm	-0.17±0.02	-0.15±0.04	-0.5±0.01	-0.36±0.07	-0.61±0.06
uci_naval	0.71±0.19	0.52±0.2	-0.26±0.03	-0.13±0.32	-2.26±0.08
uci_protein	-1.16±0.01	-1.12±0.01	-1.13±0.02	-1.42±0.38	-1.13±0.05
uci_superconduct	-0.38±0.02	-0.35±0.02	-0.52±0.02	-1.73±1.71	-0.66±0.04
uci_wine_red	-1.91±0.07	-2.13±0.21	-1.67±0.18	-7.77±6.39	-2560.95±5395.69
uci_wine_white	-1.72±0.06	-1.75±0.1	-1.49±0.17	-305.7±549.73	-27.69±48.8
uci_yacht	0.9±0.02	1.33±0.11	-0.24±0.0	0.63±0.59	-0.59±0.11

	f_ens	f_mcd	john	BBB+NCP	BBB	Det
uci_boston	-0.68±0.04	-0.81±0.51	-0.18±0.19	-1.39±0.33	-248.43±163.55	-96.5±187.47
uci_carbon	-3.71±0.04	0.29±1.08	1.13±0.51	na	na	na
uci_ccpp	-0.61±0.05	-3.36±0.56	-0.18±0.12	0.21±0.04	4.06±0.69	3.17±1.46
uci_concrete	-0.65±0.04	-0.9±0.33	-0.4±0.15	0.38±0.04	3.84±0.66	2.9±2.1
uci_energy	-1.17±0.04	0.36±0.26	0.28±0.37	na	na	na
uci_kin8nm	-0.65±0.03	-0.63±0.05	-0.62±0.12	-0.68±0.08	-0.16±0.03	-0.2±0.03
uci_naval	-2.26±0.06	-0.2±0.74	-2.67±0.22	na	na	na
uci_protein	-1.05±0.01	-7.41±0.27	-1.54±0.74	-1.02±0.01	-0.96±0.02	-0.96±0.03
uci_superconduct	-0.68±0.03	-1.72±0.25	-0.96±0.18	-0.2±0.19	-0.04±0.06	-0.03±0.14
uci_wine_red	-1.24±0.08	-4.24±0.91	-1.14±0.04	0.16±0.04	3.76±0.39	3.8±0.42
uci_wine_white	-1.16±0.08	-5.86±1.08	-1.4±0.58	0.29±0.06	3.76±0.82	4.05±0.37
uci_yacht	-0.58±0.04	0.33±0.69	0.4±0.14	0.63±0.1	1.57±0.6	0.41±1.53

Table 3: UCI benchmarks - RMSE $[y, \mu(x)]$

	d-VV	VV (no PIG)	VV (Gaussian Noise)	VV (no prior)	f_mvn
uci_boston	0.33±0.09	0.33±0.08	0.38±0.04	0.38±0.09	0.35±0.06
uci_carbon	0.03±0.02	0.03±0.02	0.03±0.01	0.03±0.02	0.75±0.01
uci_ccpp	0.23±0.01	0.23±0.01	0.23±0.01	0.23±0.01	0.23±0.01
uci_concrete	0.29±0.04	0.29±0.04	0.27±0.03	0.33±0.04	0.29±0.01
uci_energy	0.08±0.01	0.08±0.01	0.07±0.02	0.3±0.03	0.13±0.01
uci_kin8nm	0.26±0.01	0.26±0.01	0.26±0.01	0.28±0.01	0.27±0.01
uci_naval	0.09±0.06	0.1±0.03	0.1±0.05	0.33±0.07	0.72±0.01
uci_protein	0.71±0.01	0.71±0.0	0.7±0.01	0.75±0.01	0.71±0.01
uci_superconduct	0.35±0.01	0.35±0.01	0.34±0.01	0.4±0.02	0.35±0.01
uci_wine_red	0.89±0.01	0.9±0.04	0.84±0.06	0.77±0.07	1.13±0.15
uci_wine_white	0.9±0.03	0.88±0.02	0.83±0.06	0.82±0.06	0.85±0.05
uci_yacht	0.05±0.02	0.04±0.02	0.04±0.01	0.82±0.12	0.05±0.01

	f_ens	f_mcd	john	BBB+NCP	BBB	Det
uci_boston	0.29±0.05	0.33±0.05	0.3±0.07	0.47±0.06	0.5±0.04	0.52±0.04
uci_carbon	0.75±0.01	0.08±0.0	0.09±0.08	na	na	na
uci_ccpp	0.23±0.01	0.24±0.01	0.27±0.03	0.07±0.03	0.0±0.0	0.01±0.0
uci_concrete	0.27±0.01	0.28±0.02	0.36±0.07	0.08±0.02	0.0±0.0	0.01±0.0
uci_energy	0.13±0.01	0.13±0.02	0.22±0.08	na	na	na
uci_kin8nm	0.26±0.01	0.33±0.01	0.44±0.06	0.4±0.07	0.29±0.0	0.3±0.01
uci_naval	0.72±0.01	0.2±0.07	0.86±0.08	na	na	na
uci_protein	0.69±0.01	0.7±0.01	1.12±0.73	0.76±0.02	0.73±0.01	0.73±0.01
uci_superconduct	0.32±0.01	0.33±0.01	0.67±0.22	0.44±0.03	0.41±0.01	0.41±0.01
uci_wine_red	0.84±0.06	0.77±0.05	0.76±0.02	0.1±0.04	0.01±0.0	0.01±0.0
uci_wine_white	0.77±0.06	0.79±0.05	0.93±0.39	0.08±0.04	0.01±0.0	0.01±0.0
uci_yacht	0.05±0.01	0.11±0.04	0.09±0.06	0.09±0.02	0.16±0.03	0.12±0.05