Global Performance Heatmap Across All Configurations

Tasks & Constraints	DYNAMICAL_TRANSFORMER	ESN	GRU	LSTM	MAMBA	TRANSFORMER_DECODER	TRANSFORMER_ENCDEC	XLSTM
adding_problem (MD)	0.00	0.69	0.00	0.00	0.00	0.00	0.00	0.00
adding_problem (SM)	0.00	0.00	0.00	0.02	0.01	0.00	0.00	0.00
associative_recall (MD)	0.88	0.86	0.91	0.91	0.91	0.85	0.88	0.90
associative_recall (SM)	0.87	0.86	0.86	0.86	0.84	0.87	0.87	0.87
bracket_matching (MD)	0.02	0.13	0.01	0.00	0.02	0.04	0.01	0.00
bracket_matching (SM)	0.01	0.13	0.18	0.21	0.29	0.19	0.27	0.06
chaotic_forecasting (MD)	0.00	0.00	0.00	0.00	0.00	0.08	0.07	0.00
chaotic_forecasting (SM)	0.03	0.00	0.00	0.01	0.00	0.04	0.04	0.01
continuous_pattern_completion (MD)	0.01	0.06	0.01	0.01	0.02	0.01	0.00	0.01
continuous_pattern_completion (SM)	0.01	0.02	0.01	0.01	0.01	0.03	0.06	0.01
continuous_postcasting (MD)	0.00	0.00	0.00	0.00	0.02	0.00	0.00	0.00
continuous_postcasting (SM)	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
cross_situation (MD)	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
cross_situation (SM)	0.02	0.00	0.03	0.05	0.05	0.03	0.04	0.03
discrete_pattern_completion (MD)	0.07	0.09	0.18	0.11	0.14	0.07	0.00	0.18
discrete_pattern_completion (SM)	0.04	0.06	0.06	0.05	0.06	0.07	0.36	0.06
discrete_postcasting (MD)	0.00	0.00	0.00	0.00	0.08	0.00	0.00	0.00
discrete_postcasting (SM)	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
induction_heads (MD)	0.00	0.73	0.47	0.77	0.12	0.00	0.00	0.25
induction_heads (SM)	0.00	0.01	0.35	0.37	0.48	0.00	0.00	0.39
selective_copy (MD)	0.00	0.70	0.23	0.86	0.00	0.00	0.01	0.23
selective_copy (SM)	0.00	0.41	0.22	0.52	0.21	0.05	0.17	0.34
simple_copy (MD)	0.00	0.70	0.64	0.84	0.13	0.00	0.00	0.51
simple_copy (SM)	0.00	0.00	0.49	0.57	0.42	0.00	0.00	0.42
sinus_forecasting (MD)	0.00	0.00	0.01	0.00	0.00	0.07	0.07	0.01
sinus_forecasting (SM)	0.01	0.00	0.01	0.01	0.01	0.02	0.02	0.01
sorting_problem (MD)	0.02	0.71	0.17	0.56	0.00	0.00	0.00	0.06
sorting_problem (SM)	0.13	0.43	0.42	0.49	0.51	0.00	0.00	0.05

Architectures

Mean Metric (Lower is usually better)

0.0 0.2 0.4 0.6 0.8