auto-sktime: Automated Time Series Forecasting Supplementary Materials

1 Raw Experiment Results

Table 1: Raw experiment results for all evaluated frameworks and used datasets using MASE.

Dataset	Auto-PyTorch		auto-sktime		AutoGluon		AutoTS		DeepAR		ETS		HyperTS		pmdarima		PyAF		TFT	
	MASE	Duration	MASE	Duration	MASE	Duration	MASE	Duration	MASE	Duration	MASE	Duration	MASE	Duration	MASE	Duration	MASE	Duration	MASE	Duration
AEDemand	–	310	0.18	320	1.09	100	1.11	308	1.04	6	0.94	0	1.00	305	0.00	1	1.15	21	1.07	197
AEP_hourly	5.59	317	3.24	341	–	0	2.97	1256	–	0	–	0	3.05	563	–	0	–	0	–	0
AirPassengers	0.78	321	0.93	323	0.59	99	1.54	310	3.68	3	3.16	0	0.82	312	1.86	1	3.00	4	2.79	4
DAYTON_hourly	5.81	370	3.91	327	–	0	3.20	1198	–	0	–	0	3.13	571	–	0	–	0	–	0
DEOK_hourly	5.38	313	2.03	322	–	0	2.98	798	–	0	–	0	2.71	473	–	0	–	0	–	0
DM	3.39	321	2.38	321	2.92	298	2.99	215	–	27	3.39	0	1.94	354	3.41	7	3.39	55	3.85	300
DOM_hourly	4.05	312	1.81	319	–	0	4.04	1045	–	0	–	0	1.98	520	–	0	–	0	–	0
DUQ_hourly	5.21	303	0.81	329	–	0	2.13	1286	–	0	–	0	2.05	525	–	0	–	0	–	0
EKPC_hourly	4.25	302	1.93	430	–	0	2.27	408	–	0	–	0	2.55	379	–	0	–	0	–	0
FE_hourly	7.06	303	3.19	327	–	0	3.89	790	–	0	–	0	3.34	420	–	0	–	0	–	0
M3_Monthly	–	6	–	328	–	944	–	359	–	4	–	49	–	8	–	1966	–	17	–	4
M3_Quarterly	–	3	–	326	–	1623	–	359	–	2	–	27	–	3	–	1063	–	34	–	1
M3_Yearly	–	3	9.78	336	–	3	17.57	2601	–	1	5.96	24	–	3	–	358	–	9	–	1
NI_hourly	–	0	2.84	499	–	0	3.29	741	–	0	–	0	4.99	397	–	0	–	0	–	0
NN3-Final-Dataset	–	0	1.01	336	0.94	10029	–	1599	1.15	18	1.00	3	0.81	327	0.89	136	–	411	–	0
NN5-Final-Dataset	–	0	0.44	345	–	0	–	1599	–	0	–	6	–	327	–	0	–	0	–	0
PJME_hourly	5.31	304	1.62	334	–	0	4.42	369	–	0	–	1	2.97	559	–	0	–	0	–	0
PJMW_hourly	4.50	304	0.87	339	–	0	1.93	400	–	0	–	1	2.66	539	–	0	–	0	–	0
PJM_Load_hourly	–	595	1.68	389	–	0	1.66	365	–	0	–	0	1.98	345	–	0	–	0	–	0
Twitter_volume_AAPL	0.18	309	0.33	333	0.35	3718	0.38	403	0.99	300	0.24	0	0.75	328	0.60	333	0.24	11	0.27	300
Twitter_volume_AMZN	0.88	308	0.58	335	1.03	4126	1.07	390	1.01	300	1.06	0	1.24	339	1.03	48	0.64	13	0.41	300
Twitter_volume_GOOG	1.67	311	1.67	330	1.68	3801	2.05	412	1.82	300	1.63	0	1.80	340	1.71	174	7.31	13	2.83	300
a10	1.30	312	1.40	315	1.27	117	2.20	317	4.50	6	2.76	0	1.26	312	2.69	1	1.69	3	3.12	30
ausbeer	0.26	316	0.20	342	0.22	102	0.26	310	0.61	6	0.56	0	0.19	310	0.21	5	0.18	6	–	0
auscafe	1.59	320	1.08	322	0.98	98	1.10	338	–	15	1.54	0	0.74	317	1.49	4	1.55	6	1.57	73
departures	3.23	322	1.28	336	0.58	518	0.83	422	2.29	182	2.57	0	2.35	1048	2.23	81	0.98	34	2.39	76
ec2_cpu_24ae8d	0.03	325	0.01	360	0.03	654	0.02	335	–	85	0.16	0	0.14	306	0.16	8	0.02	9	0.03	300
ec2_cpu_53ea38	0.52	326	0.28	339	0.31	650	0.29	330	14.15	85	0.63	0	0.31	326	0.59	84	0.31	10	0.55	300
ec2_cpu_5f5533	0.19	318	0.18	334	0.13	651	0.25	326	0.17	300	0.17	0	2.17	309	0.36	27	0.13	9	0.19	300
ec2_cpu_77c1ca	0.00	321	0.00	336	0.00	638	0.07	312	–	159	0.00	0	0.00	305	1.21	122	0.00	9	0.02	132
ec2_cpu_825cc2	1.53	302	1.29	344	–	0	0.89	350	–	0	–	0	0.54	308	–	0	–	0	–	0
ec2_cpu_ac20cd	–	0	6.32	326	–	0	0.16	340	–	0	–	0	1.48	307	–	0	–	0	–	0
ec2_cpu_c6585a	0.35	320	0.42	338	0.68	656	0.40	314	88.68	75	0.61	0	0.50	305	0.60	16	0.34	8	0.35	230
ec2_cpu_fe7f93	0.09	318	0.10	344	0.11	577	0.18	323	0.21	300	0.12	0	0.34	310	0.48	61	0.15	10	0.10	300
ec2_network_in_257a54	0.02	302	0.01	331	–	0	0.03	353	–	0	–	0	0.01	309	–	0	–	0	–	0
ec2_network_in_5abac7	0.00	316	0.00	331	0.00	678	0.00	313	0.00	300	0.27	0	0.00	304	0.50	9	0.00	10	0.52	108
elb_request_count_8c0756	0.61	302	2.42	323	–	0	1.39	320	–	0	–	0	0.45	305	–	0	–	0	–	0
elecdaily	0.66	318	1.10	320	1.42	113	1.67	407	1.97	143	0.85	0	1.99	307	0.80	7	2.55	13	1.78	135
elecdemand	2.93	319	1.78	333	0.70	876	1.46	524	–	300	1.17	0	5.16	309	1.28	690	2.22	92	2.12	300
exchange-2_cpc_results	–	0	0.93	418	–	0	1.57	339	–	0	–	0	1.33	306	–	0	–	0	–	0
exchange-2_cpm_results	–	0	1.12	402	–	0	1.60	335	–	0	–	0	1.33	307	–	0	–	0	–	0
exchange-4_cpc_results	1.75	302	1.43	384	–	0	1.43	397	–	0	–	0	1.28	312	–	0	–	0	–	0
exchange-4_cpm_results	1.90	302	1.30	425	–	0	1.32	330	–	0	–	0	1.09	307	–	0	–	0	–	0
exchange_rate	0.94	320	1.32	322	1.14	1329	1.87	979	–	300	0.94	0	0.90	379	0.93	203	0.94	232	1.40	300
gasoline	2.01	325	1.40	438	2.04	103	1.83	324	–	37	1.66	0	1.76	307	1.66	4	1.69	18	2.92	300
h02	0.48	328	0.84	308	0.81	99	0.84	307	1.78	4	2.01	0	0.80	310	1.98	0	0.54	6	1.69	10
hts_dataset	–	0	0.46	323	0.39	797	0.65	304	0.64	2	0.63	0	0.39	672	0.42	19	0.43	29	–	0
hyndsight	0.56	320	0.54	325	0.62	103	0.72	311	1.29	46	1.40	0	0.76	306	0.61	18	0.47	8	–	0
infant_gts	–	0	0.62	320	0.71	1550	2.13	934	6.56	3	1.50	0	0.66	343	1.07	13	1.20	66	–	0
insurance	1.55	321	3.69	289	3.15	105	83.11	313	–	0	4.30	0	4.43	306	1.86	1	2.78	6	–	0
melsyd	–	0	0.78	316	–	0	1.32	175	–	0	–	0	1.53	378	–	0	–	0	–	0
ozone-la	0.49	327	0.64	315	0.73	102	0.97	307	0.99	4	1.74	0	0.73	305	1.79	0	0.59	6	0.99	11
petrol	1.86	322	1.04	318	1.60	400	1.73	263	1.01	42	1.05	0	0.98	458	0.87	22	1.34	26	1.36	109
prison	–	0	2.45	328	2.63	111	–	476	11.18	3	2.84	1	3.84	366	2.53	27	3.88	116	–	0
prisonLF	–	0	2.94	336	2.92	1118	9.99	415	–	0	2.98	1	–	0	2.79	29	–	6	–	0
qcement	1.49	327	0.53	316	1.69	6	1.81	330	–	6	1.97	0	1.11	313	1.03	15	2.55	6	1.22	28
qgas	0.30	317	0.24	326	0.31	2	0.22	348	1.41	3	1.26	0	0.76	308	0.23	4	0.26	6	1.33	4
rds_cpu_cc0c53	1.39	302	1.06	343	–	0	3.32	341	–	0	–	0	1.09	305	–	0	–	0	–	0
rds_cpu_e47b3b	0.48	317	0.64	328	0.43	406	0.41	348	0.47	300	0.62	0	0.69	313	0.68	28	1.06	10	0.49	299
rossman	–	0	–	0	–	0	–	0	–	0	–	0	–	0	–	0	–	0	–	0
sunspots	0.89	306	0.82	318	0.69	4	1.77	320	2.11	163	2.68	0	0.97	306	0.55	8	–	0	2.12	95
uschange	0.49	306	0.22	310	0.38	9	0.63	303	0.68	3	0.36	0	0.49	303	0.16	3	0.41	10	0.65	11
usmelec	0.55	314	0.36	346	1.12	5	0.55	314	1.71	47	1.37	0	0.34	305	1.36	14	0.72	7	1.39	29
visnights	0.72	312	0.63	331	0.67	119	1.03	814	–	24	0.77	0	0.43	409	0.64	39	0.58	114	1.04	7

Table 2: Raw experiment results for all evaluated frameworks and used datasets using RMSE. Predictions with consistent failures are marked with –.

Dataset	Auto-PyTorch		auto-sktime		AutoGluon		AutoTS		DeepAR		ETS		HyperTS		pmdarima		PyAF		TFT	
	RMSE	Duration	RMSE	Duration	RMSE	Duration	RMSE	Duration	RMSE	Duration	RMSE	Duration	RMSE	Duration	RMSE	Duration	RMSE	Duration	RMSE	Duration
AEDemand	–	310	0.48	320	2.51	100	2.45	308	2.22	6	2.26	0	2.41	305	0.00	1	2.28	21	2.50	197
AEP_hourly	2807.39	317	1758.63	341	–	0	1526.48	1256	–	0	–	0	1517.44	563	–	0	–	0	–	0
AirPassengers	24.15	321	26.08	323	18.05	99	44.83	310	111.65	3	102.98	0	22.84	312	53.29	1	79.36	4	91.43	4
DAYTON_hourly	442.30	370	317.72	327	–	0	242.41	1198	–	0	–	0	247.90	571	–	0	–	0	–	0
DEOK_hourly	607.46	313	255.92	322	–	0	335.39	798	–	0	–	0	301.04	473	–	0	–	0	–	0
DM	0.10	321	0.07	321	0.08	298	0.09	215	653.97	27	0.10	0	0.06	354	0.10	7	0.10	55	0.10	300
DOM_hourly	1877.51	312	1038.07	319	–	0	2203.63	1045	–	0	–	0	926.73	520	–	0	–	0	–	0
DUQ_hourly	293.03	303	50.01	329	–	0	120.02	1286	–	0	–	0	124.57	525	–	0	–	0	–	0
EKPC_hourly	288.85	302	138.17	430	–	0	158.11	408	–	0	–	0	178.45	379	–	0	–	0	–	0
FE_hourly	1860.53	303	863.76	327	–	0	1066.99	790	–	0	–	0	838.23	420	–	0	–	0	–	0
M3_Monthly	–	6	–	328	–	944	–	359	–	4	–	49	–	8	–	1966	–	17	–	4
M3_Quarterly	–	3	–	326	–	1623	–	359	–	2	–	27	–	3	–	1063	–	34	–	1
M3_Yearly	–	3	2897.38	336	–	3	–	2601	–	1	1919.97	24	–	3	–	358	–	9	–	1
NI_hourly	–	0	1166.77	499	–	0	–	741	–	0	–	0	2028.23	397	–	0	–	0	–	0
NN3-Final-Dataset	–	0	1024.01	336	981.39	10029	–	1599	1119.09	18	1035.86	3	859.34	327	917.25	136	–	411	–	0
NN5-Final-Dataset	–	0	4.18	345	–	0	–	1599	–	0	–	6	10.08	327	–	0	–	0	–	0
PJME_hourly	6571.53	304	2010.92	334	–	0	6083.82	369	–	0	–	1	4199.94	559	–	0	–	0	–	0
PJMW_hourly	863.96	304	179.42	339	–	0	399.87	400	–	0	–	1	531.26	539	–	0	–	0	–	0
PJM_Load_hourly	–	595	2185.96	389	–	0	2078.34	365	–	0	–	0	2512.01	345	–	0	–	0	–	0
Twitter_vol_AAPL	8.27	309	13.38	333	14.27	3718	15.16	403	33.12	300	10.92	0	25.29	328	22.30	333	11.04	11	9.71	300
Twitter_vol_AMZN	11.53	308	7.96	335	13.31	4126	14.50	390	13.93	300	13.75	0	15.95	339	13.31	48	8.79	13	5.72	300
Twitter_vol_GOOG	15.44	311	15.60	330	15.43	3801	19.54	412	16.66	300	15.18	0	16.35	340	15.84	174	58.75	13	25.12	300
a10	2.11	312	2.09	315	1.92	117	3.33	317	6.36	6	4.14	0	1.83	312	4.02	1	2.43	3	4.51	30
ausbeer	15.30	316	13.18	342	16.34	102	17.63	310	41.16	6	40.99	0	13.96	310	13.15	5	11.25	6	–	0
auscafe	0.16	320	0.11	322	0.10	98	0.10	338	–	15	0.18	0	0.07	317	0.17	4	0.15	6	0.18	73
departures	72.22	322	28.45	336	14.90	518	20.12	422	54.45	182	61.96	0	54.07	1048	52.34	81	22.13	34	56.82	76
ec2_cpu_24ae8d	0.00	325	0.00	360	0.00	654	0.00	335	24.24	85	0.01	0	0.01	306	0.01	8	0.00	9	0.00	300
ec2_cpu_53ea38	0.07	326	0.04	339	0.04	650	0.04	330	–	85	0.09	0	0.04	326	0.08	84	0.04	10	0.09	300
ec2_cpu_5f5533	1.07	318	0.85	334	0.74	651	1.28	326	0.98	300	0.97	0	8.53	309	1.47	27	0.58	9	1.07	300
ec2_cpu_77c1ca	0.01	321	0.00	336	0.00	638	0.70	312	–	159	0.00	0	0.01	305	7.16	122	0.00	9	0.10	132
ec2_cpu_825cc2	3.07	302	2.64	344	–	0	2.05	350	–	0	–	0	1.19	308	–	0	–	0	–	0
ec2_cpu_ac20cd	–	0	14.31	326	–	0	0.39	340	–	0	–	0	3.39	307	–	0	–	0	–	0
ec2_cpu_c6585a	0.03	320	0.03	338	0.04	656	0.03	314	4.06	75	0.03	0	0.03	305	0.03	16	0.03	8	0.03	230
ec2_cpu_fe7f93	0.37	318	0.38	344	0.43	577	0.65	323	0.66	300	0.47	0	0.98	310	1.45	61	0.50	10	0.45	300
ec2_net_in_257a54	12114.54	302	6700.46	331	–	0	31374.28	353	–	0	–	0	5597.40	309	–	0	–	0	–	0
ec2_net_in_5abac7	18.05	316	162.29	331	17.75	678	53.30	313	69.77	300	–	0	28.94	304	–	9	33.42	10	–	108
elb_req_count_8c0756	37.31	302	–	323	–	0	89.88	320	–	0	–	0	29.69	305	–	0	–	0	–	0
elecdaily	12.73	318	19.86	320	23.29	113	28.51	407	32.10	143	16.25	0	34.01	307	17.94	7	43.85	13	29.82	135
elecdemand	0.38	319	0.24	333	0.10	876	0.23	524	–	300	0.17	0	0.62	309	0.21	690	0.31	92	0.28	300
exchange-2_cpc_res	–	0	0.02	418	–	0	0.03	339	–	0	–	0	0.02	306	–	0	–	0	–	0
exchange-2_cpm_res	–	0	0.10	402	–	0	0.14	335	–	0	–	0	0.12	307	–	0	–	0	–	0
exchange-4_cpc_res	0.06	302	0.05	384	–	0	0.05	397	–	0	–	0	0.05	312	–	0	–	0	–	0
exchange-4_cpm_res	0.40	302	0.28	425	–	0	0.28	330	–	0	–	0	0.23	307	–	0	–	0	–	0
exchange_rate	0.00	320	0.00	322	0.00	1329	0.01	979	298.26	300	0.00	0	0.00	379	0.00	203	0.00	232	0.00	300
gasoline	0.63	325	0.46	438	0.67	103	0.60	324	–	37	0.56	0	0.56	307	0.56	4	0.57	18	0.84	300
h02	0.07	328	0.10	308	0.10	99	0.10	307	0.19	4	0.24	0	0.09	310	0.23	0	0.07	6	0.19	10
hts_dataset	–	0	1000.83	323	818.11	797	1421.04	304	1350.41	2	1361.60	0	829.78	672	941.17	19	941.53	29	–	0
hyndsight	193.28	320	202.10	325	225.78	103	257.50	311	404.28	46	442.27	0	249.45	306	207.86	18	181.17	8	–	0
infant_gts	–	0	15.99	320	18.47	1550	–	934	145.42	3	35.54	0	16.95	343	26.05	13	29.31	66	–	0
insurance	2.43	321	5.73	289	4.74	105	168.43	313	–	0	6.27	0	6.64	306	2.71	1	4.45	6	–	0
melsyd	–	0	0.66	316	–	0	1.08	175	–	0	–	0	1.11	378	–	0	–	0	–	0
ozone-la	0.50	327	0.64	315	0.74	102	0.96	307	1.02	4	1.72	0	0.76	305	1.75	0	0.58	6	0.92	11
petrol	4.68	322	2.75	318	3.98	400	4.57	263	2.71	42	2.61	0	2.80	458	2.34	22	3.45	26	3.60	109
prison	–	0	61.85	328	66.80	111	–	476	267.26	3	72.57	1	98.48	366	64.42	27	97.15	116	–	0
prisonLF	–	0	161.69	336	106.92	1118	–	415	–	0	119.49	1	–	0	104.25	29	–	6	–	0
qcement	0.19	327	0.07	316	0.21	6	0.27	330	–	6	0.28	0	0.14	313	0.14	15	0.33	6	0.16	28
qgas	4.76	317	4.25	326	5.03	2	4.31	348	27.13	3	20.48	0	12.19	308	4.21	4	4.72	6	24.51	4
rds_cpu_cc0c53	0.86	302	0.62	343	–	0	2.15	341	–	0	–	0	0.64	305	–	0	–	0	–	0
rds_cpu_e47b3b	0.45	317	0.63	328	0.43	406	0.46	348	0.56	300	0.62	0	0.69	313	0.64	28	0.94	10	0.55	299
rossman	–	0	–	0	–	0	–	0	–	0	–	0	–	0	–	0	–	0	–	0
sunspots	17.83	306	19.77	318	18.03	4	41.95	320	44.69	163	62.53	0	22.48	306	15.36	8	–	0	43.32	95
uschange	0.31	306	0.16	310	0.25	9	0.41	303	0.44	3	0.24	0	0.30	303	0.12	3	0.33	10	0.43	11
usmelec	11.78	314	8.53	346	26.16	5	12.70	314	37.58	47	35.01	0	9.36	305	29.38	14	15.20	7	35.51	29
visnights	0.72	312	0.61	331	0.66	119	148.88	814	247.60	24	0.80	0	0.41	409	0.64	39	0.54	114	1.01	7

Table 3: Raw experiment results for all evaluated frameworks and used datasets using sMAPE. Predictions with consistent failures are marked with –.

Dataset	Auto-PyTorch		auto-sktime		AutoGluon		AutoTS		DeepAR		ETS		HyperTS		pmdarima		PyAF		TFT	
	sMAPE	Duration	sMAPE	Duration	sMAPE	Duration	sMAPE	Duration	sMAPE	Duration	sMAPE	Duration	sMAPE	Duration	sMAPE	Duration	sMAPE	Duration	sMAPE	Duration
AEDemand	–	310	0.00	320	0.02	100	0.02	308	0.02	6	0.02	0	0.02	305	0.00	1	0.02	21	0.02	197
AEP_hourly	0.16	317	0.09	341	–	0	0.08	1256	–	0	–	0	0.09	563	–	0	–	0	–	0
AirPassengers	0.04	321	0.05	323	0.03	99	0.08	310	0.19	3	0.16	0	0.04	312	0.09	1	0.16	4	0.14	4
DAYTON_hourly	0.18	370	0.12	327	–	0	0.10	1198	–	0	–	0	0.10	571	–	0	–	0	–	0
DEOK_hourly	0.16	313	0.06	322	–	0	0.09	798	–	0	–	0	0.08	473	–	0	–	0	–	0
DM	0.04	321	0.03	321	0.04	298	0.04	215	2.08	27	0.04	0	0.03	354	0.04	7	0.04	55	0.05	300
DOM_hourly	0.13	312	0.05	319	–	0	0.12	1045	–	0	–	0	0.07	520	–	0	–	0	–	0
DUQ_hourly	0.15	303	0.02	329	–	0	0.06	1286	–	0	–	0	0.06	525	–	0	–	0	–	0
EKPC_hourly	0.18	302	0.08	430	–	0	0.09	408	–	0	–	0	0.11	379	–	0	–	0	–	0
FE_hourly	0.19	303	0.08	327	–	0	0.10	790	–	0	–	0	0.09	420	–	0	–	0	–	0
M3_Monthly	–	6	–	328	–	944	–	359	–	4	–	49	–	8	–	1966	–	17	–	4
M3_Quarterly	–	3	–	326	–	1623	–	359	–	2	–	27	–	3	–	1063	–	34	–	1
M3_Yearly	–	3	0.36	336	–	3	0.70	2601	–	1	0.31	24	–	3	–	358	–	9	–	1
NI_hourly	–	0	0.10	499	–	0	0.29	741	–	0	–	0	0.17	397	–	0	–	0	–	0
NN3-Final-Dataset	–	0	0.19	336	0.18	10029	–	1599	0.21	18	0.20	3	0.16	327	0.17	136	0.28	411	–	0
NN5-Final-Dataset	–	0	0.20	345	–	0	–	1599	–	0	–	6	0.40	327	–	0	–	0	–	0
PJME_hourly	0.15	304	0.05	334	–	0	0.11	369	–	0	–	1	0.08	559	–	0	–	0	–	0
PJMW_hourly	0.13	304	0.02	339	–	0	0.05	400	–	0	–	1	0.08	539	–	0	–	0	–	0
PJM_Load_hourly	–	595	0.06	389	–	0	0.05	365	–	0	–	0	0.07	345	–	0	–	0	–	0
Twitter_volume_AAPL	0.15	309	0.25	333	0.25	3718	0.29	403	0.58	300	0.19	0	0.47	328	0.40	333	0.20	11	0.21	300
Twitter_volume_AMZN	0.20	308	0.13	335	0.22	4126	0.23	390	0.22	300	0.23	0	0.26	339	0.22	48	0.15	13	0.10	300
Twitter_volume_GOOG	0.24	311	0.24	330	0.24	3801	0.27	412	0.26	300	0.23	0	0.26	340	0.25	174	2.00	13	0.44	300
a10	0.08	312	0.08	315	0.07	117	0.11	317	0.27	6	0.15	0	0.07	312	0.15	1	0.09	3	0.18	30
ausbeer	0.04	316	0.03	342	0.03	102	0.04	310	0.08	6	0.07	0	0.03	310	0.03	5	0.03	6	–	0
auscafe	0.03	320	0.02	322	0.02	98	0.02	338	2.10	15	0.03	0	0.02	317	0.03	4	0.03	6	0.03	73
departures	0.18	322	0.12	336	0.09	518	0.11	422	0.18	182	0.18	0	0.18	1048	0.18	81	0.08	34	0.18	76
ec2_cpu_24ae8d	0.01	325	0.00	360	0.01	654	0.01	335	2.09	85	0.06	0	0.05	306	0.06	8	0.01	9	0.01	300
ec2_cpu_53ea38	0.03	326	0.02	339	0.02	650	0.02	330	1.80	85	0.04	0	0.02	326	0.04	84	0.02	10	0.03	300
ec2_cpu_5f5533	0.02	318	0.02	334	0.01	651	0.03	326	0.02	300	0.02	0	0.42	309	0.04	27	0.01	9	0.02	300
ec2_cpu_77c1ca	0.08	321	0.03	336	0.02	638	0.33	312	2.10	159	0.01	0	0.07	305	1.93	122	0.02	9	2.00	132
ec2_cpu_825cc2	0.03	302	0.03	344	–	0	0.02	350	–	0	–	0	0.01	308	–	0	–	0	–	0
ec2_cpu_ac20cd	–	0	0.17	326	–	0	0.00	340	–	0	–	0	0.03	307	–	0	–	0	–	0
ec2_cpu_c6585a	0.15	320	0.19	338	0.28	656	0.17	314	1.95	75	0.29	0	0.23	305	0.28	16	0.14	8	0.15	230
ec2_cpu_fe7f93	0.09	318	0.11	344	0.12	577	0.17	323	0.21	300	0.13	0	0.31	310	0.40	61	0.15	10	0.11	300
ec2_network_in_257a54	0.05	302	0.02	331	–	0	0.06	353	–	0	–	0	0.02	309	–	0	–	0	–	0
ec2_network_in_5abac7	0.23	316	0.56	331	0.23	678	0.44	313	0.63	300	2.00	0	0.37	304	2.00	9	0.38	10	–	108
elb_request_count_8c0756	1.43	302	–	323	–	0	1.24	320	–	0	–	0	0.62	305	–	0	–	0	–	0
elecdaily	0.06	318	0.09	320	0.11	113	0.13	407	0.16	143	0.07	0	0.16	307	0.06	7	0.19	13	0.14	135
elecdemand	0.09	319	0.05	333	0.02	876	0.04	524	2.05	300	0.03	0	0.13	309	0.04	690	0.06	92	0.06	300
exchange-2_cpc_results	–	0	0.13	418	–	0	0.22	339	–	0	–	0	0.18	306	–	0	–	0	–	0
exchange-2_cpm_results	–	0	0.22	402	–	0	0.33	335	–	0	–	0	0.24	307	–	0	–	0	–	0
exchange-4_cpc_results	0.75	302	0.67	384	–	0	0.66	397	–	0	–	0	0.63	312	–	0	–	0	–	0
exchange-4_cpm_results	0.86	302	0.70	425	–	0	0.70	330	–	0	–	0	0.63	307	–	0	–	0	–	0
exchange_rate	0.00	320	0.00	322	0.00	1329	0.01	979	1.38	300	0.00	0	0.00	379	0.00	203	0.00	232	0.01	300
gasoline	0.06	325	0.04	438	0.06	103	0.05	324	2.10	37	0.05	0	0.05	307	0.05	4	0.05	18	0.08	300
h02	0.05	328	0.10	308	0.09	99	0.10	307	0.18	4	0.21	0	0.09	310	0.20	0	0.06	6	0.17	10
hts_dataset	–	0	0.09	323	0.08	797	0.13	304	0.12	2	0.12	0	0.08	672	0.09	19	0.10	29	–	0
hyndsight	0.10	320	0.09	325	0.10	103	0.12	311	0.21	46	0.23	0	0.13	306	0.09	18	0.07	8	–	0
infant_gts	–	0	0.23	320	0.24	1550	0.43	934	0.83	3	0.37	0	0.24	343	0.33	13	0.50	66	–	0
insurance	0.15	321	0.47	289	0.36	105	1.03	313	–	0	0.51	0	0.54	306	0.19	1	0.30	6	–	0
melsyd	–	0	0.08	316	–	0	0.12	175	–	0	–	0	0.08	378	–	0	–	0	–	0
ozone-la	0.15	327	0.21	315	0.23	102	0.37	307	0.32	4	0.66	0	0.24	305	0.69	0	0.23	6	0.33	11
petrol	0.17	322	0.08	318	0.13	400	0.18	263	0.09	42	0.09	0	0.08	458	0.08	22	0.11	26	0.11	109
prison	–	0	0.12	328	0.13	111	–	476	0.38	3	0.13	1	0.12	366	0.13	27	0.15	116	–	0
prisonLF	–	0	0.14	336	0.12	1118	0.19	415	–	0	0.13	1	–	0	0.13	29	–	6	–	0
qcement	0.08	327	0.03	316	0.09	6	0.10	330	2.10	6	0.10	0	0.06	313	0.05	15	0.13	6	0.06	28
qgas	0.02	317	0.02	326	0.02	2	0.01	348	0.09	3	0.08	0	0.05	308	0.01	4	0.02	6	0.09	4
rds_cpu_cc0c53	0.05	302	0.04	343	–	0	0.13	341	–	0	–	0	0.04	305	–	0	–	0	–	0
rds_cpu_e47b3b	0.02	317	0.03	328	0.02	406	0.02	348	0.02	300	0.03	0	0.03	313	0.03	28	0.05	10	0.02	299
rossman	–	0	–	0	–	0	–	0	–	0	–	0	–	0	–	0	–	0	–	0
sunspots	0.46	306	0.42	318	0.38	4	0.74	320	0.73	163	1.23	0	0.38	306	0.35	8	–	0	0.74	95
uschange	0.38	306	0.18	310	0.32	9	0.59	303	0.75	3	0.30	0	0.38	303	0.12	3	0.35	10	0.71	11
usmelec	0.03	314	0.02	346	0.06	5	0.03	314	0.09	47	0.08	0	0.02	305	0.08	14	0.04	7	0.08	29
visnights	0.17	312	0.14	331	0.14	119	0.28	814	0.36	24	0.17	0	0.09	409	0.16	39	0.15	114	0.25	7

Table 4: Raw experiment results for all evaluated frameworks in the ablation study and used datasets using MASE. Predictions with consistent failures are marked with –.

Dataset	Auto-PyTorch		Multi-Fidelity		Templates		Warm-Starting		auto-sktime	
	MASE	Duration	MASE	Duration	MASE	Duration	MASE	Duration	MASE	Duration
AEDemand	–	310	0.18	320	0.18	311	0.20	309	0.02	307
AEP_hourly	5.59	317	3.24	341	3.24	416	2.31	320	2.31	323
AirPassengers	0.78	321	0.93	323	0.94	317	0.78	315	0.84	317
DAYTON_hourly	5.81	370	3.91	327	3.91	406	2.80	327	2.72	330
DEOK_hourly	5.38	313	2.03	322	2.03	348	2.84	336	1.91	334
DM	3.39	321	2.38	321	3.55	317	3.05	339	2.32	359
DOM_hourly	4.05	312	1.81	319	1.81	326	1.74	321	1.66	316
DUQ_hourly	5.21	303	0.81	329	0.90	327	0.84	324	0.81	326
EKPC_hourly	4.25	302	1.93	430	2.59	371	3.00	356	1.89	372
FE_hourly	7.06	303	3.19	327	3.11	323	3.07	341	2.99	360
M3_Monthly	–	6	–	328	–	337	–	326	–	588
M3_Quarterly	–	3	–	326	–	313	–	327	–	342
M3_Yearly	–	3	9.78	336	12.53	333	6.18	399	6.58	362
NI_hourly	–	0	2.84	499	3.05	334	3.73	341	3.22	359
NN3-Final-Dataset	–	0	1.01	336	1.04	345	1.85	375	1.08	347
NN5-Final-Dataset	–	0	0.44	345	0.44	343	0.93	343	0.45	346
PJME_hourly	5.31	304	1.62	334	2.09	323	1.59	325	1.59	328
PJMW_hourly	4.50	304	0.87	339	0.96	343	1.51	320	1.00	322
PJM_Load_hourly	–	595	1.68	389	1.68	393	1.90	335	1.58	330
Twitter_volume_AAPL	0.18	309	0.33	333	0.52	431	0.51	343	0.43	345
Twitter_volume_AMZN	0.88	308	0.58	335	0.76	503	0.52	342	0.52	344
Twitter_volume_GOOG	1.67	311	1.67	330	2.31	615	3.63	348	3.63	357
a10	1.30	312	1.40	315	1.37	309	1.64	313	1.57	316
ausbeer	0.26	316	0.20	342	0.21	370	0.19	366	0.19	366
auscafe	1.59	320	1.08	322	1.20	321	1.30	316	1.23	321
departures	3.23	322	1.28	336	1.37	363	4.18	350	1.40	389
ec2_cpu_24ae8d	0.03	325	0.01	360	0.01	337	0.06	315	0.04	319
ec2_cpu_53ea38	0.52	326	0.28	339	0.33	363	0.32	351	0.31	352
ec2_cpu_5f5533	0.19	318	0.18	334	0.18	334	0.39	333	0.27	369
ec2_cpu_77c1ca	0.00	321	0.00	336	0.00	329	0.00	343	0.00	354
ec2_cpu_825cc2	1.53	302	1.29	344	1.65	395	1.45	337	0.86	345
ec2_cpu_ac20cd	–	0	6.32	326	9.76	316	3.86	349	2.88	349
ec2_cpu_c6585a	0.35	320	0.42	338	1.25	330	0.47	343	0.36	351
ec2_cpu_fe7f93	0.09	318	0.10	344	0.11	349	0.11	352	0.11	351
ec2_network_in_257a54	0.02	302	0.01	331	0.01	346	0.01	349	0.01	342
ec2_network_in_5abac7	0.00	316	0.00	331	0.00	323	0.00	355	0.00	356
elb_request_count_8c0756	0.61	302	0.64	323	0.53	320	0.45	340	0.45	342
elecdaily	0.66	318	1.10	320	1.20	321	1.40	327	1.05	325
elecdemand	2.93	319	1.78	333	2.98	398	1.85	362	1.59	382
exchange-2_cpc_results	–	0	0.93	418	0.93	418	0.96	325	0.95	332
exchange-2_cpm_results	–	0	1.12	402	1.13	403	1.20	323	1.14	320
exchange-4_cpc_results	1.75	302	1.43	384	1.43	384	1.50	326	1.54	368
exchange-4_cpm_results	1.90	302	1.30	425	1.29	404	1.22	372	1.42	371
exchange_rate	0.94	320	1.32	322	1.32	338	68.54	329	2.85	333
gasoline	2.01	325	1.40	438	1.55	323	1.64	338	1.38	411
h02	0.48	328	0.84	308	0.84	302	0.68	309	0.66	310
hts_dataset	–	0	0.46	323	0.46	321	0.62	344	0.46	354
hyndsight	0.56	320	0.54	325	0.53	322	0.52	323	0.52	312
infant_gts	–	0	0.62	320	0.62	320	1.34	333	0.76	329
insurance	1.55	321	3.69	289	3.58	286	3.93	302	4.00	312
melsyd	–	0	0.78	316	0.93	317	6.76	326	1.38	314
ozone-la	0.49	327	0.64	315	0.67	310	0.46	305	0.41	319
petrol	1.86	322	1.04	318	1.04	316	2.81	313	1.25	312
prison	–	0	2.45	328	2.45	340	5.01	320	4.32	317
prisonLF	–	0	2.94	336	3.83	321	4.97	329	3.64	326
qcement	1.49	327	0.53	316	0.55	307	0.61	317	0.44	304
qgas	0.30	317	0.24	326	0.28	337	0.69	312	0.56	323
rds_cpu_cc0c53	1.39	302	1.06	343	1.08	408	1.08	356	0.88	321
rds_cpu_e47b3b	0.48	317	0.64	328	1.65	326	0.90	337	0.90	345
rossman	–	0	–	0	–	0	–	0	–	0
sunspots	0.89	306	0.82	318	1.01	344	0.90	306	0.78	315
uschange	0.49	306	0.22	310	0.19	304	0.21	304	0.22	312
usmelec	0.55	314	0.36	346	0.39	349	0.38	336	0.36	336
visnights	0.72	312	0.63	331	0.61	364	0.93	339	0.62	338

Table 5: Raw experiment results for all evaluated frameworks in the ablation study and used datasets using RMSE. Predictions with consistent failures are marked with –.

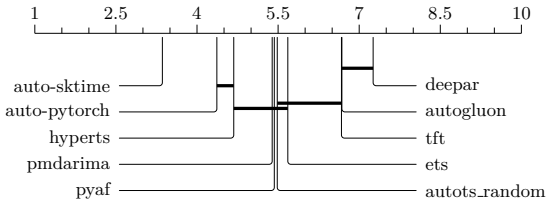
Dataset	Auto-PyTorch		Multi-Fidelity		Templates		Warm-Starting		auto-sktime	
	RMSE	Duration	RMSE	Duration	RMSE	Duration	RMSE	Duration	RMSE	Duration
AEDemand	–	310	0.48	320	0.48	311	0.49	309	0.04	307
AEP_hourly	2807.39	317	1758.63	341	2074.84	416	1222.00	320	1222.00	323
AirPassengers	24.15	321	26.08	323	25.89	317	22.50	315	23.84	317
DAYTON_hourly	442.30	370	317.72	327	317.72	406	225.38	327	219.94	330
DEOK_hourly	607.46	313	255.92	322	255.92	348	332.32	336	231.93	334
DM	0.10	321	0.07	321	0.10	317	0.09	339	0.07	359
DOM_hourly	1877.51	312	1038.07	319	1038.07	326	1013.58	321	955.58	316
DUQ_hourly	293.03	303	50.01	329	54.40	327	50.31	324	48.59	326
EKPC_hourly	288.85	302	138.17	430	181.67	371	204.01	356	131.88	372
FE_hourly	1860.53	303	863.76	327	848.06	323	830.96	341	806.40	360
M3_Monthly	–	6	–	328	–	337	–	326	–	588
M3_Quarterly	–	3	–	326	–	313	–	327	–	342
M3_Yearly	–	3	2897.38	336	3259.23	333	1611.30	399	1682.68	362
NI_hourly	–	0	1166.77	499	1278.55	334	1528.85	341	1295.72	359
NN3-Final-Dataset	–	0	1024.01	336	1062.72	345	1751.30	375	1099.58	347
NN5-Final-Dataset	–	0	4.18	345	4.18	343	7.93	343	4.24	346
PJME_hourly	6571.53	304	2010.92	334	2668.06	323	1988.94	325	1988.94	328
PJMW_hourly	863.96	304	179.42	339	209.63	343	315.02	320	217.85	322
PJM_Load_hourly	–	595	2185.96	389	2271.94	393	2454.68	335	2237.83	330
Twitter_volume_AAPL	8.27	309	13.38	333	19.62	431	18.74	343	16.13	345
Twitter_volume_AMZN	11.53	308	7.96	335	10.22	503	7.26	342	7.26	344
Twitter_volume_GOOG	15.44	311	15.60	330	22.27	615	31.34	348	31.34	357
a10	2.11	312	2.09	315	2.10	309	2.43	313	2.40	316
ausbeer	15.30	316	13.18	342	13.75	370	12.04	366	11.79	366
auscafe	0.16	320	0.11	322	0.12	321	0.12	316	0.12	321
departures	72.22	322	28.45	336	31.17	363	89.80	350	31.53	389
ec2_cpu_24ae8d	0.00	325	0.00	360	0.00	337	0.00	315	0.00	319
ec2_cpu_53ea38	0.07	326	0.04	339	0.05	363	0.04	351	0.04	352
ec2_cpu_5f5533	1.07	318	0.85	334	0.84	334	1.72	333	1.17	369
ec2_cpu_77c1ca	0.01	321	0.00	336	0.00	329	0.00	343	0.00	354
ec2_cpu_825cc2	3.07	302	2.64	344	3.31	395	2.92	337	1.81	345
ec2_cpu_ac20cd	–	0	14.31	326	21.96	316	8.76	349	6.56	349
ec2_cpu_c6585a	0.03	320	0.03	338	0.06	330	0.03	343	0.03	351
ec2_cpu_fe7f93	0.37	318	0.38	344	0.42	349	0.43	352	0.42	351
ec2_network_in_257a54	12114.54	302	6700.46	331	6700.46	346	9152.59	349	8415.57	342
ec2_network_in_5abac7	18.05	316	162.29	331	265.72	323	23.92	355	23.05	356
elb_request_count_8c0756	37.31	302	–	323	32.19	320	27.74	340	27.74	342
elecdaily	12.73	318	19.86	320	20.39	321	23.45	327	17.90	325
elecdemand	0.38	319	0.24	333	0.37	398	0.25	362	0.22	382
exchange-2_cpc_results	–	0	0.02	418	0.02	418	0.02	325	0.02	332
exchange-2_cpm_results	–	0	0.10	402	0.10	403	0.10	323	0.10	320
exchange-4_cpc_results	0.06	302	0.05	384	0.05	384	0.05	326	0.06	368
exchange-4_cpm_results	0.40	302	0.28	425	0.28	404	0.26	372	0.31	371
exchange_rate	0.00	320	0.00	322	0.00	338	0.19	329	0.01	333
gasoline	0.63	325	0.46	438	0.50	323	0.55	338	0.42	411
h02	0.07	328	0.10	308	0.10	302	0.08	309	0.08	310
hts_dataset	–	0	1000.83	323	999.59	321	1367.25	344	1025.17	354
hyndsight	193.28	320	202.10	325	195.96	322	200.35	323	199.60	312
infant_gts	–	0	15.99	320	15.99	320	31.80	333	19.16	329
insurance	2.43	321	5.73	289	5.47	286	6.00	302	5.99	312
melsyd	–	0	0.66	316	0.76	317	4.75	326	1.10	314
ozone-la	0.50	327	0.64	315	0.64	310	0.47	305	0.42	319
petrol	4.68	322	2.75	318	2.76	316	6.76	313	3.22	312
prison	–	0	61.85	328	61.79	340	122.74	320	106.34	317
prisonLF	–	0	161.69	336	175.21	321	251.98	329	164.27	326
qcement	0.19	327	0.07	316	0.08	307	0.09	317	0.07	304
qgas	4.76	317	4.25	326	4.99	337	11.97	312	9.68	323
rds_cpu_cc0c53	0.86	302	0.62	343	0.63	408	0.66	356	0.53	321
rds_cpu_e47b3b	0.45	317	0.63	328	1.47	326	0.80	337	0.80	345
rossman	–	0	–	0	–	0	–	0	–	0
sunspots	17.83	306	19.77	318	24.97	344	21.34	306	19.33	315
uschange	0.31	306	0.16	310	0.15	304	0.15	304	0.16	312
usmelec	11.78	314	8.53	346	8.99	349	8.09	336	7.78	336
visnights	0.72	312	0.61	331	0.59	364	0.89	339	0.59	338

Table 6: Raw experiment results for all evaluated frameworks in the ablation study and used datasets using sMAPE. Predictions with consistent failures are marked with –.

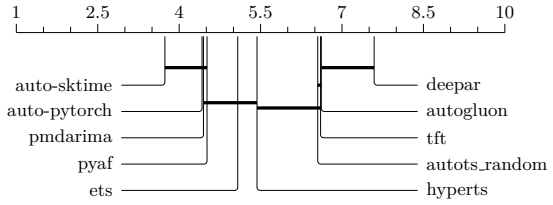
Dataset	Auto-PyTorch		Multi-Fidelity		Templates		Warm-Starting		auto-sktime	
	sMAPE	Duration	sMAPE	Duration	sMAPE	Duration	sMAPE	Duration	sMAPE	Duration
AEDemand	–	310	0.00	320	0.00	311	0.00	309	0.00	307
AEP_hourly	0.16	317	0.09	341	0.11	416	0.06	320	0.06	323
AirPassengers	0.04	321	0.05	323	0.05	317	0.04	315	0.04	317
DAYTON_hourly	0.18	370	0.12	327	0.12	406	0.08	327	0.08	330
DEOK_hourly	0.16	313	0.06	322	0.06	348	0.08	336	0.05	334
DM	0.04	321	0.03	321	0.05	317	0.04	339	0.03	359
DOM_hourly	0.13	312	0.05	319	0.05	326	0.05	321	0.05	316
DUQ_hourly	0.15	303	0.02	329	0.03	327	0.02	324	0.02	326
EKPC_hourly	0.18	302	0.08	430	0.11	371	0.12	356	0.08	372
FE_hourly	0.19	303	0.08	327	0.08	323	0.08	341	0.08	360
M3_Monthly	–	6	–	328	–	337	–	326	–	588
M3_Quarterly	–	3	–	326	–	313	–	327	–	342
M3_Yearly	–	3	0.36	336	0.42	333	0.25	399	0.26	362
NI_hourly	–	0	0.10	499	0.11	334	0.13	341	0.12	359
NN3-Final-Dataset	–	0	0.19	336	0.19	345	0.39	375	0.20	347
NN5-Final-Dataset	–	0	0.20	345	0.20	343	0.39	343	0.20	346
PJME_hourly	0.15	304	0.05	334	0.06	323	0.05	325	0.05	328
PJMW_hourly	0.13	304	0.02	339	0.03	343	0.04	320	0.03	322
PJM_Load_hourly	–	595	0.06	389	0.06	393	0.06	335	0.06	330
Twitter_volume_AAPL	0.15	309	0.25	333	0.34	431	0.35	343	0.30	345
Twitter_volume_AMZN	0.20	308	0.13	335	0.17	503	0.12	342	0.12	344
Twitter_volume_GOOG	0.24	311	0.24	330	0.41	615	0.65	348	0.65	357
a10	0.08	312	0.08	315	0.08	309	0.09	313	0.09	316
ausbeer	0.04	316	0.03	342	0.03	370	0.03	366	0.03	366
auscafe	0.03	320	0.02	322	0.03	321	0.03	316	0.03	321
departures	0.18	322	0.12	336	0.12	363	0.41	350	0.13	389
ec2_cpu_24ae8d	0.01	325	0.00	360	0.00	337	0.02	315	0.01	319
ec2_cpu_53ea38	0.03	326	0.02	339	0.02	363	0.02	351	0.02	352
ec2_cpu_5f5533	0.02	318	0.02	334	0.02	334	0.04	333	0.03	369
ec2_cpu_77c1ca	0.08	321	0.03	336	0.03	329	0.03	343	0.02	354
ec2_cpu_825cc2	0.03	302	0.03	344	0.03	395	0.03	337	0.02	345
ec2_cpu_ac20cd	–	0	0.17	326	0.26	316	0.09	349	0.07	349
ec2_cpu_c6585a	0.15	320	0.19	338	0.36	330	0.20	343	0.15	351
ec2_cpu_fe7f93	0.09	318	0.11	344	0.11	349	0.12	352	0.12	351
ec2_network_in_257a54	0.05	302	0.02	331	0.02	346	0.03	349	0.03	342
ec2_network_in_5abac7	0.23	316	0.56	331	0.64	323	0.30	355	0.29	356
elb_request_count_8c0756	1.43	302	–	323	1.13	320	0.78	340	0.78	342
elecdaily	0.06	318	0.09	320	0.10	321	0.11	327	0.09	325
elecdemand	0.09	319	0.05	333	0.08	398	0.05	362	0.04	382
exchange-2_cpc_results	–	0	0.13	418	0.13	418	0.14	325	0.14	332
exchange-2_cpm_results	–	0	0.22	402	0.23	403	0.24	323	0.23	320
exchange-4_cpc_results	0.75	302	0.67	384	0.67	384	0.69	326	0.70	368
exchange-4_cpm_results	0.86	302	0.70	425	0.70	404	0.67	372	0.77	371
exchange_rate	0.00	320	0.00	322	0.00	338	0.46	329	0.01	333
gasoline	0.06	325	0.04	438	0.04	323	0.05	338	0.04	411
h02	0.05	328	0.10	308	0.10	302	0.08	309	0.08	310
hts_dataset	–	0	0.09	323	0.09	321	0.12	344	0.09	354
hyndsight	0.10	320	0.09	325	0.08	322	0.08	323	0.08	312
infant_gts	–	0	0.23	320	0.23	320	0.41	333	0.28	329
insurance	0.15	321	0.47	289	0.46	286	0.51	302	0.52	312
melsyd	–	0	0.08	316	0.08	317	0.80	326	0.22	314
ozone-la	0.15	327	0.21	315	0.23	310	0.15	305	0.13	319
petrol	0.17	322	0.08	318	0.08	316	0.20	313	0.09	312
prison	–	0	0.12	328	0.12	340	0.17	320	0.14	317
prisonLF	–	0	0.14	336	0.15	321	0.21	329	0.14	326
qcement	0.08	327	0.03	316	0.03	307	0.03	317	0.02	304
qgas	0.02	317	0.02	326	0.02	337	0.04	312	0.04	323
rds_cpu_cc0c53	0.05	302	0.04	343	0.04	408	0.04	356	0.03	321
rds_cpu_e47b3b	0.02	317	0.03	328	0.09	326	0.04	337	0.04	345
rossman	–	0	–	0	–	0	–	0	–	0
sunspots	0.46	306	0.42	318	0.45	344	0.42	306	0.40	315
uschange	0.38	306	0.18	310	0.15	304	0.16	304	0.18	312
usmelec	0.03	314	0.02	346	0.02	349	0.02	336	0.02	336
visnights	0.17	312	0.14	331	0.14	364	0.21	339	0.14	338

2 Ranking of Evaluated Methods

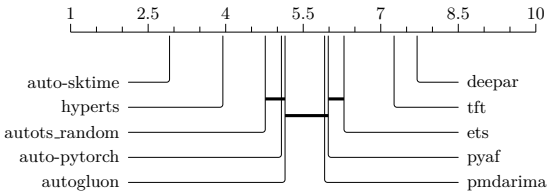
Figure 1 contains critical difference plots for all evaluated scenarios in Section 4.2 and 4.3. The average ranking of each evaluated method is displayed, and whether observed values differ significantly from each other according to a t-test with significance level $p = 0.05$ and Bonferroni correction for multiple hypothesis tests.



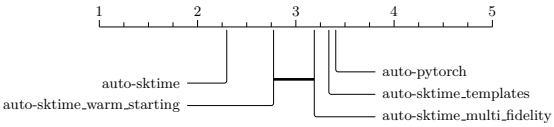
(a) All time series while enforcing timeouts.



(b) Time series without missing values with timeouts.



(c) All time series without enforcing timeouts.



(d) All time series while enforcing timeouts.

Figure 1: Critical difference diagram of all evaluated framework combinations.

3 Used Computing Hardware

All experiments were conducted on standard hardware, namely *e2-standard-4* virtual machines on Google Cloud Platform equipped with Intel Xeon E5 processors with four cores and 16 GB memory. Even though multiple tested frameworks construct neural networks, no graphics card was used. Some frameworks make extensive usage of parallelization of workload over all available cores, while other frameworks do not do this. To ensure fair comparisons, all frameworks were limited to exactly one CPU core to ensure identical computational budgets.