Supplementary Material: Leveraging Probabilistic Circuits for Nonparametric Multi-Output Regression

Zhongjie Yu[1] **Mingye Zhu**[2] **Martin Trapp**[3] **Arseny Skryagin**[1] **Kristian Kersting**[1,4]

[1]Department of Computer Science, TU Darmstadt, Darmstadt, Germany
[2]Department of Automation Engineering, Nanjing University of Aeronautics and Astronautics, Nanjing, China
[3]Department of Computer Science, Aalto University, Espoo, Finland
[4]Centre for Cognitive Science, TU Darmstadt, and Hessian Center for AI (hessian.AI)

A MOMOGP STRUCTURE CONSTRUCTION

Table 1 compares the results obtained from MOMoGPs constructed either using conditional independence tests or using random splitting. We construct MOMoGPs using conditional independence tests for the splitting of the output space, we employed the randomized conditional correlation test Strobl et al. [2019]. In all experiments, we used a p-value of 0.5. We see that the use of a conditional independence test for the structure construction results in an overall improvement of the performance of MOMoGPs with respect to the RMSE and the MAE.

Data Set		Random	CI
Parkinsons	RMSE	**0.775**	**0.775**
	MAE	**0.605**	**0.605**
	NLPD	**2.208**	**2.208**
scm20d	RMSE	0.820	**0.816**
	MAE	0.630	**0.627**
	NLPD	**11.416**	11.470
WindTurbine	RMSE	0.143	**0.137**
	MAE	0.073	**0.072**
	NLPD	**-7.467**	-7.478
Energy	RMSE	0.556	**0.516**
	MAE	0.398	**0.358**
	NLPD	**8.610**	9.402
usFlight	RMSE	0.934	**0.914**
	MAE	0.505	**0.485**
	NLPD	2.091	**2.075**

Table 1: Comparison of results obtained for MOMoGPs constructed using conditional independence tests (CI) or using random splitting.

B MULTI-OUTPUT REGRESSION BENCHMARK RESULTS

We train and test GP, DSMGP, sumGP, and MOMoGP on all data sets (except for usFlight) five times with different random seeds. For usFlight, the above models are trained twice. Table 2 shows both average and standard deviation of the multiple runs. MOGP and MOSVGP have only been trained once for all data sets, thus, their results are not compared in Table 2.

References

Eric V Strobl, Kun Zhang, and Shyam Visweswaran. Approximate kernel-based conditional independence tests for fast non-parametric causal discovery. *Journal of Causal Inference*, 7(1), 2019.

Data Set		GP		DSMGP		sumGP		**MOMoGP**	
		mean	std	mean	std	mean	std	mean	std
Parkinsons	RMSE	0.783	4×10^{-4}	**0.774**	1×10^{-4}	0.784	3×10^{-4}	0.775↓	5×10^{-4}
	MAE	0.610	3×10^{-4}	**0.604**	9×10^{-4}	0.610	4×10^{-4}	0.605↓	9×10^{-4}
	NLPD	2.389	1.0×10^{-3}	2.319	3.0×10^{-3}	2.388	1.0×10^{-3}	**2.208**↑	3.0×10^{-3}
scm20d	RMSE	0.332	1.80×10^{-2}	0.323	2.7×10^{-4}	**0.322**	1.6×10^{-4}	0.324↓	1.1×10^{-3}
	MAE	0.195	1.59×10^{-2}	0.188	4.1×10^{-4}	**0.187**	3.2×10^{-4}	**0.187**↑	1.2×10^{-3}
	NLPD	0.543	1.90×10^{-2}	1.341	5.15×10^{-2}	8.097	1.77×10^{-1}	**-0.201**↑	1.90×10^{-1}
WindTurbine	RMSE	**0.133**	0	0.143	1×10^{-4}	**0.133**	2×10^{-4}	0.143	0
	MAE	0.074	3×10^{-4}	**0.073**	0	0.074	2×10^{-4}	**0.073**	0
	NLPD	-2.649	4.34×10^{-2}	**-8.749**	2.3×10^{-3}	-2.627	2.89×10^{-2}	-7.467↓	3.0×10^{-3}
Energy	RMSE	NA	NA	**0.547**	6.0×10^{-3}	NA	NA	0.556↓	6.0×10^{-3}
	MAE	NA	NA	0.400	2.0×10^{-3}	NA	NA	**0.398**↑	3.0×10^{-3}
	NLPD	NA	NA	11.745	1.50×10^{-2}	NA	NA	**8.610**↑	1.00×10^{-2}
usFlight	RMSE	NA	NA	**0.927**	2×10^{-4}	NA	NA	0.934↓	3×10^{-4}
	MAE	NA	NA	**0.492**	3×10^{-4}	NA	NA	0.505↓	3×10^{-4}
	NLPD	NA	NA	2.178	1×10^{-4}	NA	NA	**2.091**↑	3.57×10^{-2}

Table 2: Mean and standard deviation of Root Mean Square Error (RMSE), Mean Absolute Error (MAE), and Negative Log Predictive Density (NLPD) of state-of-the-art approaches and MOMoGP (our work) on benchmark data sets. Smaller values are better. Best result is indicated in **bold** and comparison of MOMoGP to DSMGP is indicated using arrows ↑ / ↓.