Appendix

Table 5.1: Results of the ablation study of the Machine learning architecture.

ID	Config.	MSE	RMSE	MAE	MAPE	R^2	CVRMSE	Training Time (s)	Overfit Gap	Loss Ratio
1	Linear Regression:	141776.29	376.53	252.09	0.87	1	1.26	0.88	-60.543908	0.86
2	Ridge Regression: alpha=0.01	141877.73	376.67	252.12	0.87	1	1.26	0.19	-60.413898	0.86
3	Ridge Regression: alpha=0.1	143002.98	378.16	252.65	0.88	1	1.26	0.2	-59.291826	0.86
4	Ridge Regression: alpha=1	158254.53	397.81	263.42	0.91	1	1.31	0.21	-53.272749	0.88
5	Ridge Regression: alpha=10	265867.72	515.62	346.11	1.2	0.99	1.66	0.18	-45.975812	0.92
6	Lasso Regression: alpha=0.01	4796216.96	2190.03	1859.82	7.23	0.87	7.17	3.27	15.625081	1.01
7	Lasso Regression: alpha=0.1	51073484.12	7146.57	6004.68	23.43	-0.34	24.64	0.19	-20.620023	1
8	Lasso Regression: alpha=1	51073484.12	7146.57	6004.68	23.43	-0.34	24.64	0.13	-20.620023	1
9	Random Forest: max_depth=5, min_samples_split=2, n_estimators=50	363156.9	602.63	435.79	1.48	0.99	1.94	155.66	-36.752922	0.94
10	Random Forest: max_depth=5, min_samples_split=2, n_estimators=100	366169.09	605.12	435.92	1.48	0.99	1.95	289.49	-33.181601	0.95
11	Random Forest: max_depth=5, min_samples_split=2, n_estimators=200	358232.25	598.53	431.07	1.46	0.99	1.93	597.8	-38.577704	0.94
12	Random Forest: max_depth=5, min_samples_split=5, n_estimators=50	361047.26	600.87	432.52	1.47	0.99	1.94	149.53	-35.486951	0.94
13	Random Forest: max_depth=5, min_samples_split=5, n_estimators=100	361284.02	601.07	433.51	1.47	0.99	1.94	293.46	-36.394789	0.94
14	Random Forest: max_depth=5, min_samples_split=5, n_estimators=200	356002.1	596.66	429.94	1.46	0.99	1.93	598.08	-40.202596	0.94
15	Random Forest: max_depth=5, min_samples_split=10, n_estimators=50	357328.77	597.77	428.27	1.45	0.99	1.91	147.14	-36.857328	0.94
16	Random Forest: max_depth=5, min_samples_split=10, n_estimators=100	367999.61	606.63	436.9	1.48	0.99	1.95	290.59	-32.860473	0.95
17	Random Forest: max_depth=5, min_samples_split=10, n_estimators=200	357672.33	598.06	429.28	1.45	0.99	1.92	627.38	-37.410473	0.94
18	Random Forest: max_depth=10, min_samples_split=2, n_estimators=50	150481.85	387.92	267.59	0.93	1	1.29	269.58	99.428901	1.34
19	Random Forest: max_depth=10, min_samples_split=2, n_estimators=100	150345.15	387.74	267.28	0.93	1	1.28	529.02	101.082790	1.35

Table 5.1: Results of the ablation study of the Machine learning architecture.

ID	Config.	MSE	RMSE	MAE	MAPE	R^2	CVRMSE	Training Time (s)	Overfit Gap	Loss Ratio
20	Random Forest: max__depth=10, min__samples__split=2, n__estimators=200	148797.62	385.74	265.66	0.92	1	1.29	1066.13	99.690336	1.35
21	Random Forest: max__depth=10, min__samples__split=5, n__estimators=50	151491.27	389.22	267.65	0.93	1	1.3	261.08	96.190329	1.33
22	Random Forest: max__depth=10, min__samples__split=5, n__estimators=100	149222.53	386.29	266.1	0.92	1	1.29	527.68	94.224497	1.32
23	Random Forest: max__depth=10, min__samples__split=5, n__estimators=200	149435.28	386.57	266.92	0.92	1	1.29	1058.31	95.272409	1.33
24	Random Forest: max__depth=10, min__samples__split=10, n__estimators=50	150048.05	387.36	266.87	0.92	1	1.29	274.54	91.887739	1.31
25	Random Forest: max__depth=10, min__samples__split=10, n__estimators=100	149296.06	386.39	265.57	0.92	1	1.29	548.87	92.885181	1.32
26	Random Forest: max__depth=10, min__samples__split=10, n__estimators=200	148903.82	385.88	265.81	0.92	1	1.29	1080.86	91.662446	1.31
27	Random Forest: max__depth=15, min__samples__split=2, n__estimators=50	142872.21	377.98	260.28	0.9	1	1.27	369.28	200.962790	2.14
28	Random Forest: max__depth=15, min__samples__split=2, n__estimators=100	143612.18	378.96	260.68	0.9	1	1.27	759.03	203.683264	2.16
29	Random Forest: max__depth=15, min__samples__split=2, n__estimators=200	142150.81	377.03	258.96	0.9	1	1.26	1498.74	202.137978	2.16
30	Random Forest: max__depth=15, min__samples__split=5, n__estimators=50	144728.8	380.43	262.72	0.91	1	1.28	367.5	188.572266	1.98
31	Random Forest: max__depth=15, min__samples__split=5, n__estimators=100	143400.99	378.68	261.35	0.91	1	1.27	711.89	189.268092	2
32	Random Forest: max__depth=15, min__samples__split=5, n__estimators=200	142892.33	378.01	259.45	0.9	1	1.26	1378.22	187.875811	1.99
33	Random Forest: max__depth=15, min__samples__split=10, n__estimators=50	146923.82	383.31	263.56	0.91	1	1.28	335.28	176.322837	1.85
34	Random Forest: max__depth=15, min__samples__split=10, n__estimators=100	144592.12	380.25	260.97	0.9	1	1.27	673.28	175.653349	1.86
35	Random Forest: max__depth=15, min__samples__split=10, n__estimators=200	143289.95	378.54	260.08	0.9	1	1.26	1340.55	172.847759	1.84
36	Gradient Boosting: learning__rate=0.001, max__depth=3, n__estimators=50, subsample=0.7	46380107.07	6810.29	5719.46	22.32	-0.22	23.5	98.95	-18.215639	1
37	Gradient Boosting: learning__rate=0.001, max__depth=3, n__estimators=50, subsample=0.8	46376460.17	6810.03	5719.32	22.32	-0.22	23.5	114.19	-18.472051	1
38	Gradient Boosting: learning__rate=0.001, max__depth=3, n__estimators=50, subsample=1.0	46375660.23	6809.97	5719.24	22.32	-0.22	23.5	142.6	-18.558061	1

Table 5.1: Results of the ablation study of the Machine learning architecture.

ID	Config.	MSE	RMSE	MAE	MAPE	R^2	CVRMSE	Training Time (s)	Overfit Gap	Loss Ratio
39	Gradient Boosting: learning_rate=0.001, max_depth=3, n_estimators=100, subsample=0.7	42129249.48	6490.7	5448.18	21.27	-0.11	22.42	198.37	-15.783533	1
40	Gradient Boosting: learning_rate=0.001, max_depth=3, n_estimators=100, subsample=0.8	42128905.9	6490.68	5448.14	21.27	-0.11	22.42	226.88	-15.840599	1
41	Gradient Boosting: learning_rate=0.001, max_depth=3, n_estimators=100, subsample=1.0	42131596.04	6490.89	5448.18	21.27	-0.11	22.42	284.46	-15.639223	1
42	Gradient Boosting: learning_rate=0.001, max_depth=3, n_estimators=200, subsample=0.7	34816123.75	5900.52	4946.47	19.32	0.09	20.42	395.49	-8.176807	1
43	Gradient Boosting: learning_rate=0.001, max_depth=3, n_estimators=200, subsample=0.8	34827886.64	5901.52	4947.01	19.32	0.08	20.43	453.63	-7.272416	1
44	Gradient Boosting: learning_rate=0.001, max_depth=3, n_estimators=200, subsample=1.0	34824996.06	5901.27	4946.88	19.32	0.08	20.43	567.59	-7.442802	1
45	Gradient Boosting: learning_rate=0.001, max_depth=5, n_estimators=50, subsample=0.7	46246616.06	6800.49	5715.25	22.3	-0.22	23.44	155.66	-20.350767	1
46	Gradient Boosting: learning_rate=0.001, max_depth=5, n_estimators=50, subsample=0.8	46242079.24	6800.15	5714.98	22.3	-0.22	23.44	179.49	-20.683194	1
47	Gradient Boosting: learning_rate=0.001, max_depth=5, n_estimators=50, subsample=1.0	46241654.53	6800.12	5714.9	22.3	-0.22	23.44	225.12	-20.736567	1
48	Gradient Boosting: learning_rate=0.001, max_depth=5, n_estimators=100, subsample=0.7	41880911.38	6471.55	5440.06	21.22	-0.1	22.3	312.85	-19.965498	1
49	Gradient Boosting: learning_rate=0.001, max_depth=5, n_estimators=100, subsample=0.8	41882237.62	6471.65	5440.21	21.23	-0.1	22.3	357.9	-19.874906	1
50	Gradient Boosting: learning_rate=0.001, max_depth=5, n_estimators=100, subsample=1.0	41874776.68	6471.07	5439.7	21.22	-0.1	22.3	450.1	-20.446725	1
51	Gradient Boosting: learning_rate=0.001, max_depth=5, n_estimators=200, subsample=0.7	34364793.71	5862.15	4930.2	19.23	0.1	20.19	622.69	-18.375008	1
52	Gradient Boosting: learning_rate=0.001, max_depth=5, n_estimators=200, subsample=0.8	34365204.85	5862.18	4930.23	19.23	0.1	20.19	715.86	-18.351626	1

Table 5.1: Results of the ablation study of the Machine learning architecture.

ID	Config.	MSE	RMSE	MAE	MAPE	R^2	CVRMSE	Training Time (s)	Overfit Gap	Loss Ratio
53	Gradient Boosting: learning_rate=0.001, max_depth=5, n_estimators=200, subsample=1.0	34363942.65	5862.08	4930.21	19.23	0.1	20.19	900.78	-18.405702	1
54	Gradient Boosting: learning_rate=0.001, max_depth=7, n_estimators=50, subsample=0.7	46241792.52	6800.13	5714.21	22.3	-0.22	23.44	209.96	-19.057124	1
55	Gradient Boosting: learning_rate=0.001, max_depth=7, n_estimators=50, subsample=0.8	46242396.82	6800.18	5714.27	22.3	-0.22	23.44	240.27	-18.987459	1
56	Gradient Boosting: learning_rate=0.001, max_depth=7, n_estimators=50, subsample=1.0	46242661.21	6800.2	5714.28	22.3	-0.22	23.45	301.72	-18.959817	1
57	Gradient Boosting: learning_rate=0.001, max_depth=7, n_estimators=100, subsample=0.7	41871123.43	6470.79	5438.09	21.22	-0.1	22.31	418.51	-17.392053	1
58	Gradient Boosting: learning_rate=0.001, max_depth=7, n_estimators=100, subsample=0.8	41870746.38	6470.76	5438.07	21.22	-0.1	22.31	485.57	-17.416664	1
59	Gradient Boosting: learning_rate=0.001, max_depth=7, n_estimators=100, subsample=1.0	41873405.62	6470.97	5438.15	21.22	-0.1	22.31	603.58	-17.190756	1
60	Gradient Boosting: learning_rate=0.001, max_depth=7, n_estimators=200, subsample=0.7	34335939.6	5859.69	4925.5	19.22	0.1	20.19	837.88	-14.267054	1
61	Gradient Boosting: learning_rate=0.001, max_depth=7, n_estimators=200, subsample=0.8	34337146.87	5859.79	4925.6	19.22	0.1	20.19	961.4	-14.147892	1
62	Gradient Boosting: learning_rate=0.001, max_depth=7, n_estimators=200, subsample=1.0	34340677.25	5860.09	4925.62	19.22	0.1	20.19	1210.31	-13.827036	1
63	Gradient Boosting: learning_rate=0.01, max_depth=3, n_estimators=50, subsample=0.7	19666813.34	4434.73	3698.79	14.47	0.48	15.46	99.07	9.583665	1
64	Gradient Boosting: learning_rate=0.01, max_depth=3, n_estimators=50, subsample=0.8	19647865.84	4432.59	3696.46	14.46	0.48	15.46	113.95	7.589609	1
65	Gradient Boosting: learning_rate=0.01, max_depth=3, n_estimators=50, subsample=1.0	19660227.21	4433.99	3697.76	14.47	0.48	15.46	143.21	8.856403	1
66	Gradient Boosting: learning_rate=0.01, max_depth=3, n_estimators=100, subsample=0.7	7707468.43	2776.23	2291.61	8.98	0.8	9.8	197.75	7.084020	1

Table 5.1: Results of the ablation study of the Machine learning architecture.

ID	Config.	MSE	RMSE	MAE	MAPE	R^2	CVRMSE	Training Time (s)	Overfit Gap	Loss Ratio
67	Gradient Boosting: learning_rate=0.01, max_depth=3, n_estimators=100, subsample=0.8	7710375.34	2776.76	2291.6	8.98	0.8	9.8	227.32	7.227094	1
68	Gradient Boosting: learning_rate=0.01, max_depth=3, n_estimators=100, subsample=1.0	7701978.31	2775.24	2290.29	8.98	0.8	9.8	285.47	6.065712	1
69	Gradient Boosting: learning_rate=0.01, max_depth=3, n_estimators=200, subsample=0.7	1416068.85	1189.99	957.21	3.73	0.96	4.17	396.34	-15.960600	0.99
70	Gradient Boosting: learning_rate=0.01, max_depth=3, n_estimators=200, subsample=0.8	1415283.59	1189.66	956.65	3.72	0.96	4.17	455.45	-16.590333	0.99
71	Gradient Boosting: learning_rate=0.01, max_depth=3, n_estimators=200, subsample=1.0	1417310.31	1190.51	956.73	3.73	0.96	4.18	572.3	-15.502716	0.99
72	Gradient Boosting: learning_rate=0.01, max_depth=5, n_estimators=50, subsample=0.7	18932454.48	4351.14	3664.02	14.29	0.5	14.97	156.58	-15.380127	1
73	Gradient Boosting: learning_rate=0.01, max_depth=5, n_estimators=50, subsample=0.8	18932716.55	4351.17	3664.95	14.3	0.5	14.97	179.36	-15.036305	1
74	Gradient Boosting: learning_rate=0.01, max_depth=5, n_estimators=50, subsample=1.0	18931482.81	4351.03	3663.45	14.29	0.5	14.97	226.24	-15.375018	1
75	Gradient Boosting: learning_rate=0.01, max_depth=5, n_estimators=100, subsample=0.7	7098875.68	2664.37	2245.27	8.76	0.81	9.15	323.91	-13.810648	0.99
76	Gradient Boosting: learning_rate=0.01, max_depth=5, n_estimators=100, subsample=0.8	7103358.27	2665.21	2245.33	8.76	0.81	9.15	360.31	-12.869325	1
77	Gradient Boosting: learning_rate=0.01, max_depth=5, n_estimators=100, subsample=1.0	7099672.2	2664.52	2245.52	8.76	0.81	9.15	453.55	-13.556401	0.99
78	Gradient Boosting: learning_rate=0.01, max_depth=5, n_estimators=200, subsample=0.7	1114155.74	1055.54	867.72	3.39	0.97	3.56	627.53	-15.288416	0.99
79	Gradient Boosting: learning_rate=0.01, max_depth=5, n_estimators=200, subsample=0.8	1112603.74	1054.8	865.86	3.38	0.97	3.56	721.64	-15.642747	0.99
80	Gradient Boosting: learning_rate=0.01, max_depth=5, n_estimators=200, subsample=1.0	1113624.47	1055.28	865.72	3.38	0.97	3.56	906.73	-15.216075	0.99

Table 5.1: Results of the ablation study of the Machine learning architecture.

ID	Config.	MSE	RMSE	MAE	MAPE	R^2	CVRMSE	Training Time (s)	Overfit Gap	Loss Ratio
81	Gradient Boosting: learning_rate=0.01, max_depth=7, n_estimators=50, subsample=0.7	18889500.48	4346.21	3656.32	14.27	0.5	14.96	210.14	-5.388431	1
82	Gradient Boosting: learning_rate=0.01, max_depth=7, n_estimators=50, subsample=0.8	18889224.64	4346.17	3656.49	14.27	0.5	14.96	241.35	-5.466385	1
83	Gradient Boosting: learning_rate=0.01, max_depth=7, n_estimators=50, subsample=1.0	18890316.21	4346.3	3655.96	14.27	0.5	14.96	304.13	-5.247794	1
84	Gradient Boosting: learning_rate=0.01, max_depth=7, n_estimators=100, subsample=0.7	7072433.7	2659.4	2239.47	8.75	0.81	9.14	420.28	6.839926	1
85	Gradient Boosting: learning_rate=0.01, max_depth=7, n_estimators=100, subsample=0.8	7070835.13	2659.1	2239.95	8.75	0.81	9.13	482.7	6.615332	1
86	Gradient Boosting: learning_rate=0.01, max_depth=7, n_estimators=100, subsample=1.0	7069960.87	2658.94	2239.64	8.75	0.81	9.13	607.95	6.355840	1
87	Gradient Boosting: learning_rate=0.01, max_depth=7, n_estimators=200, subsample=0.7	1112910.31	1054.95	880.15	3.43	0.97	3.59	844.2	32.092962	1.03
88	Gradient Boosting: learning_rate=0.01, max_depth=7, n_estimators=200, subsample=0.8	1116959.99	1056.86	882.4	3.44	0.97	3.6	968.28	33.778811	1.03
89	Gradient Boosting: learning_rate=0.01, max_depth=7, n_estimators=200, subsample=1.0	1120141.12	1058.37	883.43	3.44	0.97	3.6	1215.36	35.006641	1.03
90	Gradient Boosting: learning_rate=0.1, max_depth=3, n_estimators=50, subsample=0.7	245623.84	495.6	355.47	1.27	0.99	1.58	99.95	23.516872	1.05
91	Gradient Boosting: learning_rate=0.1, max_depth=3, n_estimators=50, subsample=0.8	230630.89	480.24	340.7	1.21	0.99	1.56	114.35	11.762571	1.03
92	Gradient Boosting: learning_rate=0.1, max_depth=3, n_estimators=50, subsample=1.0	242489.43	492.43	352.91	1.26	0.99	1.59	143.83	21.636576	1.05
93	Gradient Boosting: learning_rate=0.1, max_depth=3, n_estimators=100, subsample=0.7	219912.36	468.95	338.32	1.19	0.99	1.46	203.49	55.273062	1.13
94	Gradient Boosting: learning_rate=0.1, max_depth=3, n_estimators=100, subsample=0.8	228734.26	478.26	349.31	1.23	0.99	1.47	247.45	66.134598	1.16

Table 5.1: Results of the ablation study of the Machine learning architecture.

ID	Config.	MSE	RMSE	MAE	MAPE	R^2	CVRMSE	Training Time (s)	Overfit Gap	Loss Ratio
95	Gradient Boosting: learning_rate=0.1, max_depth=3, n_estimators=100, subsample=1.0	222713.04	471.92	344.19	1.22	0.99	1.48	290.58	61.225109	1.15
96	Gradient Boosting: learning_rate=0.1, max_depth=3, n_estimators=200, subsample=0.7	209865.96	458.11	330.7	1.16	0.99	1.43	411.6	85.471627	1.23
97	Gradient Boosting: learning_rate=0.1, max_depth=3, n_estimators=200, subsample=0.8	208006.25	456.08	327.13	1.14	0.99	1.41	468.07	79.509514	1.21
98	Gradient Boosting: learning_rate=0.1, max_depth=3, n_estimators=200, subsample=1.0	204400.01	452.11	328.95	1.16	0.99	1.43	583.41	81.585657	1.22
99	Gradient Boosting: learning_rate=0.1, max_depth=5, n_estimators=50, subsample=0.7	161769.73	402.21	285.06	1	1	1.31	158.63	42.316511	1.12
100	Gradient Boosting: learning_rate=0.1, max_depth=5, n_estimators=50, subsample=0.8	163486.24	404.33	288.67	1.01	1	1.33	181.83	45.607891	1.13
101	Gradient Boosting: learning_rate=0.1, max_depth=5, n_estimators=50, subsample=1.0	160477.82	400.6	286.34	1	1	1.3	227.92	45.264500	1.13
102	Gradient Boosting: learning_rate=0.1, max_depth=5, n_estimators=100, subsample=0.7	158221.79	397.77	283.76	0.99	1	1.29	324.98	72.869016	1.22
103	Gradient Boosting: learning_rate=0.1, max_depth=5, n_estimators=100, subsample=0.8	150568.08	388.03	275.59	0.95	1	1.26	372.19	66.049842	1.21
104	Gradient Boosting: learning_rate=0.1, max_depth=5, n_estimators=100, subsample=1.0	153345.71	391.59	277.08	0.95	1	1.26	467.95	70.826350	1.22
105	Gradient Boosting: learning_rate=0.1, max_depth=5, n_estimators=200, subsample=0.7	162343.78	402.92	285.58	0.99	1	1.32	664.75	105.238828	1.35
106	Gradient Boosting: learning_rate=0.1, max_depth=5, n_estimators=200, subsample=0.8	153085.4	391.26	280.08	0.97	1	1.28	764.57	99.419344	1.34
107	Gradient Boosting: learning_rate=0.1, max_depth=5, n_estimators=200, subsample=1.0	150343.05	387.74	273.53	0.94	1	1.27	959.99	93.267642	1.32
108	Gradient Boosting: learning_rate=0.1, max_depth=7, n_estimators=50, subsample=0.7	131377.87	362.46	255.67	0.9	1	1.21	212.7	78.442396	1.28

Table 5.1: Results of the ablation study of the Machine learning architecture.

ID	Config.	MSE	RMSE	MAE	MAPE	R²	CVRMSE	Training Time (s)	Overfit Gap	Loss Ratio
109	Gradient Boosting: learning_rate=0.1, max_depth=7, n_estimators=50, subsample=0.8	136219.14	369.08	259.16	0.91	1	1.23	242.84	87.598024	1.31
110	Gradient Boosting: learning_rate=0.1, max_depth=7, n_estimators=50, subsample=1.0	133422.15	365.27	256.21	0.9	1	1.22	306.42	85.556633	1.31
111	Gradient Boosting: learning_rate=0.1, max_depth=7, n_estimators=100, subsample=0.7	130675.73	361.49	253.63	0.88	1	1.21	442.19	107.118349	1.42
112	Gradient Boosting: learning_rate=0.1, max_depth=7, n_estimators=100, subsample=0.8	141560.33	376.25	263.75	0.93	1	1.29	508.52	123.492412	1.49
113	Gradient Boosting: learning_rate=0.1, max_depth=7, n_estimators=100, subsample=1.0	133698.02	365.65	255.23	0.9	1	1.24	636.38	117.666924	1.47
114	Gradient Boosting: learning_rate=0.1, max_depth=7, n_estimators=200, subsample=0.7	137305.8	370.55	260.63	0.91	1	1.27	915.43	143.758091	1.63
115	Gradient Boosting: learning_rate=0.1, max_depth=7, n_estimators=200, subsample=0.8	129290.16	359.57	254.01	0.89	1	1.21	1049.77	134.280218	1.6
116	Gradient Boosting: learning_rate=0.1, max_depth=7, n_estimators=200, subsample=1.0	133901.32	365.93	255.71	0.9	1	1.25	1328.25	139.773314	1.62
117	XGBoost: gamma=0, learning_rate=0.001, max_depth=3, n_estimators=50, subsample=0.7	46377423.23	6810.1	5719.34	22.32	-0.22	23.5	0.92	-18.572252	1
118	XGBoost: gamma=0, learning_rate=0.001, max_depth=3, n_estimators=50, subsample=0.8	46377438.74	6810.1	5719.31	22.32	-0.22	23.5	0.42	-18.549472	1
119	XGBoost: gamma=0, learning_rate=0.001, max_depth=3, n_estimators=50, subsample=1.0	46376202.29	6810.01	5719.26	22.32	-0.22	23.5	0.4	-18.625938	1
120	XGBoost: gamma=0, learning_rate=0.001, max_depth=3, n_estimators=100, subsample=0.7	42132163.09	6490.93	5448.27	21.27	-0.11	22.42	0.63	-15.856118	1
121	XGBoost: gamma=0, learning_rate=0.001, max_depth=3, n_estimators=100, subsample=0.8	42135314.46	6491.17	5448.39	21.27	-0.11	22.42	0.6	-15.578925	1
122	XGBoost: gamma=0, learning_rate=0.001, max_depth=3, n_estimators=100, subsample=1.0	42135198.8	6491.16	5448.3	21.27	-0.11	22.42	0.66	-15.554479	1

Table 5.1: Results of the ablation study of the Machine learning architecture.

ID	Config.	MSE	RMSE	MAE	MAPE	R^2	CVRMSE	Training Time (s)	Overfit Gap	Loss Ratio
123	XGBoost; gamma=0, learning_rate=0.001, max_depth=3, n_estimators=200, subsample=0.7	34822301.79	5901.04	4946.72	19.32	0.08	20.43	1.02	-8.245035	1
124	XGBoost; gamma=0, learning_rate=0.001, max_depth=3, n_estimators=200, subsample=0.8	34825344.05	5901.3	4946.8	19.32	0.08	20.43	1.02	-7.878553	1
125	XGBoost; gamma=0, learning_rate=0.001, max_depth=3, n_estimators=200, subsample=1.0	34827754.74	5901.5	4946.81	19.32	0.08	20.43	1	-7.587614	1
126	XGBoost; gamma=0, learning_rate=0.001, max_depth=5, n_estimators=50, subsample=0.7	46251683.66	6800.86	5715.44	22.3	-0.22	23.44	0.63	-20.471576	1
127	XGBoost; gamma=0, learning_rate=0.001, max_depth=5, n_estimators=50, subsample=0.8	46250699.65	6800.79	5715.42	22.3	-0.22	23.44	0.61	-20.502796	1
128	XGBoost; gamma=0, learning_rate=0.001, max_depth=5, n_estimators=50, subsample=1.0	46247328.94	6800.54	5715.18	22.3	-0.22	23.44	0.66	-20.697797	1
129	XGBoost; gamma=0, learning_rate=0.001, max_depth=5, n_estimators=100, subsample=0.7	41892493.6	6472.44	5440.74	21.23	-0.1	22.31	1.07	-20.007426	1
130	XGBoost; gamma=0, learning_rate=0.001, max_depth=5, n_estimators=100, subsample=0.8	41890915	6472.32	5440.68	21.23	-0.1	22.31	1.01	-20.043158	1
131	XGBoost; gamma=0, learning_rate=0.001, max_depth=5, n_estimators=100, subsample=1.0	41888858.16	6472.16	5440.59	21.23	-0.1	22.31	1.02	-20.102339	1
132	XGBoost; gamma=0, learning_rate=0.001, max_depth=5, n_estimators=200, subsample=0.7	34376909.66	5863.18	4931.06	19.24	0.1	20.2	2.71	-19.093021	1
133	XGBoost; gamma=0, learning_rate=0.001, max_depth=5, n_estimators=200, subsample=0.8	34374207.15	5862.95	4930.89	19.24	0.1	20.2	1.94	-19.141245	1
134	XGBoost; gamma=0, learning_rate=0.001, max_depth=5, n_estimators=200, subsample=1.0	34374238.45	5862.95	4931.1	19.24	0.1	20.19	1.82	-18.896324	1
135	XGBoost; gamma=0, learning_rate=0.001, max_depth=7, n_estimators=50, subsample=0.7	46239824.58	6799.99	5714.01	22.3	-0.22	23.44	1.06	-20.422145	1
136	XGBoost; gamma=0, learning_rate=0.001, max_depth=7, n_estimators=50, subsample=0.8	46238207.37	6799.87	5713.93	22.29	-0.22	23.44	1.06	-20.442522	1

Table 5.1: Results of the ablation study of the Machine learning architecture.

ID	Config.	MSE	RMSE	MAE	MAPE	R^2	CVRMSE	Training Time (s)	Overfit Gap	Loss Ratio
137	XGBoost; gamma=0, learning_rate=0.001, max_depth=7, n_estimators=50, subsample=1.0	46242483.74	6800.18	5713.99	22.3	-0.22	23.44	1.11	-19.981506	1
138	XGBoost; gamma=0, learning_rate=0.001, max_depth=7, n_estimators=100, subsample=0.7	41864503.77	6470.28	5437.59	21.22	-0.1	22.3	1.89	-20.280690	1
139	XGBoost; gamma=0, learning_rate=0.001, max_depth=7, n_estimators=100, subsample=0.8	41862803.83	6470.15	5437.46	21.22	-0.1	22.3	1.88	-20.231068	1
140	XGBoost; gamma=0, learning_rate=0.001, max_depth=7, n_estimators=100, subsample=1.0	41867102.09	6470.48	5437.61	21.22	-0.1	22.3	1.92	-19.624430	1
141	XGBoost; gamma=0, learning_rate=0.001, max_depth=7, n_estimators=200, subsample=0.7	34321408.35	5858.45	4924.45	19.22	0.1	20.19	3.58	-20.043108	1
142	XGBoost; gamma=0, learning_rate=0.001, max_depth=7, n_estimators=200, subsample=0.8	34320044.83	5858.33	4924.31	19.22	0.1	20.19	3.54	-19.785785	1
143	XGBoost; gamma=0, learning_rate=0.001, max_depth=7, n_estimators=200, subsample=1.0	34322699.07	5858.56	4923.96	19.22	0.1	20.19	3.7	-18.995445	1
144	XGBoost; gamma=0, learning_rate=0.01, max_depth=3, n_estimators=50, subsample=0.7	19666296.28	4434.67	3698.43	14.47	0.48	15.46	0.39	8.451086	1
145	XGBoost; gamma=0, learning_rate=0.01, max_depth=3, n_estimators=50, subsample=0.8	19647481.14	4432.55	3697	14.46	0.48	15.45	0.4	6.480059	1
146	XGBoost; gamma=0, learning_rate=0.01, max_depth=3, n_estimators=50, subsample=1.0	19681157.54	4436.35	3698.55	14.47	0.48	15.47	0.41	10.309046	1
147	XGBoost; gamma=0, learning_rate=0.01, max_depth=3, n_estimators=100, subsample=0.7	7705914.11	2775.95	2290.37	8.98	0.8	9.8	0.61	5.042698	1
148	XGBoost; gamma=0, learning_rate=0.01, max_depth=3, n_estimators=100, subsample=0.8	7710686.95	2776.81	2291.52	8.98	0.8	9.8	0.6	5.807418	1
149	XGBoost; gamma=0, learning_rate=0.01, max_depth=3, n_estimators=100, subsample=1.0	7708988.56	2776.51	2290.6	8.98	0.8	9.8	0.61	5.759622	1
150	XGBoost; gamma=0, learning_rate=0.01, max_depth=3, n_estimators=200, subsample=0.7	1417867.47	1190.74	956.73	3.73	0.96	4.18	1.96	-18.627706	0.98

Table 5.1: Results of the ablation study of the Machine learning architecture.

ID	Config.	MSE	RMSE	MAE	MAPE	R^2	CVRMSE	Training Time (s)	Overfit Gap	Loss Ratio
151	XGBoost; gamma=0, learning_rate=0.01, max_depth=3, n_estimators=200, subsample=0.8	1417431.86	1190.56	956.58	3.72	0.96	4.18	1.06	-18.837539	0.98
152	XGBoost; gamma=0, learning_rate=0.01, max_depth=3, n_estimators=200, subsample=1.0	1418742.66	1191.11	957	3.73	0.96	4.18	1.01	-18.100379	0.99
153	XGBoost; gamma=0, learning_rate=0.01, max_depth=5, n_estimators=50, subsample=0.7	18935015.99	4351.44	3664.04	14.29	0.5	14.96	0.6	-18.463796	1
154	XGBoost; gamma=0, learning_rate=0.01, max_depth=5, n_estimators=50, subsample=0.8	18933980.74	4351.32	3664.08	14.29	0.5	14.96	0.63	-18.254487	1
155	XGBoost; gamma=0, learning_rate=0.01, max_depth=5, n_estimators=50, subsample=1.0	18934824.65	4351.42	3664.44	14.29	0.5	14.96	0.62	-17.757541	1
156	XGBoost; gamma=0, learning_rate=0.01, max_depth=5, n_estimators=100, subsample=0.7	7098912.49	2664.38	2245.73	8.76	0.81	9.14	1.02	-18.742297	0.99
157	XGBoost; gamma=0, learning_rate=0.01, max_depth=5, n_estimators=100, subsample=0.8	7094694.92	2663.59	2244.21	8.75	0.81	9.13	1.02	-19.230075	0.99
158	XGBoost; gamma=0, learning_rate=0.01, max_depth=5, n_estimators=100, subsample=1.0	7098347.79	2664.27	2245.8	8.76	0.81	9.14	1.01	-17.975113	0.99
159	XGBoost; gamma=0, learning_rate=0.01, max_depth=5, n_estimators=200, subsample=0.7	1116174.35	1056.49	869.86	3.39	0.97	3.57	1.8	-21.074717	0.98
160	XGBoost; gamma=0, learning_rate=0.01, max_depth=5, n_estimators=200, subsample=0.8	1114230.85	1055.57	867.34	3.39	0.97	3.56	1.82	-21.714502	0.98
161	XGBoost; gamma=0, learning_rate=0.01, max_depth=5, n_estimators=200, subsample=1.0	1113514.59	1055.23	866.61	3.38	0.97	3.56	1.77	-21.579264	0.98
162	XGBoost; gamma=0, learning_rate=0.01, max_depth=7, n_estimators=50, subsample=0.7	18871005.54	4344.08	3653.82	14.27	0.5	14.96	1.08	-16.624806	1
163	XGBoost; gamma=0, learning_rate=0.01, max_depth=7, n_estimators=50, subsample=0.8	18872180.22	4344.21	3653.8	14.27	0.5	14.96	1.07	-15.672513	1
164	XGBoost; gamma=0, learning_rate=0.01, max_depth=7, n_estimators=50, subsample=1.0	18864218.75	4343.3	3652.17	14.26	0.5	14.96	1.07	-15.514100	1

Table 5.1: Results of the ablation study of the Machine learning architecture.

ID	Config.	MSE	RMSE	MAE	MAPE	R^2	CVRMSE	Training Time (s)	Overfit Gap	Loss Ratio
165	XGBoost; gamma=0, learning_rate=0.01, max_depth=7, n_estimators=100, subsample=0.7	7068466.96	2658.66	2237.64	8.74	0.81	9.14	1.98	-6.706627	1
166	XGBoost; gamma=0, learning_rate=0.01, max_depth=7, n_estimators=100, subsample=0.8	7070326.91	2659.01	2237.57	8.74	0.81	9.14	1.99	-5.293910	1
167	XGBoost; gamma=0, learning_rate=0.01, max_depth=7, n_estimators=100, subsample=1.0	7081937.11	2661.19	2239.3	8.75	0.81	9.15	2.09	-1.618217	1
168	XGBoost; gamma=0, learning_rate=0.01, max_depth=7, n_estimators=200, subsample=0.7	1133146.45	1064.49	884.38	3.45	0.97	3.64	3.78	25.978215	1.03
169	XGBoost; gamma=0, learning_rate=0.01, max_depth=7, n_estimators=200, subsample=0.8	1136027.57	1065.85	885.8	3.46	0.97	3.65	4.82	28.505839	1.03
170	XGBoost; gamma=0, learning_rate=0.01, max_depth=7, n_estimators=200, subsample=1.0	1135954.27	1065.81	885.89	3.46	0.97	3.65	3.8	29.931985	1.03
171	XGBoost; gamma=0, learning_rate=0.1, max_depth=3, n_estimators=50, subsample=0.7	246423.43	496.41	353.91	1.26	0.99	1.6	0.4	19.467531	1.04
172	XGBoost; gamma=0, learning_rate=0.1, max_depth=3, n_estimators=50, subsample=0.8	239886.84	489.78	345.4	1.21	0.99	1.56	0.39	11.818338	1.02
173	XGBoost; gamma=0, learning_rate=0.1, max_depth=3, n_estimators=50, subsample=1.0	234618.74	484.37	343.61	1.22	0.99	1.57	0.38	6.974202	1.01
174	XGBoost; gamma=0, learning_rate=0.1, max_depth=3, n_estimators=100, subsample=0.7	231046.41	480.67	346.92	1.22	0.99	1.51	0.55	60.112151	1.14
175	XGBoost; gamma=0, learning_rate=0.1, max_depth=3, n_estimators=100, subsample=0.8	227851.28	477.34	342.71	1.19	0.99	1.48	0.59	56.652358	1.13
176	XGBoost; gamma=0, learning_rate=0.1, max_depth=3, n_estimators=100, subsample=1.0	217122.25	465.96	341.46	1.21	0.99	1.48	0.57	48.219801	1.12
177	XGBoost; gamma=0, learning_rate=0.1, max_depth=3, n_estimators=200, subsample=0.7	204772.93	452.52	327.99	1.16	0.99	1.43	0.88	69.495350	1.18
178	XGBoost; gamma=0, learning_rate=0.1, max_depth=3, n_estimators=200, subsample=0.8	206477.53	454.4	329.81	1.15	0.99	1.42	0.89	71.982471	1.19

Table 5.1: Results of the ablation study of the Machine learning architecture.

ID	Config.	MSE	RMSE	MAE	MAPE	R^2	CVRMSE	Training Time (s)	Overfit Gap	Loss Ratio
179	XGBoost; gamma=0, learning_rate=0.1, max_depth=3, n_estimators=200, subsample=1.0	198571.28	445.61	325.09	1.15	0.99	1.42	0.88	63.930992	1.17
180	XGBoost; gamma=0, learning_rate=0.1, max_depth=5, n_estimators=50, subsample=0.7	161150.53	401.44	286.42	1	1	1.33	0.58	30.081291	1.08
181	XGBoost; gamma=0, learning_rate=0.1, max_depth=5, n_estimators=50, subsample=0.8	164122.04	405.12	289.74	1.02	1	1.34	0.6	33.222582	1.09
182	XGBoost; gamma=0, learning_rate=0.1, max_depth=5, n_estimators=50, subsample=1.0	159129.15	398.91	285.51	1	1	1.31	0.58	28.342860	1.08
183	XGBoost; gamma=0, learning_rate=0.1, max_depth=5, n_estimators=100, subsample=0.7	153836.18	392.22	280.41	0.97	1	1.29	0.89	57.534747	1.17
184	XGBoost; gamma=0, learning_rate=0.1, max_depth=5, n_estimators=100, subsample=0.8	155902.15	394.84	282	0.98	1	1.3	0.87	59.001192	1.18
185	XGBoost; gamma=0, learning_rate=0.1, max_depth=5, n_estimators=100, subsample=1.0	153915.49	392.32	280.85	0.97	1	1.28	0.86	59.825262	1.18
186	XGBoost; gamma=0, learning_rate=0.1, max_depth=5, n_estimators=200, subsample=0.7	149458.13	386.6	275.95	0.96	1	1.28	1.45	82.204145	1.27
187	XGBoost; gamma=0, learning_rate=0.1, max_depth=5, n_estimators=200, subsample=0.8	154769.02	393.41	281.08	0.98	1	1.3	1.43	90.743298	1.3
188	XGBoost; gamma=0, learning_rate=0.1, max_depth=5, n_estimators=200, subsample=1.0	149696.21	386.91	273.98	0.95	1	1.28	1.41	83.804228	1.28
189	XGBoost; gamma=0, learning_rate=0.1, max_depth=7, n_estimators=50, subsample=0.7	142512.46	377.51	264.71	0.94	1	1.28	1.09	74.577352	1.25
190	XGBoost; gamma=0, learning_rate=0.1, max_depth=7, n_estimators=50, subsample=0.8	142924.98	378.05	266.93	0.95	1	1.31	1.08	78.481240	1.26
191	XGBoost; gamma=0, learning_rate=0.1, max_depth=7, n_estimators=50, subsample=1.0	139218.4	373.12	260.17	0.92	1	1.27	1.08	76.405197	1.26
192	XGBoost; gamma=0, learning_rate=0.1, max_depth=7, n_estimators=100, subsample=0.7	135909.04	368.66	255.65	0.9	1	1.26	1.74	102.778209	1.39

Table 5.1: Results of the ablation study of the Machine learning architecture.

ID	Config.	MSE	RMSE	MAE	MAPE	R^2	CVRMSE	Training Time (s)	Overfit Gap	Loss Ratio
193	XGBoost; gamma=0, learning_rate=0.1, max_depth=7, n_estimators=100, subsample=0.8	137366.73	370.63	260.02	0.92	1	1.28	2.68	107.827440	1.41
194	XGBoost; gamma=0, learning_rate=0.1, max_depth=7, n_estimators=100, subsample=1.0	133542.83	365.44	251.9	0.88	1	1.26	1.77	104.634426	1.4
195	XGBoost; gamma=0, learning_rate=0.1, max_depth=7, n_estimators=200, subsample=0.7	136673.97	369.69	257.28	0.9	1	1.26	2.88	136.850338	1.59
196	XGBoost; gamma=0, learning_rate=0.1, max_depth=7, n_estimators=200, subsample=0.8	137564.29	370.9	260.64	0.92	1	1.28	2.87	141.399707	1.62
197	XGBoost; gamma=0, learning_rate=0.1, max_depth=7, n_estimators=200, subsample=1.0	134405.14	366.61	252.98	0.88	1	1.26	2.77	136.372939	1.59
198	XGBoost; gamma=0.1, learning_rate=0.001, max_depth=3, n_estimators=50, subsample=0.7	46377423.23	6810.1	5719.34	22.32	-0.22	23.5	0.4	-18.572252	1
199	XGBoost; gamma=0.1, learning_rate=0.001, max_depth=3, n_estimators=50, subsample=0.8	46377438.74	6810.1	5719.31	22.32	-0.22	23.5	0.41	-18.549472	1
200	XGBoost; gamma=0.1, learning_rate=0.001, max_depth=3, n_estimators=50, subsample=1.0	46376202.29	6810.01	5719.26	22.32	-0.22	23.5	0.39	-18.625938	1
201	XGBoost; gamma=0.1, learning_rate=0.001, max_depth=3, n_estimators=100, subsample=0.7	42132163.09	6490.93	5448.27	21.27	-0.11	22.42	0.6	-15.856118	1
202	XGBoost; gamma=0.1, learning_rate=0.001, max_depth=3, n_estimators=100, subsample=0.8	42135314.46	6491.17	5448.39	21.27	-0.11	22.42	0.6	-15.578925	1
203	XGBoost; gamma=0.1, learning_rate=0.001, max_depth=3, n_estimators=100, subsample=1.0	42135198.8	6491.16	5448.3	21.27	-0.11	22.42	0.63	-15.554479	1
204	XGBoost; gamma=0.1, learning_rate=0.001, max_depth=3, n_estimators=200, subsample=0.7	34822301.75	5901.04	4946.72	19.32	0.08	20.43	1.02	-8.245035	1
205	XGBoost; gamma=0.1, learning_rate=0.001, max_depth=3, n_estimators=200, subsample=0.8	34825344.05	5901.3	4946.8	19.32	0.08	20.43	1.01	-7.878553	1
206	XGBoost; gamma=0.1, learning_rate=0.001, max_depth=3, n_estimators=200, subsample=1.0	34827754.74	5901.5	4946.81	19.32	0.08	20.43	1.04	-7.587614	1

Table 5.1: Results of the ablation study of the Machine learning architecture.

ID	Config.	MSE	RMSE	MAE	MAPE	R^2	CVRMSE	Training Time (s)	Overfit Gap	Loss Ratio
207	XGBoost; gamma=0.1, learning_rate=0.001, max_depth=5, n_estimators=50, subsample=0.7	46251646.44	6800.86	5715.43	22.3	-0.22	23.44	0.61	-20.477036	1
208	XGBoost; gamma=0.1, learning_rate=0.001, max_depth=5, n_estimators=50, subsample=0.8	46250691.91	6800.79	5715.42	22.3	-0.22	23.44	0.61	-20.503898	1
209	XGBoost; gamma=0.1, learning_rate=0.001, max_depth=5, n_estimators=50, subsample=1.0	46247328.94	6800.54	5715.18	22.3	-0.22	23.44	0.63	-20.697797	1
210	XGBoost; gamma=0.1, learning_rate=0.001, max_depth=5, n_estimators=100, subsample=0.7	41892282.34	6472.42	5440.71	21.23	-0.1	22.31	1.14	-20.035027	1
211	XGBoost; gamma=0.1, learning_rate=0.001, max_depth=5, n_estimators=100, subsample=0.8	41891000.35	6472.33	5440.69	21.23	-0.1	22.31	1.11	-20.037673	1
212	XGBoost; gamma=0.1, learning_rate=0.001, max_depth=5, n_estimators=100, subsample=1.0	41888858.10	6472.16	5440.59	21.23	-0.1	22.31	1.07	-20.102339	1
213	XGBoost; gamma=0.1, learning_rate=0.001, max_depth=5, n_estimators=200, subsample=0.7	34376138.53	5863.12	4930.92	19.24	0.1	20.2	1.87	-19.195708	1
214	XGBoost; gamma=0.1, learning_rate=0.001, max_depth=5, n_estimators=200, subsample=0.8	34374083.8	5862.94	4930.86	19.24	0.1	20.2	2.77	-19.155851	1
215	XGBoost; gamma=0.1, learning_rate=0.001, max_depth=5, n_estimators=200, subsample=1.0	34374238.45	5862.95	4931.1	19.24	0.1	20.19	1.89	-18.896324	1
216	XGBoost; gamma=0.1, learning_rate=0.001, max_depth=7, n_estimators=50, subsample=0.7	46249134.6	6800.67	5715.17	22.3	-0.22	23.45	0.97	-20.250470	1
217	XGBoost; gamma=0.1, learning_rate=0.001, max_depth=7, n_estimators=50, subsample=0.8	46248769.13	6800.64	5715.14	22.3	-0.22	23.44	0.93	-20.128805	1
218	XGBoost; gamma=0.1, learning_rate=0.001, max_depth=7, n_estimators=50, subsample=1.0	46242150.99	6800.16	5714.64	22.3	-0.22	23.44	0.94	-20.401812	1
219	XGBoost; gamma=0.1, learning_rate=0.001, max_depth=7, n_estimators=100, subsample=0.7	41885669.47	6471.91	5440.12	21.23	-0.1	22.31	1.6	-19.739786	1
220	XGBoost; gamma=0.1, learning_rate=0.001, max_depth=7, n_estimators=100, subsample=0.8	41882261.7	6471.65	5439.74	21.22	-0.1	22.31	1.6	-19.711504	1

Table 5.1: Results of the ablation study of the Machine learning architecture.

ID	Config.	MSE	RMSE	MAE	MAPE	R^2	CVRMSE	Training Time (s)	Overfit Gap	Loss Ratio
221	XGBoost; gamma=0.1, learning_rate=0.001, max_depth=7, n_estimators=100, subsample=1.0	41870080.33	6470.71	5438.83	21.22	-0.1	22.3	1.64	-20.238807	1
222	XGBoost; gamma=0.1, learning_rate=0.001, max_depth=7, n_estimators=200, subsample=0.7	34359300.23	5861.68	4929.57	19.23	0.1	20.2	2.9	-19.142773	1
223	XGBoost; gamma=0.1, learning_rate=0.001, max_depth=7, n_estimators=200, subsample=0.8	34355933.71	5861.39	4929.13	19.23	0.1	20.19	2.97	-18.876707	1
224	XGBoost; gamma=0.1, learning_rate=0.001, max_depth=7, n_estimators=200, subsample=1.0	34343081.07	5860.3	4927.97	19.23	0.1	20.19	3.03	-19.115755	1
225	XGBoost; gamma=0.1, learning_rate=0.01, max_depth=3, n_estimators=50, subsample=0.7	19666296.28	4434.67	3698.43	14.47	0.48	15.46	0.39	8.451086	1
226	XGBoost; gamma=0.1, learning_rate=0.01, max_depth=3, n_estimators=50, subsample=0.8	19647481.14	4432.55	3697	14.46	0.48	15.45	0.39	6.480059	1
227	XGBoost; gamma=0.1, learning_rate=0.01, max_depth=3, n_estimators=50, subsample=1.0	19681157.54	4436.35	3698.55	14.47	0.48	15.47	0.39	10.309046	1
228	XGBoost; gamma=0.1, learning_rate=0.01, max_depth=3, n_estimators=100, subsample=0.7	7705914.11	2775.95	2290.37	8.98	0.8	9.8	0.63	5.042698	1
229	XGBoost; gamma=0.1, learning_rate=0.01, max_depth=3, n_estimators=100, subsample=0.8	7710686.95	2776.81	2291.52	8.98	0.8	9.8	0.62	5.807418	1
230	XGBoost; gamma=0.1, learning_rate=0.01, max_depth=3, n_estimators=100, subsample=1.0	7708988.56	2776.51	2290.6	8.98	0.8	9.8	0.62	5.759622	1
231	XGBoost; gamma=0.1, learning_rate=0.01, max_depth=3, n_estimators=200, subsample=0.7	1417867.47	1190.74	956.73	3.73	0.96	4.18	1.06	-18.627706	0.98
232	XGBoost; gamma=0.1, learning_rate=0.01, max_depth=3, n_estimators=200, subsample=0.8	1417431.86	1190.56	956.58	3.72	0.96	4.18	1.02	-18.837539	0.98
233	XGBoost; gamma=0.1, learning_rate=0.01, max_depth=3, n_estimators=200, subsample=1.0	1418742.66	1191.11	957	3.73	0.96	4.18	1.02	-18.100379	0.99
234	XGBoost; gamma=0.1, learning_rate=0.01, max_depth=5, n_estimators=50, subsample=0.7	18926746.29	4350.49	3661.66	14.29	0.5	14.97	0.62	-20.192375	1

Table 5.1: Results of the ablation study of the Machine learning architecture.

ID	Config.	MSE	RMSE	MAE	MAPE	R^2	CVRMSE	Training Time (s)	Overfit Gap	Loss Ratio
235	XGBoost; gamma=0.1, learning_rate=0.01, max_depth=5, n_estimators=50, subsample=0.8	18930165.94	4350.88	3662.98	14.29	0.5	14.97	0.62	-19.037527	1
236	XGBoost; gamma=0.1, learning_rate=0.01, max_depth=5, n_estimators=50, subsample=1.0	18935912.62	4351.54	3664.46	14.29	0.5	14.97	0.66	-17.687920	1
237	XGBoost; gamma=0.1, learning_rate=0.01, max_depth=5, n_estimators=100, subsample=0.7	7119047.2	2668.15	2244.56	8.76	0.81	9.17	1.06	-19.659681	0.99
238	XGBoost; gamma=0.1, learning_rate=0.01, max_depth=5, n_estimators=100, subsample=0.8	7107441.29	2665.98	2242.88	8.75	0.81	9.16	1.08	-20.338987	0.99
239	XGBoost; gamma=0.1, learning_rate=0.01, max_depth=5, n_estimators=100, subsample=1.0	7102290.97	2665.01	2244.16	8.75	0.81	9.15	1.03	-19.620134	0.99
240	XGBoost; gamma=0.1, learning_rate=0.01, max_depth=5, n_estimators=200, subsample=0.7	1163395.55	1078.61	884.9	3.45	0.97	3.64	1.76	-25.685012	0.98
241	XGBoost; gamma=0.1, learning_rate=0.01, max_depth=5, n_estimators=200, subsample=0.8	1157426.32	1075.84	883.04	3.44	0.97	3.63	1.82	-24.081997	0.98
242	XGBoost; gamma=0.1, learning_rate=0.01, max_depth=5, n_estimators=200, subsample=1.0	1145450.24	1070.26	878.63	3.43	0.97	3.6	1.74	-23.345004	0.98
243	XGBoost; gamma=0.1, learning_rate=0.01, max_depth=7, n_estimators=50, subsample=0.7	18890390.8	4346.31	3656.68	14.27	0.5	14.95	0.83	-21.407167	1
244	XGBoost; gamma=0.1, learning_rate=0.01, max_depth=7, n_estimators=50, subsample=0.8	18897938.09	4347.18	3658.32	14.28	0.5	14.96	0.91	-19.228515	1
245	XGBoost; gamma=0.1, learning_rate=0.01, max_depth=7, n_estimators=50, subsample=1.0	18902314.52	4347.68	3660.3	14.28	0.5	14.96	0.92	-16.925616	1
246	XGBoost; gamma=0.1, learning_rate=0.01, max_depth=7, n_estimators=100, subsample=0.7	7093068.54	2663.28	2237.31	8.74	0.81	9.14	1.37	-21.832961	0.99
247	XGBoost; gamma=0.1, learning_rate=0.01, max_depth=7, n_estimators=100, subsample=0.8	7078789.41	2660.6	2235.03	8.73	0.81	9.13	1.43	-21.988012	0.99
248	XGBoost; gamma=0.1, learning_rate=0.01, max_depth=7, n_estimators=100, subsample=1.0	7052200.29	2655.6	2230.97	8.72	0.81	9.11	1.49	-23.066939	0.99

Table 5.1: Results of the ablation study of the Machine learning architecture.

ID	Config.	MSE	RMSE	MAE	MAPE	R^2	CVRMSE	Training Time (s)	Overfit Gap	Loss Ratio
249	XGBoost: gamma=0.1, learning_rate=0.01, max_depth=7, n_estimators=200, subsample=0.7	1160918.64	1077.46	883.18	3.44	0.97	3.63	2.62	-25.467600	0.98
250	XGBoost: gamma=0.1, learning_rate=0.01, max_depth=7, n_estimators=200, subsample=0.8	1156044.4	1075.2	882.78	3.44	0.97	3.62	2.38	-22.908478	0.98
251	XGBoost: gamma=0.1, learning_rate=0.01, max_depth=7, n_estimators=200, subsample=1.0	1140176.31	1067.79	875.9	3.42	0.97	3.59	2.54	-22.784651	0.98
252	XGBoost: gamma=0.1, learning_rate=0.1, max_depth=3, n_estimators=50, subsample=0.7	249810.15	499.81	352.66	1.25	0.99	1.61	0.43	-5.344989	0.99
253	XGBoost: gamma=0.1, learning_rate=0.1, max_depth=3, n_estimators=50, subsample=0.8	248620.79	498.62	353.21	1.25	0.99	1.6	0.42	-0.078202	1
254	XGBoost: gamma=0.1, learning_rate=0.1, max_depth=3, n_estimators=50, subsample=1.0	246236.51	496.22	350.49	1.24	0.99	1.6	0.43	3.338862	1.01
255	XGBoost: gamma=0.1, learning_rate=0.1, max_depth=3, n_estimators=100, subsample=0.7	247266.8	497.26	349.76	1.24	0.99	1.6	0.49	-7.003853	0.99
256	XGBoost: gamma=0.1, learning_rate=0.1, max_depth=3, n_estimators=100, subsample=0.8	248794.79	498.79	355.52	1.26	0.99	1.6	0.5	2.690003	1.01
257	XGBoost: gamma=0.1, learning_rate=0.1, max_depth=3, n_estimators=100, subsample=1.0	246671.24	496.66	352.24	1.25	0.99	1.59	0.52	5.227642	1.01
258	XGBoost: gamma=0.1, learning_rate=0.1, max_depth=3, n_estimators=200, subsample=0.7	247343.91	497.34	349.87	1.24	0.99	1.6	0.73	-6.926169	0.99
259	XGBoost: gamma=0.1, learning_rate=0.1, max_depth=3, n_estimators=200, subsample=0.8	248678.02	498.68	355.35	1.26	0.99	1.6	0.71	2.572762	1.01
260	XGBoost: gamma=0.1, learning_rate=0.1, max_depth=3, n_estimators=200, subsample=1.0	246674.71	496.66	352.25	1.25	0.99	1.59	0.76	5.231110	1.01
261	XGBoost: gamma=0.1, learning_rate=0.1, max_depth=5, n_estimators=50, subsample=0.7	246961.9	496.95	361.38	1.28	0.99	1.57	0.58	23.018749	1.05
262	XGBoost: gamma=0.1, learning_rate=0.1, max_depth=5, n_estimators=50, subsample=0.8	234893.06	484.66	350.44	1.23	0.99	1.54	0.51	17.307546	1.04

Table 5.1: Results of the ablation study of the Machine learning architecture.

ID	Config.	MSE	RMSE	MAE	MAPE	R^2	CVRMSE	Training Time (s)	Overfit Gap	Loss Ratio
263	XGBoost; gamma=0.1, learning_rate=0.1, max_depth=5, n_estimators=50, subsample=1.0	236059.27	485.86	355.05	1.24	0.99	1.53	0.53	28.627266	1.06
264	XGBoost; gamma=0.1, learning_rate=0.1, max_depth=5, n_estimators=100, subsample=0.7	246704.71	496.69	361.02	1.27	0.99	1.57	0.64	22.759798	1.05
265	XGBoost; gamma=0.1, learning_rate=0.1, max_depth=5, n_estimators=100, subsample=0.8	232565.76	482.25	347.77	1.22	0.99	1.53	0.63	15.771287	1.03
266	XGBoost; gamma=0.1, learning_rate=0.1, max_depth=5, n_estimators=100, subsample=1.0	236066.07	485.87	355.06	1.24	0.99	1.53	0.6	28.634245	1.06
267	XGBoost; gamma=0.1, learning_rate=0.1, max_depth=5, n_estimators=200, subsample=0.7	244366	494.33	358.56	1.26	0.99	1.56	0.84	21.353447	1.05
268	XGBoost; gamma=0.1, learning_rate=0.1, max_depth=5, n_estimators=200, subsample=0.8	232440.12	482.12	347.59	1.22	0.99	1.53	0.81	15.640821	1.03
269	XGBoost; gamma=0.1, learning_rate=0.1, max_depth=5, n_estimators=200, subsample=1.0	236069.75	485.87	355.07	1.24	0.99	1.53	0.84	28.637919	1.06
270	XGBoost; gamma=0.1, learning_rate=0.1, max_depth=7, n_estimators=50, subsample=0.7	239311.66	489.19	353.71	1.24	0.99	1.56	0.56	17.114508	1.04
271	XGBoost; gamma=0.1, learning_rate=0.1, max_depth=7, n_estimators=50, subsample=0.8	242235.15	492.17	358.13	1.26	0.99	1.56	0.57	23.351680	1.05
272	XGBoost; gamma=0.1, learning_rate=0.1, max_depth=7, n_estimators=50, subsample=1.0	237311.97	487.15	354.15	1.24	0.99	1.54	0.58	30.144776	1.07
273	XGBoost; gamma=0.1, learning_rate=0.1, max_depth=7, n_estimators=100, subsample=0.7	236613.64	486.43	350.69	1.23	0.99	1.55	0.68	15.291527	1.03
274	XGBoost; gamma=0.1, learning_rate=0.1, max_depth=7, n_estimators=100, subsample=0.8	242101.95	492.04	357.95	1.26	0.99	1.56	0.7	23.216435	1.05
275	XGBoost; gamma=0.1, learning_rate=0.1, max_depth=7, n_estimators=100, subsample=1.0	237322.19	487.16	354.16	1.24	0.99	1.54	0.75	30.155271	1.07
276	XGBoost; gamma=0.1, learning_rate=0.1, max_depth=7, n_estimators=200, subsample=0.7	236633.53	486.45	350.72	1.23	0.99	1.55	0.89	15.312016	1.03

Table 5.1: Results of the ablation study of the Machine learning architecture.

ID	Config.	MSE	RMSE	MAE	MAPE	R^2	CVRMSE	Training Time (s)	Overfit Gap	Loss Ratio
277	XGBoost; gamma=0.1, learning_rate=0.1, max_depth=7, n_estimators=200, subsample=0.8	241958.45	491.89	357.75	1.26	0.99	1.56	0.9	23.070425	1.05
278	XGBoost; gamma=0.1, learning_rate=0.1, max_depth=7, n_estimators=200, subsample=1.0	237326.57	487.16	354.17	1.24	0.99	1.54	0.94	30.159722	1.07
279	XGBoost; gamma=0.5, learning_rate=0.001, max_depth=3, n_estimators=50, subsample=0.7	46377423.23	6810.1	5719.34	22.32	-0.22	23.5	0.42	-18.572252	1
280	XGBoost; gamma=0.5, learning_rate=0.001, max_depth=3, n_estimators=50, subsample=0.8	46377438.74	6810.1	5719.31	22.32	-0.22	23.5	0.42	-18.549472	1
281	XGBoost; gamma=0.5, learning_rate=0.001, max_depth=3, n_estimators=50, subsample=1.0	46376202.29	6810.01	5719.26	22.32	-0.22	23.5	0.4	-18.625938	1
282	XGBoost; gamma=0.5, learning_rate=0.001, max_depth=3, n_estimators=100, subsample=0.7	42132163.09	6490.93	5448.27	21.27	-0.11	22.42	0.6	-15.856118	1
283	XGBoost; gamma=0.5, learning_rate=0.001, max_depth=3, n_estimators=100, subsample=0.8	42135314.46	6491.17	5448.39	21.27	-0.11	22.42	0.63	-15.578925	1
284	XGBoost; gamma=0.5, learning_rate=0.001, max_depth=3, n_estimators=100, subsample=1.0	42135198.8	6491.16	5448.3	21.27	-0.11	22.42	0.63	-15.554479	1
285	XGBoost; gamma=0.5, learning_rate=0.001, max_depth=3, n_estimators=200, subsample=0.7	34822301.79	5901.04	4946.72	19.32	0.08	20.43	1.05	-8.245035	1
286	XGBoost; gamma=0.5, learning_rate=0.001, max_depth=3, n_estimators=200, subsample=0.8	34825344.09	5901.3	4946.8	19.32	0.08	20.43	1.04	-7.878553	1
287	XGBoost; gamma=0.5, learning_rate=0.001, max_depth=3, n_estimators=200, subsample=1.0	34827754.74	5901.5	4946.81	19.32	0.08	20.43	1.15	-7.587614	1
288	XGBoost; gamma=0.5, learning_rate=0.001, max_depth=5, n_estimators=50, subsample=0.7	46271192.08	6802.29	5714.77	22.3	-0.22	23.46	0.61	-20.492637	1
289	XGBoost; gamma=0.5, learning_rate=0.001, max_depth=5, n_estimators=50, subsample=0.8	46272321.05	6802.38	5714.87	22.3	-0.22	23.46	0.63	-20.327532	1
290	XGBoost; gamma=0.5, learning_rate=0.001, max_depth=5, n_estimators=50, subsample=1.0	46259746.8	6801.45	5714.18	22.3	-0.22	23.45	0.64	-20.804088	1

Table 5.1: Results of the ablation study of the Machine learning architecture.

ID	Config.	MSE	RMSE	MAE	MAPE	R^2	CVRMSE	Training Time (s)	Overfit Gap	Loss Ratio
291	XGBoost; gamma=0.5, learning_rate=0.001, max_depth=5, n_estimators=100, subsample=0.7	41922165.65	6474.73	5438.92	21.22	-0.1	22.33	1.06	-20.530021	1
292	XGBoost; gamma=0.5, learning_rate=0.001, max_depth=5, n_estimators=100, subsample=0.8	41926609.04	6475.08	5439.23	21.22	-0.1	22.33	1.96	-20.000141	1
293	XGBoost; gamma=0.5, learning_rate=0.001, max_depth=5, n_estimators=100, subsample=1.0	41902769.76	6473.23	5437.75	21.22	-0.1	22.32	1.08	-21.074220	1
294	XGBoost; gamma=0.5, learning_rate=0.001, max_depth=5, n_estimators=200, subsample=0.7	34426459.64	5867.41	4927.36	19.23	0.1	20.25	1.93	-20.159756	1
295	XGBoost; gamma=0.5, learning_rate=0.001, max_depth=5, n_estimators=200, subsample=0.8	34433817.65	5868.03	4927.9	19.23	0.1	20.25	1.88	-19.196271	1
296	XGBoost; gamma=0.5, learning_rate=0.001, max_depth=5, n_estimators=200, subsample=1.0	34409036.04	5865.92	4926.34	19.22	0.1	20.24	1.83	-20.063315	1
297	XGBoost; gamma=0.5, learning_rate=0.001, max_depth=7, n_estimators=50, subsample=0.7	46271192.08	6802.29	5714.77	22.3	-0.22	23.46	0.64	-20.492637	1
298	XGBoost; gamma=0.5, learning_rate=0.001, max_depth=7, n_estimators=50, subsample=0.8	46272321.05	6802.38	5714.87	22.3	-0.22	23.46	0.65	-20.327532	1
299	XGBoost; gamma=0.5, learning_rate=0.001, max_depth=7, n_estimators=50, subsample=1.0	46259746.8	6801.45	5714.18	22.3	-0.22	23.45	0.75	-20.804088	1
300	XGBoost; gamma=0.5, learning_rate=0.001, max_depth=7, n_estimators=100, subsample=0.7	41922165.65	6474.73	5438.92	21.22	-0.1	22.33	1.05	-20.530021	1
301	XGBoost; gamma=0.5, learning_rate=0.001, max_depth=7, n_estimators=100, subsample=0.8	41926609.04	6475.08	5439.23	21.22	-0.1	22.33	1.08	-20.000141	1
302	XGBoost; gamma=0.5, learning_rate=0.001, max_depth=7, n_estimators=100, subsample=1.0	41902769.76	6473.23	5437.75	21.22	-0.1	22.32	1.14	-21.074220	1
303	XGBoost; gamma=0.5, learning_rate=0.001, max_depth=7, n_estimators=200, subsample=0.7	34426459.64	5867.41	4927.36	19.23	0.1	20.25	1.99	-20.159756	1
304	XGBoost; gamma=0.5, learning_rate=0.001, max_depth=7, n_estimators=200, subsample=0.8	34433817.65	5868.03	4927.9	19.23	0.1	20.25	1.98	-19.196271	1

Table 5.1: Results of the ablation study of the Machine learning architecture.

ID	Config.	MSE	RMSE	MAE	MAPE	R^2	CVRMSE	Training Time (s)	Overfit Gap	Loss Ratio
305	XGBoost; gamma=0.5, learning_rate=0.001, max_depth=7, n_estimators=200, subsample=1.0	34409036.04	5865.92	4926.34	19.22	0.1	20.24	2.05	-20.063315	1
306	XGBoost; gamma=0.5, learning_rate=0.01, max_depth=3, n_estimators=50, subsample=0.7	19666296.28	4434.67	3698.43	14.47	0.48	15.46	0.42	8.451086	1
307	XGBoost; gamma=0.5, learning_rate=0.01, max_depth=3, n_estimators=50, subsample=0.8	19647481.14	4432.55	3697	14.46	0.48	15.45	0.42	6.480059	1
308	XGBoost; gamma=0.5, learning_rate=0.01, max_depth=3, n_estimators=50, subsample=1.0	19681157.54	4436.35	3698.55	14.47	0.48	15.47	0.41	10.309046	1
309	XGBoost; gamma=0.5, learning_rate=0.01, max_depth=3, n_estimators=100, subsample=0.7	7705914.11	2775.95	2290.37	8.98	0.8	9.8	0.61	5.042698	1
310	XGBoost; gamma=0.5, learning_rate=0.01, max_depth=3, n_estimators=100, subsample=0.8	7710686.95	2776.81	2291.52	8.98	0.8	9.8	0.63	5.807418	1
311	XGBoost; gamma=0.5, learning_rate=0.01, max_depth=3, n_estimators=100, subsample=1.0	7708988.56	2776.51	2290.6	8.98	0.8	9.8	0.63	5.759622	1
312	XGBoost; gamma=0.5, learning_rate=0.01, max_depth=3, n_estimators=200, subsample=0.7	1420569.29	1191.88	954.26	3.72	0.96	4.2	1.04	-21.385434	0.98
313	XGBoost; gamma=0.5, learning_rate=0.01, max_depth=3, n_estimators=200, subsample=0.8	1416307.71	1190.09	955.18	3.72	0.96	4.18	1.05	-20.998093	0.98
314	XGBoost; gamma=0.5, learning_rate=0.01, max_depth=3, n_estimators=200, subsample=1.0	1418544.45	1191.03	956.89	3.73	0.96	4.18	1.04	-18.617856	0.98
315	XGBoost; gamma=0.5, learning_rate=0.01, max_depth=5, n_estimators=50, subsample=0.7	19064801.04	4366.33	3665.24	14.3	0.5	15.08	0.65	-17.164667	1
316	XGBoost; gamma=0.5, learning_rate=0.01, max_depth=5, n_estimators=50, subsample=0.8	19058907.42	4365.65	3664.07	14.3	0.5	15.07	1.57	-16.566855	1
317	XGBoost; gamma=0.5, learning_rate=0.01, max_depth=5, n_estimators=50, subsample=1.0	19033740.19	4362.77	3662.28	14.29	0.5	15.06	0.66	-17.892872	1
318	XGBoost; gamma=0.5, learning_rate=0.01, max_depth=5, n_estimators=100, subsample=0.7	7275643.78	2697.34	2254.95	8.8	0.81	9.35	1.02	-23.590356	0.99

Table 5.1: Results of the ablation study of the Machine learning architecture.

ID	Config.	MSE	RMSE	MAE	MAPE	R^2	CVRMSE	Training Time (s)	Overfit Gap	Loss Ratio
319	XGBoost; gamma=0.5, learning_rate=0.01, max_depth=5, n_estimators=100, subsample=0.8	7264129	2695.2	2256.63	8.8	0.81	9.32	1.01	-22.184563	0.99
320	XGBoost; gamma=0.5, learning_rate=0.01, max_depth=5, n_estimators=100, subsample=1.0	7236700.67	2690.11	2255.55	8.8	0.81	9.28	1.03	-21.098389	0.99
321	XGBoost; gamma=0.5, learning_rate=0.01, max_depth=5, n_estimators=200, subsample=0.7	1322026.14	1149.79	924.35	3.6	0.97	4.01	1.59	-33.861982	0.97
322	XGBoost; gamma=0.5, learning_rate=0.01, max_depth=5, n_estimators=200, subsample=0.8	1306245.04	1142.91	920.9	3.58	0.97	3.97	1.57	-33.386408	0.97
323	XGBoost; gamma=0.5, learning_rate=0.01, max_depth=5, n_estimators=200, subsample=1.0	1278248.37	1130.6	916.1	3.56	0.97	3.9	1.54	-33.185714	0.97
324	XGBoost; gamma=0.5, learning_rate=0.01, max_depth=7, n_estimators=50, subsample=0.7	19064801.04	4366.33	3665.24	14.3	0.5	15.08	0.62	-17.164667	1
325	XGBoost; gamma=0.5, learning_rate=0.01, max_depth=7, n_estimators=50, subsample=0.8	19058907.42	4365.65	3664.07	14.3	0.5	15.07	0.63	-16.566855	1
326	XGBoost; gamma=0.5, learning_rate=0.01, max_depth=7, n_estimators=50, subsample=1.0	19033740.14	4362.77	3662.28	14.29	0.5	15.06	0.67	-17.892872	1
327	XGBoost; gamma=0.5, learning_rate=0.01, max_depth=7, n_estimators=100, subsample=0.7	7275643.78	2697.34	2254.95	8.8	0.81	9.35	1	-23.590356	0.99
328	XGBoost; gamma=0.5, learning_rate=0.01, max_depth=7, n_estimators=100, subsample=0.8	7264129	2695.2	2256.63	8.8	0.81	9.32	1.02	-22.184563	0.99
329	XGBoost; gamma=0.5, learning_rate=0.01, max_depth=7, n_estimators=100, subsample=1.0	7236700.67	2690.11	2255.55	8.8	0.81	9.28	1.05	-21.098389	0.99
330	XGBoost; gamma=0.5, learning_rate=0.01, max_depth=7, n_estimators=200, subsample=0.7	1322026.14	1149.79	924.35	3.6	0.97	4.01	1.59	-33.861982	0.97
331	XGBoost; gamma=0.5, learning_rate=0.01, max_depth=7, n_estimators=200, subsample=0.8	1306245.04	1142.91	920.9	3.58	0.97	3.97	1.63	-33.386408	0.97
332	XGBoost; gamma=0.5, learning_rate=0.01, max_depth=7, n_estimators=200, subsample=1.0	1278248.37	1130.6	916.1	3.56	0.97	3.9	1.68	-33.185714	0.97

Table 5.1: Results of the ablation study of the Machine learning architecture.

ID	Config.	MSE	RMSE	MAE	MAPE	R^2	CVRMSE	Training Time (s)	Overfit Gap	Loss Ratio
333	XGBoost; gamma=0.5, learning_rate=0.1, max_depth=3, n_estimators=50, subsample=0.7	335652.14	579.35	411.75	1.51	0.99	1.97	0.39	-37.425675	0.94
334	XGBoost; gamma=0.5, learning_rate=0.1, max_depth=3, n_estimators=50, subsample=0.8	315679.17	561.85	395.77	1.44	0.99	1.9	0.39	-44.366390	0.93
335	XGBoost; gamma=0.5, learning_rate=0.1, max_depth=3, n_estimators=50, subsample=1.0	305129.8	552.39	389.2	1.41	0.99	1.87	0.36	-30.170200	0.95
336	XGBoost; gamma=0.5, learning_rate=0.1, max_depth=3, n_estimators=100, subsample=0.7	329307.44	573.85	406.35	1.49	0.99	1.95	0.47	-39.147266	0.94
337	XGBoost; gamma=0.5, learning_rate=0.1, max_depth=3, n_estimators=100, subsample=0.8	312820.68	559.3	393.02	1.43	0.99	1.89	0.46	-43.543486	0.93
338	XGBoost; gamma=0.5, learning_rate=0.1, max_depth=3, n_estimators=100, subsample=1.0	305134.78	552.39	389.21	1.41	0.99	1.87	0.45	-30.165519	0.95
339	XGBoost; gamma=0.5, learning_rate=0.1, max_depth=3, n_estimators=200, subsample=0.7	329460.73	573.99	406.54	1.49	0.99	1.95	0.63	-39.013616	0.94
340	XGBoost; gamma=0.5, learning_rate=0.1, max_depth=3, n_estimators=200, subsample=0.8	312681.97	559.18	392.83	1.43	0.99	1.89	0.66	-43.667619	0.93
341	XGBoost; gamma=0.5, learning_rate=0.1, max_depth=3, n_estimators=200, subsample=1.0	305137.22	552.39	389.22	1.41	0.99	1.87	0.67	-30.163214	0.95
342	XGBoost; gamma=0.5, learning_rate=0.1, max_depth=5, n_estimators=50, subsample=0.7	334162.18	578.07	415.87	1.52	0.99	1.95	0.43	-30.391276	0.95
343	XGBoost; gamma=0.5, learning_rate=0.1, max_depth=5, n_estimators=50, subsample=0.8	317987.52	563.9	403.77	1.46	0.99	1.89	0.42	-31.245274	0.95
344	XGBoost; gamma=0.5, learning_rate=0.1, max_depth=5, n_estimators=50, subsample=1.0	308973.25	555.85	398.91	1.43	0.99	1.82	0.42	-19.325209	0.97
345	XGBoost; gamma=0.5, learning_rate=0.1, max_depth=5, n_estimators=100, subsample=0.7	327601.17	572.36	413.97	1.51	0.99	1.91	0.5	-32.688035	0.95
346	XGBoost; gamma=0.5, learning_rate=0.1, max_depth=5, n_estimators=100, subsample=0.8	317939.08	563.86	403.71	1.46	0.99	1.89	1.5	-31.288223	0.95

Table 5.1: Results of the ablation study of the Machine learning architecture.

ID	Config.	MSE	RMSE	MAE	MAPE	R^2	CVRMSE	Training Time (s)	Overfit Gap	Loss Ratio
347	XGBoost: gamma=0.5, learning_rate=0.1, max_depth=5, n_estimators=100, subsample=1.0	308979.62	555.86	398.91	1.43	0.99	1.82	0.55	-19.319544	0.97
348	XGBoost: gamma=0.5, learning_rate=0.1, max_depth=5, n_estimators=200, subsample=0.7	325082.36	570.16	411.21	1.5	0.99	1.9	0.72	-31.118747	0.95
349	XGBoost: gamma=0.5, learning_rate=0.1, max_depth=5, n_estimators=200, subsample=0.8	317745.23	563.69	403.47	1.46	0.99	1.89	0.75	-31.460312	0.95
350	XGBoost: gamma=0.5, learning_rate=0.1, max_depth=5, n_estimators=200, subsample=1.0	308983.13	555.86	398.92	1.43	0.99	1.82	0.73	-19.316470	0.97
351	XGBoost: gamma=0.5, learning_rate=0.1, max_depth=7, n_estimators=50, subsample=0.7	334162.18	578.07	415.87	1.52	0.99	1.95	0.45	-30.391276	0.95
352	XGBoost: gamma=0.5, learning_rate=0.1, max_depth=7, n_estimators=50, subsample=0.8	317987.52	563.9	403.77	1.46	0.99	1.89	0.43	-31.245274	0.95
353	XGBoost: gamma=0.5, learning_rate=0.1, max_depth=7, n_estimators=50, subsample=1.0	308973.25	555.85	398.91	1.43	0.99	1.82	0.45	-19.325209	0.97
354	XGBoost: gamma=0.5, learning_rate=0.1, max_depth=7, n_estimators=100, subsample=0.7	327601.17	572.36	413.97	1.51	0.99	1.91	0.57	-32.688035	0.95
355	XGBoost: gamma=0.5, learning_rate=0.1, max_depth=7, n_estimators=100, subsample=0.8	317939.08	563.86	403.71	1.46	0.99	1.89	0.55	-31.288223	0.95
356	XGBoost: gamma=0.5, learning_rate=0.1, max_depth=7, n_estimators=100, subsample=1.0	308979.62	555.86	398.91	1.43	0.99	1.82	0.57	-19.319544	0.97
357	XGBoost: gamma=0.5, learning_rate=0.1, max_depth=7, n_estimators=200, subsample=0.7	325082.36	570.16	411.21	1.5	0.99	1.9	0.74	-31.118747	0.95
358	XGBoost: gamma=0.5, learning_rate=0.1, max_depth=7, n_estimators=200, subsample=0.8	317745.23	563.69	403.47	1.46	0.99	1.89	0.75	-31.460312	0.95
359	XGBoost: gamma=0.5, learning_rate=0.1, max_depth=7, n_estimators=200, subsample=1.0	308983.13	555.86	398.92	1.43	0.99	1.82	0.77	-19.316470	0.97
360	XGBoost: gamma=1, learning_rate=0.001, max_depth=3, n_estimators=50, subsample=0.7	46377423.2[illegible]	6810.1	5719.34	22.32	-0.22	23.5	0.41	-18.572252	1

Table 5.1: Results of the ablation study of the Machine learning architecture.

ID	Config.	MSE	RMSE	MAE	MAPE	R^2	CVRMSE	Training Time (s)	Overfit Gap	Loss Ratio
361	XGBoost; gamma=1, learning_rate=0.001, max_depth=3, n_estimators=50, subsample=0.8	46377438.74	6810.1	5719.31	22.32	-0.22	23.5	0.4	-18.549472	1
362	XGBoost; gamma=1, learning_rate=0.001, max_depth=3, n_estimators=50, subsample=1.0	46376202.29	6810.01	5719.26	22.32	-0.22	23.5	0.39	-18.625938	1
363	XGBoost; gamma=1, learning_rate=0.001, max_depth=3, n_estimators=100, subsample=0.7	42132163.09	6490.93	5448.27	21.27	-0.11	22.42	0.6	-15.856118	1
364	XGBoost; gamma=1, learning_rate=0.001, max_depth=3, n_estimators=100, subsample=0.8	42135314.44	6491.17	5448.39	21.27	-0.11	22.42	0.62	-15.578925	1
365	XGBoost; gamma=1, learning_rate=0.001, max_depth=3, n_estimators=100, subsample=1.0	42135198.8	6491.16	5448.3	21.27	-0.11	22.42	0.62	-15.554479	1
366	XGBoost; gamma=1, learning_rate=0.001, max_depth=3, n_estimators=200, subsample=0.7	34822301.79	5901.04	4946.72	19.32	0.08	20.43	1.03	-8.245035	1
367	XGBoost; gamma=1, learning_rate=0.001, max_depth=3, n_estimators=200, subsample=0.8	34825344.05	5901.3	4946.8	19.32	0.08	20.43	1	-7.878553	1
368	XGBoost; gamma=1, learning_rate=0.001, max_depth=3, n_estimators=200, subsample=1.0	34827754.74	5901.5	4946.81	19.32	0.08	20.43	1.03	-7.587614	1
369	XGBoost; gamma=1, learning_rate=0.001, max_depth=5, n_estimators=50, subsample=0.7	46269980.1	6802.2	5714.99	22.3	-0.22	23.46	0.62	-20.793185	1
370	XGBoost; gamma=1, learning_rate=0.001, max_depth=5, n_estimators=50, subsample=0.8	46264758.31	6801.82	5714.46	22.3	-0.22	23.46	0.62	-21.046813	1
371	XGBoost; gamma=1, learning_rate=0.001, max_depth=5, n_estimators=50, subsample=1.0	46266891	6801.98	5714.6	22.3	-0.22	23.46	0.61	-20.854686	1
372	XGBoost; gamma=1, learning_rate=0.001, max_depth=5, n_estimators=100, subsample=0.7	41923719.07	6474.85	5439.77	21.22	-0.1	22.33	1.01	-21.061268	1
373	XGBoost; gamma=1, learning_rate=0.001, max_depth=5, n_estimators=100, subsample=0.8	41916075.68	6474.26	5438.89	21.22	-0.1	22.33	1.02	-21.332800	1
374	XGBoost; gamma=1, learning_rate=0.001, max_depth=5, n_estimators=100, subsample=1.0	41916612.41	6474.3	5438.52	21.22	-0.1	22.33	1.03	-21.111167	1

Table 5.1: Results of the ablation study of the Machine learning architecture.

ID	Config.	MSE	RMSE	MAE	MAPE	R^2	CVRMSE	Training Time (s)	Overfit Gap	Loss Ratio
375	XGBoost; gamma=1, learning_rate=0.001, max_depth=5, n_estimators=200, subsample=0.7	34440386.96	5868.59	4929.15	19.23	0.09	20.25	1.78	-20.844060	1
376	XGBoost; gamma=1, learning_rate=0.001, max_depth=5, n_estimators=200, subsample=0.8	34429255.01	5867.64	4928.4	19.23	0.1	20.25	1.81	-21.088303	1
377	XGBoost; gamma=1, learning_rate=0.001, max_depth=5, n_estimators=200, subsample=1.0	34422179.75	5867.04	4927.2	19.23	0.1	20.25	2.72	-21.037104	1
378	XGBoost; gamma=1, learning_rate=0.001, max_depth=7, n_estimators=50, subsample=0.7	46269980.1	6802.2	5714.99	22.3	-0.22	23.46	0.63	-20.793185	1
379	XGBoost; gamma=1, learning_rate=0.001, max_depth=7, n_estimators=50, subsample=0.8	46264758.31	6801.82	5714.46	22.3	-0.22	23.46	0.62	-21.046813	1
380	XGBoost; gamma=1, learning_rate=0.001, max_depth=7, n_estimators=50, subsample=1.0	46266891	6801.98	5714.6	22.3	-0.22	23.46	0.62	-20.854686	1
381	XGBoost; gamma=1, learning_rate=0.001, max_depth=7, n_estimators=100, subsample=0.7	41923719.07	6474.85	5439.77	21.22	-0.1	22.33	1.01	-21.061268	1
382	XGBoost; gamma=1, learning_rate=0.001, max_depth=7, n_estimators=100, subsample=0.8	41916075.68	6474.26	5438.89	21.22	-0.1	22.33	1.02	-21.332800	1
383	XGBoost; gamma=1, learning_rate=0.001, max_depth=7, n_estimators=100, subsample=1.0	41916612.41	6474.3	5438.52	21.22	-0.1	22.33	1.02	-21.111167	1
384	XGBoost; gamma=1, learning_rate=0.001, max_depth=7, n_estimators=200, subsample=0.7	34440386.96	5868.59	4929.15	19.23	0.09	20.25	1.78	-20.844060	1
385	XGBoost; gamma=1, learning_rate=0.001, max_depth=7, n_estimators=200, subsample=0.8	34429255.01	5867.64	4928.4	19.23	0.1	20.25	1.82	-21.088303	1
386	XGBoost; gamma=1, learning_rate=0.001, max_depth=7, n_estimators=200, subsample=1.0	34422179.75	5867.04	4927.2	19.23	0.1	20.25	1.83	-21.037104	1
387	XGBoost; gamma=1, learning_rate=0.01, max_depth=3, n_estimators=50, subsample=0.7	19666296.28	4434.67	3698.43	14.47	0.48	15.46	0.41	8.451086	1
388	XGBoost; gamma=1, learning_rate=0.01, max_depth=3, n_estimators=50, subsample=0.8	19647481.14	4432.55	3697	14.46	0.48	15.45	0.39	6.480059	1

Table 5.1: Results of the ablation study of the Machine learning architecture.

ID	Config.	MSE	RMSE	MAE	MAPE	R^2	CVRMSE	Training Time (s)	Overfit Gap	Loss Ratio
389	XGBoost; gamma=1, learning_rate=0.01, max_depth=3, n_estimators=50, subsample=1.0	19681157.54	4436.35	3698.55	14.47	0.48	15.47	0.4	10.309046	1
390	XGBoost; gamma=1, learning_rate=0.01, max_depth=3, n_estimators=100, subsample=0.7	7705914.11	2775.95	2290.37	8.98	0.8	9.8	0.61	5.042698	1
391	XGBoost; gamma=1, learning_rate=0.01, max_depth=3, n_estimators=100, subsample=0.8	7710686.95	2776.81	2291.52	8.98	0.8	9.8	0.61	5.807418	1
392	XGBoost; gamma=1, learning_rate=0.01, max_depth=3, n_estimators=100, subsample=1.0	7708988.56	2776.51	2290.6	8.98	0.8	9.8	0.62	5.759622	1
393	XGBoost; gamma=1, learning_rate=0.01, max_depth=3, n_estimators=200, subsample=0.7	1501368.39	1225.3	969.16	3.79	0.96	4.42	1.01	-18.830883	0.98
394	XGBoost; gamma=1, learning_rate=0.01, max_depth=3, n_estimators=200, subsample=0.8	1470544.99	1212.66	964.25	3.76	0.96	4.34	1.01	-24.836076	0.98
395	XGBoost; gamma=1, learning_rate=0.01, max_depth=3, n_estimators=200, subsample=1.0	1442083.53	1200.87	956.27	3.73	0.96	4.28	1.02	-26.018782	0.98
396	XGBoost; gamma=1, learning_rate=0.01, max_depth=5, n_estimators=50, subsample=0.7	19134864.96	4374.34	3667.73	14.31	0.5	15.13	0.56	-17.693049	1
397	XGBoost; gamma=1, learning_rate=0.01, max_depth=5, n_estimators=50, subsample=0.8	19100721.84	4370.44	3665.18	14.31	0.5	15.12	0.6	-19.164637	1
398	XGBoost; gamma=1, learning_rate=0.01, max_depth=5, n_estimators=50, subsample=1.0	19075841.93	4367.59	3662.45	14.3	0.5	15.1	0.6	-18.810314	1
399	XGBoost; gamma=1, learning_rate=0.01, max_depth=5, n_estimators=100, subsample=0.7	7376792.54	2716.03	2262.21	8.83	0.81	9.47	0.89	-22.911903	0.99
400	XGBoost; gamma=1, learning_rate=0.01, max_depth=5, n_estimators=100, subsample=0.8	7364929.56	2713.84	2262.01	8.83	0.81	9.45	0.91	-22.188066	0.99
401	XGBoost; gamma=1, learning_rate=0.01, max_depth=5, n_estimators=100, subsample=1.0	7327140.78	2706.87	2258.26	8.81	0.81	9.41	0.92	-23.810163	0.99
402	XGBoost; gamma=1, learning_rate=0.01, max_depth=5, n_estimators=200, subsample=0.7	1466449.83	1210.97	960.45	3.75	0.96	4.35	1.39	-24.868947	0.98

Table 5.1: Results of the ablation study of the Machine learning architecture.

ID	Config.	MSE	RMSE	MAE	MAPE	R^2	CVRMSE	Training Time (s)	Overfit Gap	Loss Ratio
403	XGBoost; gamma=1, learning_rate=0.01, max_depth=5, n_estimators=200, subsample=0.8	1434508.99	1197.71	953.86	3.71	0.96	4.28	1.41	-29.698369	0.98
404	XGBoost; gamma=1, learning_rate=0.01, max_depth=5, n_estimators=200, subsample=1.0	1396281.21	1181.64	944.24	3.67	0.96	4.19	2.36	-32.025608	0.97
405	XGBoost; gamma=1, learning_rate=0.01, max_depth=7, n_estimators=50, subsample=0.7	19134864.96	4374.34	3667.73	14.31	0.5	15.13	0.58	-17.693049	1
406	XGBoost; gamma=1, learning_rate=0.01, max_depth=7, n_estimators=50, subsample=0.8	19100721.82	4370.44	3665.18	14.31	0.5	15.12	0.6	-19.164637	1
407	XGBoost; gamma=1, learning_rate=0.01, max_depth=7, n_estimators=50, subsample=1.0	19075841.93	4367.59	3662.45	14.3	0.5	15.1	0.6	-18.810314	1
408	XGBoost; gamma=1, learning_rate=0.01, max_depth=7, n_estimators=100, subsample=0.7	7376792.54	2716.03	2262.21	8.83	0.81	9.47	0.89	-22.911903	0.99
409	XGBoost; gamma=1, learning_rate=0.01, max_depth=7, n_estimators=100, subsample=0.8	7364929.56	2713.84	2262.01	8.83	0.81	9.45	0.9	-22.188066	0.99
410	XGBoost; gamma=1, learning_rate=0.01, max_depth=7, n_estimators=100, subsample=1.0	7327140.78	2706.87	2258.26	8.81	0.81	9.41	0.94	-23.810163	0.99
411	XGBoost; gamma=1, learning_rate=0.01, max_depth=7, n_estimators=200, subsample=0.7	1466449.83	1210.97	960.45	3.75	0.96	4.35	1.4	-24.868947	0.98
412	XGBoost; gamma=1, learning_rate=0.01, max_depth=7, n_estimators=200, subsample=0.8	1434508.99	1197.71	953.86	3.71	0.96	4.28	1.42	-29.698369	0.98
413	XGBoost; gamma=1, learning_rate=0.01, max_depth=7, n_estimators=200, subsample=1.0	1396281.21	1181.64	944.24	3.67	0.96	4.19	1.47	-32.025608	0.97
414	XGBoost; gamma=1, learning_rate=0.1, max_depth=3, n_estimators=50, subsample=0.7	429225.97	655.15	461.84	1.75	0.99	2.37	0.38	-38.065057	0.95
415	XGBoost; gamma=1, learning_rate=0.1, max_depth=3, n_estimators=50, subsample=0.8	400752.14	633.05	445.33	1.67	0.99	2.26	0.4	-40.126254	0.94
416	XGBoost; gamma=1, learning_rate=0.1, max_depth=3, n_estimators=50, subsample=1.0	381728.43	617.84	438.78	1.64	0.99	2.18	0.39	-32.908594	0.95

Table 5.1: Results of the ablation study of the Machine learning architecture.

ID	Config.	MSE	RMSE	MAE	MAPE	R^2	CVRMSE	Training Time (s)	Overfit Gap	Loss Ratio
417	XGBoost; gamma=1, learning_rate=0.1, max_depth=3, n_estimators=100, subsample=0.7	428802.79	654.83	461.42	1.75	0.99	2.37	0.49	-38.388234	0.94
418	XGBoost; gamma=1, learning_rate=0.1, max_depth=3, n_estimators=100, subsample=0.8	400665.34	632.98	445.24	1.67	0.99	2.26	0.44	-40.194807	0.94
419	XGBoost; gamma=1, learning_rate=0.1, max_depth=3, n_estimators=100, subsample=1.0	381733.37	617.85	438.79	1.64	0.99	2.18	0.49	-32.904624	0.95
420	XGBoost; gamma=1, learning_rate=0.1, max_depth=3, n_estimators=200, subsample=0.7	428092.45	654.29	460.7	1.75	0.99	2.37	0.62	-32.646724	0.95
421	XGBoost; gamma=1, learning_rate=0.1, max_depth=3, n_estimators=200, subsample=0.8	399364.65	631.95	443.84	1.67	0.99	2.26	0.65	-35.352635	0.95
422	XGBoost; gamma=1, learning_rate=0.1, max_depth=3, n_estimators=200, subsample=1.0	381735.77	617.85	438.79	1.64	0.99	2.18	0.65	-32.902545	0.95
423	XGBoost; gamma=1, learning_rate=0.1, max_depth=5, n_estimators=50, subsample=0.7	408550.62	639.18	445.51	1.69	0.99	2.31	0.42	-47.392631	0.93
424	XGBoost; gamma=1, learning_rate=0.1, max_depth=5, n_estimators=50, subsample=0.8	404672.21	636.14	447.49	1.68	0.99	2.28	0.39	-36.399037	0.95
425	XGBoost; gamma=1, learning_rate=0.1, max_depth=5, n_estimators=50, subsample=1.0	381499.18	617.66	440.38	1.63	0.99	2.15	0.41	-32.377064	0.95
426	XGBoost; gamma=1, learning_rate=0.1, max_depth=5, n_estimators=100, subsample=0.7	408069.6	638.8	444.99	1.68	0.99	2.31	0.48	-47.769072	0.93
427	XGBoost; gamma=1, learning_rate=0.1, max_depth=5, n_estimators=100, subsample=0.8	404652.21	636.12	447.47	1.68	0.99	2.28	0.5	-36.414736	0.95
428	XGBoost; gamma=1, learning_rate=0.1, max_depth=5, n_estimators=100, subsample=1.0	381505.04	617.66	440.39	1.63	0.99	2.15	0.49	-32.372436	0.95
429	XGBoost; gamma=1, learning_rate=0.1, max_depth=5, n_estimators=200, subsample=0.7	408346.19	639.02	445.29	1.69	0.99	2.31	0.67	-47.552450	0.93
430	XGBoost; gamma=1, learning_rate=0.1, max_depth=5, n_estimators=200, subsample=0.8	404402.78	635.93	447.2	1.68	0.99	2.28	0.7	-36.610916	0.95

Table 5.1: Results of the ablation study of the Machine learning architecture.

ID	Config.	MSE	RMSE	MAE	MAPE	R^2	CVRMSE	Training Time (s)	Overfit Gap	Loss Ratio
431	XGBoost: gamma=1, learning_rate=0.1, max_depth=5, n_estimators=200, subsample=1.0	381507.88	617.66	440.4	1.63	0.99	2.15	0.7	-32.370187	0.95
432	XGBoost: gamma=1, learning_rate=0.1, max_depth=7, n_estimators=50, subsample=0.7	408550.62	639.18	445.51	1.69	0.99	2.31	0.41	-47.392631	0.93
433	XGBoost: gamma=1, learning_rate=0.1, max_depth=7, n_estimators=50, subsample=0.8	404672.21	636.14	447.49	1.68	0.99	2.28	0.39	-36.399037	0.95
434	XGBoost: gamma=1, learning_rate=0.1, max_depth=7, n_estimators=50, subsample=1.0	381499.18	617.66	440.38	1.63	0.99	2.15	0.4	-32.377064	0.95
435	XGBoost: gamma=1, learning_rate=0.1, max_depth=7, n_estimators=100, subsample=0.7	408069.6	638.8	444.99	1.68	0.99	2.31	0.49	-47.769072	0.93
436	XGBoost: gamma=1, learning_rate=0.1, max_depth=7, n_estimators=100, subsample=0.8	404652.21	636.12	447.47	1.68	0.99	2.28	0.49	-36.414736	0.95
437	XGBoost: gamma=1, learning_rate=0.1, max_depth=7, n_estimators=100, subsample=1.0	381505.04	617.66	440.39	1.63	0.99	2.15	0.51	-32.372436	0.95
438	XGBoost: gamma=1, learning_rate=0.1, max_depth=7, n_estimators=200, subsample=0.7	408346.19	639.02	445.29	1.69	0.99	2.31	0.67	-47.552450	0.93
439	XGBoost: gamma=1, learning_rate=0.1, max_depth=7, n_estimators=200, subsample=0.8	404402.78	635.93	447.2	1.68	0.99	2.28	0.7	-36.610916	0.95
440	XGBoost: gamma=1, learning_rate=0.1, max_depth=7, n_estimators=200, subsample=1.0	381507.88	617.66	440.4	1.63	0.99	2.15	0.7	-32.370187	0.95
441	SVR: C=0.1, epsilon=0.01, kernel=linear	148216.7	384.99	258.51	0.9	1	1.28	921.78	-56.072572	0.87
442	SVR: C=0.1, epsilon=0.01, kernel=rbf	1989897.62	1410.64	1050.52	4.03	0.95	5.21	507.47	756.028301	2.15
443	SVR: C=0.1, epsilon=0.1, kernel=linear	3512111.56	1874.06	1581.55	5.76	0.91	4.8	1.66	452.917492	1.32
444	SVR: C=0.1, epsilon=0.1, kernel=rbf	6557051.84	2560.67	2081.01	7.97	0.83	8.17	6.28	582.779524	1.29
445	SVR: C=0.1, epsilon=0.2, kernel=linear	11347187.97	3368.56	2769.17	10.61	0.7	10.78	0.69	401.987520	1.14
446	SVR: C=0.1, epsilon=0.2, kernel=rbf	25705097.95	5070.02	4217.66	16.57	0.32	14.87	1.28	1158.003311	1.3

Table 5.1: Results of the ablation study of the Machine learning architecture.

ID	Config.	MSE	RMSE	MAE	MAPE	R^2	CVRMSE	Training Time (s)	Overfit Gap	Loss Ratio
447	SVR: C=1.0, epsilon=0.01, kernel=linear	143044.31	378.21	254.57	0.89	1	1.27	9932.86	-60.848095	0.86
448	SVR: C=1.0, epsilon=0.01, kernel=rbf	1688480.05	1299.42	1013.09	3.93	0.96	4.36	1309.78	906.864992	3.31
449	SVR: C=1.0, epsilon=0.1, kernel=linear	2946243.96	1716.46	1442.69	5.14	0.92	4.13	2.67	504.870654	1.42
450	SVR: C=1.0, epsilon=0.1, kernel=rbf	6914955.99	2629.63	2167.62	8.26	0.82	7.98	3.86	786.395894	1.43
451	SVR: C=1.0, epsilon=0.2, kernel=linear	5834281.08	2415.43	1983.37	7.23	0.85	8.31	0.87	-224.681832	0.91
452	SVR: C=1.0, epsilon=0.2, kernel=rbf	24779642.85	4977.92	4095.46	16.1	0.35	14.88	1.09	1112.513789	1.29

Table 5.2: Results of the ablation study of the Stacked LSTM architecture.

ID	Config.	MSE	RMSE	MAE	MAPE	R^2	CVRMSE	Training Time (s)	Overfit Gap	Loss Ratio
1	num_layers=1, dropout=True, epochs=10, batch_size=16	384695.55	620.24	443.36	1.45	0.99	2.01	332.54	-0.000123	0.68
2	num_layers=1, dropout=True, epochs=10, batch_size=32	476581.56	690.35	477.68	1.53	0.99	2.12	156.32	-0.000149	0.68
3	num_layers=1, dropout=True, epochs=10, batch_size=64	566230.43	752.48	549.87	1.72	0.99	2.08	79.88	-0.000150	0.72
4	num_layers=1, dropout=True, epochs=20, batch_size=16	330588.21	574.97	405.62	1.29	0.99	1.79	663.88	-0.000081	0.73
5	num_layers=1, dropout=True, epochs=20, batch_size=32	483506.39	695.35	508.58	1.69	0.99	2.27	312.9	-0.000039	0.89
6	num_layers=1, dropout=True, epochs=20, batch_size=64	441020.31	664.09	466.95	1.47	0.99	2.11	159.94	-0.000087	0.78
7	num_layers=1, dropout=True, epochs=50, batch_size=16	317355.3	563.34	412	1.36	0.99	1.83	1652.25	-0.000036	0.86
8	num_layers=1, dropout=True, epochs=50, batch_size=32	262201.1	512.06	349.15	1.12	0.99	1.66	780.65	-0.000059	0.75
9	num_layers=1, dropout=True, epochs=50, batch_size=64	3064671.95	1750.62	1438.97	4.06	0.94	3.23	393.7	0.001847	8.76
10	num_layers=1, dropout=False, epochs=10, batch_size=16	313059.99	559.52	384.83	1.24	0.99	1.83	322.14	-0.000050	0.81
11	num_layers=1, dropout=False, epochs=10, batch_size=32	385994.43	621.28	442.6	1.46	0.99	1.99	153.75	-0.000002	0.99

Table 5.2: Results of the ablation study of the Stacked LSTM architecture.

ID	Config.	MSE	RMSE	MAE	MAPE	R^2	CVRMSE	Training Time (s)	Overfit Gap	Loss Ratio
12	num_layers=1, dropout=False, epochs=10, batch_size=64	440168.68	663.45	449.03	1.44	0.99	2.12	78.27	-0.000026	0.92
13	num_layers=1, dropout=False, epochs=20, batch_size=16	293412.63	541.68	387.6	1.23	0.99	1.64	644.8	-0.000003	0.98
14	num_layers=1, dropout=False, epochs=20, batch_size=32	282089.41	531.12	363.25	1.16	0.99	1.71	310.56	-0.000019	0.91
15	num_layers=1, dropout=False, epochs=20, batch_size=64	410460.27	640.67	471.52	1.49	0.99	1.88	156.35	0.000041	1.17
16	num_layers=1, dropout=False, epochs=50, batch_size=16	235991.06	485.79	341.26	1.08	1	1.46	1611.98	0.000003	1.02
17	num_layers=1, dropout=False, epochs=50, batch_size=32	268392.88	518.07	370.08	1.17	0.99	1.54	762.08	0.000021	1.13
18	num_layers=1, dropout=False, epochs=50, batch_size=64	219325.22	468.32	315.39	1.01	1	1.52	389.03	-0.000018	0.89
19	num_layers=2, dropout=True, epochs=10, batch_size=16	378733.66	615.41	435.85	1.4	0.99	2.02	513.79	-0.000176	0.59
20	num_layers=2, dropout=True, epochs=10, batch_size=32	551913.8	742.91	550.4	1.85	0.99	2.48	246.34	-0.000065	0.85
21	num_layers=2, dropout=True, epochs=10, batch_size=64	590858.76	768.67	563.12	1.79	0.99	2.49	117.83	-0.000189	0.68
22	num_layers=2, dropout=True, epochs=20, batch_size=16	1117961.45	1057.34	864.66	2.82	0.98	3.35	1024.14	0.000470	2.61
23	num_layers=2, dropout=True, epochs=20, batch_size=32	445651.73	667.57	500.23	1.64	0.99	1.99	498.69	-0.000054	0.85
24	num_layers=2, dropout=True, epochs=20, batch_size=64	1927714.79	1388.42	1093.35	3.12	0.96	2.91	237.61	0.000965	3.76
25	num_layers=2, dropout=True, epochs=50, batch_size=16	305414.88	552.64	388.44	1.19	0.99	1.65	2571.78	-0.000018	0.92
26	num_layers=2, dropout=True, epochs=50, batch_size=32	3558243.37	1886.33	1525.8	4.52	0.93	5.03	1180.77	0.002209	11.24
27	num_layers=2, dropout=True, epochs=50, batch_size=64	4796429.28	2190.08	1808.18	5.04	0.91	3.1	591.34	0.003030	14.01
28	num_layers=2, dropout=False, epochs=10, batch_size=16	299566.83	547.33	376.22	1.2	0.99	1.77	495.24	-0.000056	0.79
29	num_layers=2, dropout=False, epochs=10, batch_size=32	313767.54	560.15	385.4	1.24	0.99	1.8	233.64	-0.000058	0.79
30	num_layers=2, dropout=False, epochs=10, batch_size=64	358836.9	599.03	425.95	1.37	0.99	1.9	116.63	-0.000046	0.84
31	num_layers=2, dropout=False, epochs=20, batch_size=16	254474.49	504.45	347.46	1.1	1	1.62	991.5	-0.000041	0.81

Table 5.2: Results of the ablation study of the Stacked LSTM architecture.

ID	Config.	MSE	RMSE	MAE	MAPE	R^2	CVRMSE	Training Time (s)	Overfit Gap	Loss Ratio
32	num_layers=2, dropout=False, epochs=20, batch_size=32	254250.27	504.23	344.46	1.1	1	1.64	450.79	-0.000044	0.8
33	num_layers=2, dropout=False, epochs=20, batch_size=64	367098.58	605.89	450.5	1.4	0.99	1.76	227.73	0.000019	1.08
34	num_layers=2, dropout=False, epochs=50, batch_size=16	328401.43	573.06	421.69	1.28	0.99	1.58	2442.54	0.000058	1.35
35	num_layers=2, dropout=False, epochs=50, batch_size=32	247511.72	497.51	348.96	1.1	1	1.51	1232.29	0.000006	1.04
36	num_layers=2, dropout=False, epochs=50, batch_size=64	224894.37	474.23	319.61	1.02	1	1.53	589.22	-0.000022	0.87
37	num_layers=3, dropout=True, epochs=10, batch_size=16	424040.92	651.18	490.17	1.64	0.99	2.06	656.65	-0.000188	0.61
38	num_layers=3, dropout=True, epochs=10, batch_size=32	2170386.61	1473.22	1149.8	3.23	0.96	2.7	318.26	0.001021	3.23
39	num_layers=3, dropout=True, epochs=10, batch_size=64	889141.09	942.94	700.74	2.49	0.98	3.33	161.07	-0.000021	0.97
40	num_layers=3, dropout=True, epochs=20, batch_size=16	4191812.18	2047.39	1591.24	4.36	0.92	3.37	1306.13	0.002546	9.2
41	num_layers=3, dropout=True, epochs=20, batch_size=32	7168050.56	2677.32	2264.15	6.37	0.86	3.43	633.74	0.004584	16.25
42	num_layers=3, dropout=True, epochs=20, batch_size=64	2207328.88	1485.71	1205.74	3.44	0.96	2.39	321.83	0.001150	4.22
43	num_layers=3, dropout=True, epochs=50, batch_size=16	3395723.69	1842.75	1417.6	3.87	0.93	3.27	3355.51	0.002101	10.86
44	num_layers=3, dropout=True, epochs=50, batch_size=32	2816321.41	1678.19	1307.26	3.69	0.94	3.73	1853.82	0.001682	8.09
45	num_layers=3, dropout=True, epochs=50, batch_size=64	1005851.38	1002.92	798.37	2.93	0.98	3.04	1183.04	0.000436	2.74
46	num_layers=3, dropout=False, epochs=10, batch_size=16	514056.33	716.98	560.35	1.78	0.99	1.8	1082.58	0.000084	1.31
47	num_layers=3, dropout=False, epochs=10, batch_size=32	400410.01	632.78	455.39	1.42	0.99	1.87	338.07	-0.000015	0.95
48	num_layers=3, dropout=False, epochs=10, batch_size=64	579008.27	760.93	565.08	1.82	0.99	2.2	194.01	0.000053	1.16
49	num_layers=3, dropout=False, epochs=20, batch_size=16	263879.45	513.69	361.26	1.15	0.99	1.62	1650.79	-0.000030	0.86
50	num_layers=3, dropout=False, epochs=20, batch_size=32	311075.68	557.74	403.95	1.29	0.99	1.7	837.36	-0.000015	0.93
51	num_layers=3, dropout=False, epochs=20, batch_size=64	283680.9	532.62	367.09	1.18	0.99	1.7	411.23	-0.000051	0.79

Table 5.2: Results of the ablation study of the Stacked LSTM architecture.

ID	Config.	MSE	RMSE	MAE	MAPE	R^2	CVRMSE	Training Time (s)	Overfit Gap	Loss Ratio
52	num_layers=3, dropout=False, epochs=50, batch_size=16	240325.15	490.23	337.09	1.07	1	1.55	4232.35	0.000006	1.04
53	num_layers=3, dropout=False, epochs=50, batch_size=32	378951.56	615.59	472.76	1.45	0.99	1.55	2135.71	0.000093	1.56
54	num_layers=3, dropout=False, epochs=50, batch_size=64	245321.24	495.3	340.88	1.09	1	1.58	1026.73	-0.000001	0.99
55	num_layers=4, dropout=True, epochs=10, batch_size=16	1520894.21	1233.25	938.39	2.8	0.97	3.45	1240.47	0.000626	2.53
56	num_layers=4, dropout=True, epochs=10, batch_size=32	3680606.19	1918.49	1460.02	4.02	0.93	3.54	595.7	0.002031	5.25
57	num_layers=4, dropout=True, epochs=10, batch_size=64	2602226.3	1613.14	1182.73	3.34	0.95	3.19	287.33	0.001092	2.6
58	num_layers=4, dropout=True, epochs=20, batch_size=16	2225399.42	1491.78	1283.48	4.54	0.96	4.4	1862.99	0.001208	4.91
59	num_layers=4, dropout=True, epochs=20, batch_size=32	3151510.17	1775.25	1415.61	3.96	0.94	3.57	806.86	0.001816	6.48
60	num_layers=4, dropout=True, epochs=20, batch_size=64	6735232.44	2595.23	2226.73	6.35	0.87	3.02	401.49	0.004215	12.15
61	num_layers=4, dropout=True, epochs=50, batch_size=16	5699837.37	2387.43	1920.03	5.29	0.89	3.49	5165.87	0.003646	16.3
62	num_layers=4, dropout=True, epochs=50, batch_size=32	4787794.7	2188.1	1809.74	6.34	0.91	7.08	3033.25	0.003008	12.8
63	num_layers=4, dropout=True, epochs=50, batch_size=64	3343631.94	1828.56	1419.34	3.9	0.93	3.15	1456.14	0.002033	9.15
64	num_layers=4, dropout=False, epochs=10, batch_size=16	329796.62	574.28	405.4	1.3	0.99	1.8	1021.68	-0.000034	0.87
65	num_layers=4, dropout=False, epochs=10, batch_size=32	696324.13	834.46	676.24	2.21	0.99	2.04	476.84	0.000186	1.65
66	num_layers=4, dropout=False, epochs=10, batch_size=64	422199.29	649.77	474.31	1.52	0.99	1.99	263.19	-0.000053	0.84
67	num_layers=4, dropout=False, epochs=20, batch_size=16	250323.2	500.32	338.36	1.07	1	1.6	2156.1	-0.000040	0.81
68	num_layers=4, dropout=False, epochs=20, batch_size=32	299322.6	547.1	389.53	1.25	0.99	1.68	1013.48	-0.000020	0.91
69	num_layers=4, dropout=False, epochs=20, batch_size=64	496319.53	704.5	554.41	1.74	0.99	1.73	482.08	0.000096	1.39
70	num_layers=4, dropout=False, epochs=50, batch_size=16	225944.6	475.34	330.75	1.06	1	1.54	5066.28	0.000001	1
71	num_layers=4, dropout=False, epochs=50, batch_size=32	221558.31	470.7	319.77	1.02	1	1.52	2366.15	-0.000012	0.92

Table 5.2: Results of the ablation study of the Stacked LSTM architecture.

ID	Config.	MSE	RMSE	MAE	MAPE	R^2	CVRMSE	Training Time (s)	Overfit Gap	Loss Ratio
72	num_layers=4, dropout=False, epochs=50, batch_size=64	224493.78	473.81	319.47	1.02	1	1.52	1197.92	-0.000030	0.84
73	num_layers=5, dropout=True, epochs=10, batch_size=16	3761567.16	1939.48	1606.07	4.6	0.93	4.17	1261.27	0.002182	6.72
74	num_layers=5, dropout=True, epochs=10, batch_size=32	1428606.01	1195.24	888.97	2.72	0.97	3.53	627.59	0.000484	1.99
75	num_layers=5, dropout=True, epochs=10, batch_size=64	2882271.75	1697.73	1300.61	3.75	0.94	4.2	295.42	0.001285	2.89
76	num_layers=5, dropout=True, epochs=20, batch_size=16	1666747.72	1291.03	993.54	3.34	0.97	4.49	2490.76	0.000808	3.47
77	num_layers=5, dropout=True, epochs=20, batch_size=32	1820279.06	1349.18	1074.46	3.23	0.96	3.86	1251.47	0.000902	3.66
78	num_layers=5, dropout=True, epochs=20, batch_size=64	3355491.37	1831.8	1540.28	4.66	0.93	5.25	601.12	0.001909	6.08
79	num_layers=5, dropout=True, epochs=50, batch_size=16	2127101.48	1458.46	1163.71	4.21	0.96	5	6354.87	0.001207	5.97
80	num_layers=5, dropout=True, epochs=50, batch_size=32	2488392.62	1577.46	1234.49	3.58	0.95	4.01	3134.16	0.001454	7.03
81	num_layers=5, dropout=True, epochs=50, batch_size=64	3720370.52	1928.83	1640.26	5.77	0.93	6.21	1536.97	0.002297	10.64
82	num_layers=5, dropout=False, epochs=10, batch_size=16	51070843.52	7146.39	5946.3	19.69	-0	24.06	1211.33	0.000225	1.01
83	num_layers=5, dropout=False, epochs=10, batch_size=32	313147.58	559.6	397.06	1.26	0.99	1.75	633.6	-0.000059	0.78
84	num_layers=5, dropout=False, epochs=10, batch_size=64	486057.65	697.18	522.86	1.65	0.99	2.06	298.31	0.000011	1.03
85	num_layers=5, dropout=False, epochs=20, batch_size=16	249225.3	499.22	347.99	1.11	1	1.58	2531.31	-0.000038	0.82
86	num_layers=5, dropout=False, epochs=20, batch_size=32	271780.91	521.33	365.49	1.15	0.99	1.64	1253.22	-0.000041	0.82
87	num_layers=5, dropout=False, epochs=20, batch_size=64	311775.92	558.37	394.17	1.28	0.99	1.86	604.83	-0.000038	0.85
88	num_layers=5, dropout=False, epochs=50, batch_size=16	267145.94	516.86	365.51	1.15	0.99	1.5	6359.32	0.000030	1.2
89	num_layers=5, dropout=False, epochs=50, batch_size=32	208583.58	456.71	313.34	1	1	1.46	3158.46	-0.000012	0.92
90	num_layers=5, dropout=False, epochs=50, batch_size=64	241447.74	491.37	346.76	1.11	1	1.53	1540.72	-0.000001	1

Table 5.3: Results of the ablation study of the LSTMCNN architecture.

ID	Config.	MSE	RMSE	MAE	MAPE	R^2	CVRMSE	Training Time (s)	Overfit Gap	Loss Ratio
1	num_layers=1, dropout=True, epochs=10, batch_size=16	338327.22	581.66	413.66	1.32	0.99	1.88	476.64	-0.000100	0.7
2	num_layers=1, dropout=True, epochs=10, batch_size=32	256300.33	506.26	355.71	1.15	0.99	1.66	254.99	-0.000065	0.73
3	num_layers=1, dropout=True, epochs=10, batch_size=64	256364.57	506.32	358.68	1.16	0.99	1.61	130.68	-0.000016	0.92
4	num_layers=1, dropout=True, epochs=20, batch_size=16	264072.72	513.88	365.47	1.18	0.99	1.68	996.61	-0.000022	0.89
5	num_layers=1, dropout=True, epochs=20, batch_size=32	229742.24	479.31	335.93	1.06	1	1.51	403.85	0.000001	1.01
6	num_layers=1, dropout=True, epochs=20, batch_size=64	227285.07	476.74	332.7	1.05	1	1.5	180.3	0.000024	1.18
7	num_layers=1, dropout=True, epochs=50, batch_size=16	240245.8	490.15	352.5	1.13	1	1.53	1853.75	0.000009	1.06
8	num_layers=1, dropout=True, epochs=50, batch_size=32	212389.7	460.86	327.96	1.03	1	1.45	889.15	0.000047	1.48
9	num_layers=1, dropout=True, epochs=50, batch_size=64	200016.09	447.23	314.3	1	1	1.43	450.19	0.000059	1.77
10	num_layers=1, dropout=False, epochs=10, batch_size=16	290735.93	539.2	377.8	1.21	0.99	1.75	365.02	-0.000079	0.71
11	num_layers=1, dropout=False, epochs=10, batch_size=32	260507.96	510.4	349.33	1.1	0.99	1.57	215.42	-0.000028	0.86
12	num_layers=1, dropout=False, epochs=10, batch_size=64	292076.55	540.44	391.77	1.22	0.99	1.53	117.76	0.000023	1.13
13	num_layers=1, dropout=False, epochs=20, batch_size=16	240846.25	490.76	341.15	1.09	1	1.59	868.49	-0.000020	0.89
14	num_layers=1, dropout=False, epochs=20, batch_size=32	219656.56	468.68	320.95	1.02	1	1.49	354.26	0.000002	1.01
15	num_layers=1, dropout=False, epochs=20, batch_size=64	215572.44	464.3	318.97	1.01	1	1.47	179.73	0.000022	1.18
16	num_layers=1, dropout=False, epochs=50, batch_size=16	262795.61	512.64	348.66	1.1	0.99	1.6	1843.73	0.000054	1.43
17	num_layers=1, dropout=False, epochs=50, batch_size=32	222722.33	471.93	322.81	1.02	1	1.5	1132.94	0.000059	1.64
18	num_layers=1, dropout=False, epochs=50, batch_size=64	217581.76	466.46	323.12	1.03	1	1.5	591.31	0.000092	2.62
19	num_layers=2, dropout=True, epochs=10, batch_size=16	451867.08	672.21	501.24	1.62	0.99	1.95	793.88	-0.000061	0.83
20	num_layers=2, dropout=True, epochs=10, batch_size=32	281182.72	530.27	377.68	1.21	0.99	1.66	416.39	-0.000053	0.78

Table 5.3: Results of the ablation study of the LSTMCNN architecture.

ID	Config.	MSE	RMSE	MAE	MAPE	R^2	CVRMSE	Training Time (s)	Overfit Gap	Loss Ratio
21	num_layers=2, dropout=True, epochs=10, batch_size=64	244823.43	494.8	345.26	1.09	1	1.55	226.39	-0.000037	0.82
22	num_layers=2, dropout=True, epochs=20, batch_size=16	313607.87	560.01	405.07	1.26	0.99	1.74	1675.74	-0.000002	0.99
23	num_layers=2, dropout=True, epochs=20, batch_size=32	219451.9	468.46	325.64	1.04	1	1.52	823.08	-0.000016	0.9
24	num_layers=2, dropout=True, epochs=20, batch_size=64	217581.77	466.46	325.63	1.04	1	1.51	388.48	0.000009	1.07
25	num_layers=2, dropout=True, epochs=50, batch_size=16	229610.34	479.18	337.44	1.08	1	1.52	4231.23	0.000007	1.05
26	num_layers=2, dropout=True, epochs=50, batch_size=32	216926.85	465.75	331.38	1.06	1	1.47	1625.05	0.000045	1.43
27	num_layers=2, dropout=True, epochs=50, batch_size=64	200798.33	448.11	314.22	0.99	1	1.41	782.53	0.000059	1.75
28	num_layers=2, dropout=False, epochs=10, batch_size=16	614730.07	784.05	591.04	1.8	0.99	1.82	572.64	0.000147	1.54
29	num_layers=2, dropout=False, epochs=10, batch_size=32	263973.33	513.78	351.91	1.11	0.99	1.61	278.26	-0.000020	0.9
30	num_layers=2, dropout=False, epochs=10, batch_size=64	224528.81	473.84	322.74	1.02	1	1.52	140.78	-0.000018	0.89
31	num_layers=2, dropout=False, epochs=20, batch_size=16	287447.75	536.14	377.05	1.18	0.99	1.56	1131.2	0.000017	1.1
32	num_layers=2, dropout=False, epochs=20, batch_size=32	222159.1	471.34	322.69	1.02	1	1.51	545.19	0.000009	1.06
33	num_layers=2, dropout=False, epochs=20, batch_size=64	207125.69	455.11	310.19	0.99	1	1.46	273.02	0.000021	1.18
34	num_layers=2, dropout=False, epochs=50, batch_size=16	218627.86	467.58	324.21	1.04	1	1.52	2811.4	0.000029	1.25
35	num_layers=2, dropout=False, epochs=50, batch_size=32	218120.13	467.03	331.72	1.04	1	1.46	1349.76	0.000066	1.8
36	num_layers=2, dropout=False, epochs=50, batch_size=64	199339.95	446.48	313.25	1	1	1.44	684.64	0.000084	2.61
37	num_layers=3, dropout=True, epochs=10, batch_size=16	493530.74	702.52	515.99	1.56	0.99	1.95	399.16	0.000022	1.07
38	num_layers=3, dropout=True, epochs=10, batch_size=32	244475.61	494.44	343.41	1.09	1	1.59	191.18	-0.000058	0.74
39	num_layers=3, dropout=True, epochs=10, batch_size=64	228717.5	478.24	330.03	1.05	1	1.53	97.33	-0.000020	0.89
40	num_layers=3, dropout=True, epochs=20, batch_size=16	280393.82	529.52	385.04	1.22	0.99	1.61	790.28	-0.000000	1

Table 5.3: Results of the ablation study of the LSTMCNN architecture.

ID	Config.	MSE	RMSE	MAE	MAPE	R^2	CVRMSE	Training Time (s)	Overfit Gap	Loss Ratio
41	num_layers=3, dropout=True, epochs=20, batch_size=32	272679.91	522.19	382.13	1.21	0.99	1.53	381.63	0.000037	1.25
42	num_layers=3, dropout=True, epochs=20, batch_size=64	225070.43	474.42	336.56	1.06	1	1.5	192.96	0.000044	1.4
43	num_layers=3, dropout=True, epochs=50, batch_size=16	405401.45	636.71	413.1	1.31	0.99	1.87	1976.76	0.000119	1.76
44	num_layers=3, dropout=True, epochs=50, batch_size=32	190439.42	436.39	299.22	0.95	1	1.41	955.38	0.000038	1.41
45	num_layers=3, dropout=True, epochs=50, batch_size=64	191459.48	437.56	301.53	0.96	1	1.39	482.66	0.000067	2.04
46	num_layers=3, dropout=False, epochs=10, batch_size=16	305748.94	552.95	375.03	1.19	0.99	1.77	380.77	-0.000067	0.76
47	num_layers=3, dropout=False, epochs=10, batch_size=32	305954.29	553.13	372.8	1.18	0.99	1.73	182.37	-0.000003	0.98
48	num_layers=3, dropout=False, epochs=10, batch_size=64	248499.46	498.5	346.13	1.1	1	1.58	92.63	-0.000006	0.97
49	num_layers=3, dropout=False, epochs=20, batch_size=16	281927.85	530.97	375.29	1.2	0.99	1.7	759.07	0.000002	1.01
50	num_layers=3, dropout=False, epochs=20, batch_size=32	243665.16	493.62	347.49	1.09	1	1.54	366.37	0.000012	1.08
51	num_layers=3, dropout=False, epochs=20, batch_size=64	220527.06	469.6	322.94	1.02	1	1.51	188	0.000024	1.19
52	num_layers=3, dropout=False, epochs=50, batch_size=16	298935.74	546.75	398.97	1.25	0.99	1.68	1916.87	0.000072	1.55
53	num_layers=3, dropout=False, epochs=50, batch_size=32	230795.95	480.41	331.1	1.05	1	1.53	917.79	0.000068	1.76
54	num_layers=3, dropout=False, epochs=50, batch_size=64	222837.19	472.06	331.22	1.06	1	1.5	468.64	0.000094	2.6
55	num_layers=4, dropout=True, epochs=10, batch_size=16	428328.07	654.47	486.09	1.51	0.99	1.94	540.58	-0.000077	0.79
56	num_layers=4, dropout=True, epochs=10, batch_size=32	302299.65	549.82	396	1.25	0.99	1.69	278.55	-0.000040	0.84
57	num_layers=4, dropout=True, epochs=10, batch_size=64	266604.76	516.34	364.99	1.18	0.99	1.61	133.64	-0.000017	0.92
58	num_layers=4, dropout=True, epochs=20, batch_size=16	278298.84	527.54	372.29	1.16	0.99	1.62	1085.28	-0.000024	0.89
59	num_layers=4, dropout=True, epochs=20, batch_size=32	292212.5	540.57	380.84	1.2	0.99	1.72	524.47	0.000009	1.05
60	num_layers=4, dropout=True, epochs=20, batch_size=64	207524.23	455.55	312.61	0.99	1	1.47	274.31	0.000018	1.14

Table 5.3: Results of the ablation study of the LSTMCNN architecture.

ID	Config.	MSE	RMSE	MAE	MAPE	R^2	CVRMSE	Training Time (s)	Overfit Gap	Loss Ratio
61	num_layers=4, dropout=True, epochs=50, batch_size=16	247569.56	497.56	354.23	1.16	1	1.62	2728.43	0.000015	1.1
62	num_layers=4, dropout=True, epochs=50, batch_size=32	205792.53	453.64	313.99	1	1	1.46	1349.3	0.000039	1.38
63	num_layers=4, dropout=True, epochs=50, batch_size=64	198022.2	445	310.49	0.98	1	1.41	685.59	0.000062	1.84
64	num_layers=4, dropout=False, epochs=10, batch_size=16	319114.99	564.9	398.82	1.27	0.99	1.73	529.26	-0.000074	0.75
65	num_layers=4, dropout=False, epochs=10, batch_size=32	357871.21	598.22	436.25	1.33	0.99	1.59	259.41	0.000030	1.14
66	num_layers=4, dropout=False, epochs=10, batch_size=64	274078.89	523.53	363.27	1.16	0.99	1.6	131.23	0.000012	1.07
67	num_layers=4, dropout=False, epochs=20, batch_size=16	277739.01	527.01	352.15	1.11	0.99	1.66	1055.26	-0.000003	0.98
68	num_layers=4, dropout=False, epochs=20, batch_size=32	232515.6	482.2	328.41	1.04	1	1.53	516.8	0.000003	1.02
69	num_layers=4, dropout=False, epochs=20, batch_size=64	210168.8	458.44	307.84	0.98	1	1.47	260.86	0.000016	1.12
70	num_layers=4, dropout=False, epochs=50, batch_size=16	237731.18	487.58	344.49	1.07	1	1.49	2638.88	0.000043	1.36
71	num_layers=4, dropout=False, epochs=50, batch_size=32	212653.23	461.14	324.18	1.03	1	1.43	1296.93	0.000072	1.99
72	num_layers=4, dropout=False, epochs=50, batch_size=64	204065.03	451.74	314.33	1	1	1.43	655.77	0.000090	2.85
73	num_layers=5, dropout=True, epochs=10, batch_size=16	509895.41	714.07	529.39	1.63	0.99	2.09	778.76	-0.000033	0.91
74	num_layers=5, dropout=True, epochs=10, batch_size=32	477905.9	691.31	514.08	1.54	0.99	1.92	380	0.000064	1.25
75	num_layers=5, dropout=True, epochs=10, batch_size=64	378562.52	615.27	453.63	1.42	0.99	1.95	193.3	0.000043	1.2
76	num_layers=5, dropout=True, epochs=20, batch_size=16	364388.54	603.65	455.87	1.5	0.99	1.87	1551.51	0.000017	1.07
77	num_layers=5, dropout=True, epochs=20, batch_size=32	396317.47	629.54	463.91	1.44	0.99	1.98	762.55	0.000080	1.42
78	num_layers=5, dropout=True, epochs=20, batch_size=64	359369.7	599.47	444.02	1.42	0.99	1.93	387.24	0.000080	1.48
79	num_layers=5, dropout=True, epochs=50, batch_size=16	456851.02	675.91	520.35	1.71	0.99	2.28	3883.5	0.000114	1.58
80	num_layers=5, dropout=True, epochs=50, batch_size=32	518313.86	719.94	560.56	1.79	0.99	2.33	1911.11	0.000237	3.04

Table 5.3: Results of the ablation study of the LSTMCNN architecture.

ID	Config.	MSE	RMSE	MAE	MAPE	R^2	CVRMSE	Training Time (s)	Overfit Gap	Loss Ratio
81	num_layers=5, dropout=True, epochs=50, batch_size=64	533333.41	730.3	582.93	1.94	0.99	2.43	965.25	0.000283	4.51
82	num_layers=5, dropout=False, epochs=10, batch_size=16	446857.43	668.47	496.5	1.54	0.99	1.87	729.8	0.000004	1.01
83	num_layers=5, dropout=False, epochs=10, batch_size=32	299745.76	547.49	384.03	1.26	0.99	1.85	355.97	-0.000011	0.95
84	num_layers=5, dropout=False, epochs=10, batch_size=64	272502.59	522.02	360.95	1.15	0.99	1.62	180.02	0.000012	1.07
85	num_layers=5, dropout=False, epochs=20, batch_size=16	286510.94	535.27	380.94	1.21	0.99	1.72	1459.52	-0.000001	0.99
86	num_layers=5, dropout=False, epochs=20, batch_size=32	218200.59	467.12	321.38	1.02	1	1.51	720.49	-0.000002	0.99
87	num_layers=5, dropout=False, epochs=20, batch_size=64	230579.73	480.19	334.35	1.08	1	1.53	364.1	0.000040	1.34
88	num_layers=5, dropout=False, epochs=50, batch_size=16	253615.19	503.6	362.85	1.18	1	1.6	3615.78	0.000043	1.33
89	num_layers=5, dropout=False, epochs=50, batch_size=32	195609.61	442.28	308.61	0.98	1	1.42	1833.34	0.000057	1.74
90	num_layers=5, dropout=False, epochs=50, batch_size=64	199303.86	446.43	312.51	1	1	1.44	952.49	0.000088	2.86

Table 5.4: Results of the ablation study of the Stacked LSTM architecture.

ID	Config.	MSE	RMSE	MAE	MAPE	R^2	CVRMSE	Training Time (s)	Overfit Gap	Loss Ratio
1	num_layers=1, dropout=True, epochs=10, batch_size=16	508043.99	712.77	517.91	1.63	0.99	2.22	299.65	-0.000212	0.62
2	num_layers=1, dropout=True, epochs=10, batch_size=32	874665.43	935.24	719.97	2.23	0.98	2.4	148.18	-0.000022	0.96
3	num_layers=1, dropout=True, epochs=10, batch_size=64	1078683.2	1038.6	796.51	2.38	0.98	2.73	74.4	0.000028	1.04
4	num_layers=1, dropout=True, epochs=20, batch_size=16	512159.85	715.65	511.87	1.57	0.99	2.05	598.86	-0.000075	0.82
5	num_layers=1, dropout=True, epochs=20, batch_size=32	387530.64	622.52	434.01	1.4	0.99	2.07	287.17	-0.000200	0.57
6	num_layers=1, dropout=True, epochs=20, batch_size=64	399058.48	631.71	449.94	1.46	0.99	2.06	144.6	-0.000203	0.57
7	num_layers=1, dropout=True, epochs=50, batch_size=16	446699.59	668.36	481.75	1.45	0.99	1.96	1974.53	0.000024	1.09

Table 5.4: Results of the ablation study of the Stacked LSTM architecture.

ID	Config.	MSE	RMSE	MAE	MAPE	R^2	CVRMSE	Training Time (s)	Overfit Gap	Loss Ratio
8	num_layers=1, dropout=True, epochs=50, batch_size=32	334627.18	578.47	398.04	1.24	0.99	1.81	853.2	-0.000078	0.75
9	num_layers=1, dropout=True, epochs=50, batch_size=64	345288.76	587.61	412.6	1.3	0.99	1.89	378.99	-0.000107	0.69
10	num_layers=1, dropout=False, epochs=10, batch_size=16	417110.59	645.84	486.64	1.56	0.99	1.85	284.43	-0.000005	0.98
11	num_layers=1, dropout=False, epochs=10, batch_size=32	471071.66	686.35	498.84	1.53	0.99	1.88	139.12	0.000030	1.1
12	num_layers=1, dropout=False, epochs=10, batch_size=64	360953.98	600.79	415.06	1.33	0.99	1.92	72.77	-0.000099	0.71
13	num_layers=1, dropout=False, epochs=20, batch_size=16	320375.36	566.02	395.92	1.26	0.99	1.84	575.22	-0.000017	0.93
14	num_layers=1, dropout=False, epochs=20, batch_size=32	373224.63	610.92	452.43	1.45	0.99	1.68	273.38	0.000020	1.09
15	num_layers=1, dropout=False, epochs=20, batch_size=64	388286.18	623.13	463.43	1.46	0.99	1.78	136.29	0.000020	1.08
16	num_layers=1, dropout=False, epochs=50, batch_size=16	245022	495	333.49	1.07	1	1.62	1409.65	-0.000011	0.94
17	num_layers=1, dropout=False, epochs=50, batch_size=32	295516.17	543.61	381.74	1.22	0.99	1.62	684.3	0.000020	1.11
18	num_layers=1, dropout=False, epochs=50, batch_size=64	291143.76	539.58	373.83	1.18	0.99	1.6	341.12	0.000021	1.12
19	num_layers=2, dropout=True, epochs=10, batch_size=16	628984.93	793.09	591.33	1.89	0.99	2.58	303.33	-0.000199	0.68
20	num_layers=2, dropout=True, epochs=10, batch_size=32	774379.41	879.99	646.67	2.17	0.98	3.01	151.41	-0.000158	0.77
21	num_layers=2, dropout=True, epochs=10, batch_size=64	845085.32	919.29	670.38	2.06	0.98	2.75	76.5	-0.000169	0.77
22	num_layers=2, dropout=True, epochs=20, batch_size=16	720776.27	848.99	644.95	1.95	0.99	2.48	609.55	0.000040	1.09
23	num_layers=2, dropout=True, epochs=20, batch_size=32	1512568.55	1229.87	902.35	2.54	0.97	2.82	310.63	0.000577	2.27
24	num_layers=2, dropout=True, epochs=20, batch_size=64	1261253.07	1123.06	871.86	2.55	0.97	2.77	153.13	0.000352	1.69
25	num_layers=2, dropout=True, epochs=50, batch_size=16	1160881.19	1077.44	828.96	2.46	0.98	3.06	1550.41	0.000478	2.52
26	num_layers=2, dropout=True, epochs=50, batch_size=32	1181495.93	1086.97	807.42	2.34	0.98	2.89	771.12	0.000485	2.52
27	num_layers=2, dropout=True, epochs=50, batch_size=64	2942853.24	1715.47	1254.91	3.45	0.94	3.53	384.95	0.001668	5.98

Table 5.4: Results of the ablation study of the Stacked LSTM architecture.

ID	Config.	MSE	RMSE	MAE	MAPE	R^2	CVRMSE	Training Time (s)	Overfit Gap	Loss Ratio
28	num_layers=2, dropout=False, epochs=10, batch_size=16	318744.52	564.57	388.81	1.25	0.99	1.86	276.75	-0.000085	0.72
29	num_layers=2, dropout=False, epochs=10, batch_size=32	386719.02	621.87	436.98	1.4	0.99	2.01	137.36	-0.000046	0.85
30	num_layers=2, dropout=False, epochs=10, batch_size=64	375478.17	612.76	435.01	1.38	0.99	1.94	69.63	-0.000052	0.83
31	num_layers=2, dropout=False, epochs=20, batch_size=16	300515.88	548.19	374.79	1.2	0.99	1.78	566.6	-0.000035	0.85
32	num_layers=2, dropout=False, epochs=20, batch_size=32	520345.28	721.35	561.23	1.72	0.99	1.68	278.31	0.000116	1.48
33	num_layers=2, dropout=False, epochs=20, batch_size=64	996910.24	998.45	860.17	2.72	0.98	2.03	138.86	0.000441	2.84
34	num_layers=2, dropout=False, epochs=50, batch_size=16	289117.65	537.7	374	1.17	0.99	1.61	1381.91	0.000022	1.12
35	num_layers=2, dropout=False, epochs=50, batch_size=32	240775.73	490.69	334.04	1.07	1	1.62	686.43	-0.000016	0.91
36	num_layers=2, dropout=False, epochs=50, batch_size=64	279473.69	528.65	373.09	1.16	0.99	1.63	341.91	0.000025	1.15
37	num_layers=3, dropout=True, epochs=10, batch_size=16	591300.23	768.96	546.49	1.7	0.99	2.26	413.76	-0.000203	0.66
38	num_layers=3, dropout=True, epochs=10, batch_size=32	613262.78	783.11	571.35	1.86	0.99	2.61	209.07	-0.000251	0.62
39	num_layers=3, dropout=True, epochs=10, batch_size=64	770258.49	877.64	652.48	2.05	0.98	2.78	105.07	-0.000303	0.63
40	num_layers=3, dropout=True, epochs=20, batch_size=16	433907.35	658.72	465.11	1.5	0.99	2.16	841.72	-0.000144	0.67
41	num_layers=3, dropout=True, epochs=20, batch_size=32	501137.65	707.91	519.97	1.64	0.99	2.26	405.66	-0.000103	0.77
42	num_layers=3, dropout=True, epochs=20, batch_size=64	729943.47	854.37	616.04	1.84	0.99	2.42	204.46	-0.000012	0.98
43	num_layers=3, dropout=True, epochs=50, batch_size=16	421552.89	649.27	470.45	1.49	0.99	2.05	2042.96	-0.000013	0.96
44	num_layers=3, dropout=True, epochs=50, batch_size=32	560162.63	748.44	564.21	1.83	0.99	2.08	1016.52	0.000070	1.22
45	num_layers=3, dropout=True, epochs=50, batch_size=64	1004884.38	1002.44	798.3	2.51	0.98	3.09	513.16	0.000357	2.09
46	num_layers=3, dropout=False, epochs=10, batch_size=16	321503.39	567.01	399.17	1.29	0.99	1.86	384.16	-0.000075	0.74
47	num_layers=3, dropout=False, epochs=10, batch_size=32	864406.88	929.73	762.46	2.35	0.98	1.93	188.14	0.000290	1.97

Table 5.4: Results of the ablation study of the Stacked LSTM architecture.

ID	Config.	MSE	RMSE	MAE	MAPE	R^2	CVRMSE	Training Time (s)	Overfit Gap	Loss Ratio
48	num_layers=3, dropout=False, epochs=10, batch_size=64	341600.56	584.47	401.04	1.28	0.99	1.9	95.99	-0.000093	0.71
49	num_layers=3, dropout=False, epochs=20, batch_size=16	492356.58	701.68	536.69	1.69	0.99	1.75	769.35	0.000101	1.43
50	num_layers=3, dropout=False, epochs=20, batch_size=32	371477.06	609.49	436.53	1.35	0.99	1.73	421.95	0.000007	1.03
51	num_layers=3, dropout=False, epochs=20, batch_size=64	418665.9	647.04	483.62	1.54	0.99	1.77	194.83	0.000033	1.13
52	num_layers=3, dropout=False, epochs=50, batch_size=16	235469.59	485.25	324.19	1.03	1	1.56	1841.92	-0.000014	0.92
53	num_layers=3, dropout=False, epochs=50, batch_size=32	474424.95	688.79	533.51	1.66	0.99	1.68	896.49	0.000147	1.84
54	num_layers=3, dropout=False, epochs=50, batch_size=64	353874.01	594.87	444.04	1.38	0.99	1.62	460.71	0.000054	1.29
55	num_layers=4, dropout=True, epochs=10, batch_size=16	1347342.04	1160.75	870	2.55	0.97	2.99	377.01	0.000237	1.35
56	num_layers=4, dropout=True, epochs=10, batch_size=32	602899.24	776.47	564.23	1.77	0.99	2.46	192.76	-0.000310	0.57
57	num_layers=4, dropout=True, epochs=10, batch_size=64	1178835.81	1085.74	825.5	2.52	0.98	3.19	97.72	-0.000061	0.93
58	num_layers=4, dropout=True, epochs=20, batch_size=16	628965.31	793.07	560.72	1.73	0.99	2.35	752.54	-0.000024	0.95
59	num_layers=4, dropout=True, epochs=20, batch_size=32	583591.15	763.93	581.14	1.82	0.99	2.41	378.44	-0.000097	0.8
60	num_layers=4, dropout=True, epochs=20, batch_size=64	786549.74	886.88	643.3	1.94	0.98	2.48	192.63	-0.000031	0.95
61	num_layers=4, dropout=True, epochs=50, batch_size=16	961362.69	980.49	778.4	2.36	0.98	2.89	1887.26	0.000346	2.12
62	num_layers=4, dropout=True, epochs=50, batch_size=32	630574.71	794.09	589.73	1.74	0.99	2.18	946.61	0.000109	1.34
63	num_layers=4, dropout=True, epochs=50, batch_size=64	2029474.29	1424.6	1111.83	3.27	0.96	3.92	522.59	0.001049	4.15
64	num_layers=4, dropout=False, epochs=10, batch_size=16	388524.93	623.32	451.41	1.43	0.99	1.83	493.14	-0.000042	0.86
65	num_layers=4, dropout=False, epochs=10, batch_size=32	1210524.97	1100.24	897.77	2.69	0.98	2.26	250.75	0.000493	2.49
66	num_layers=4, dropout=False, epochs=10, batch_size=64	448902.39	670	481.41	1.51	0.99	2.03	132.2	-0.000034	0.9
67	num_layers=4, dropout=False, epochs=20, batch_size=16	288569.86	537.19	374.29	1.2	0.99	1.74	1029.07	-0.000042	0.82

Table 5.4: Results of the ablation study of the Stacked LSTM architecture.

ID	Config.	MSE	RMSE	MAE	MAPE	R^2	CVRMSE	Training Time (s)	Overfit Gap	Loss Ratio
68	num_layers=4, dropout=False, epochs=20, batch_size=32	308502.4	555.43	381.5	1.22	0.99	1.8	503.63	-0.000026	0.89
69	num_layers=4, dropout=False, epochs=20, batch_size=64	308070.97	555.04	389.37	1.25	0.99	1.76	255.14	-0.000041	0.84
70	num_layers=4, dropout=False, epochs=50, batch_size=16	243290.25	493.24	338.3	1.07	1	1.6	2059.35	-0.000003	0.98
71	num_layers=4, dropout=False, epochs=50, batch_size=32	247081.5	497.07	338.25	1.08	1	1.62	900.35	-0.000007	0.96
72	num_layers=4, dropout=False, epochs=50, batch_size=64	476455.75	690.26	519.64	1.55	0.99	1.78	451.39	0.000152	1.88

Table 5.5: Results of the ablation study of the BiLSTM architecture.

ID	Config.	MSE	RMSE	MAE	MAPE	R^2	CVRMSE	Training Time (s)	Overfit Gap	Loss Ratio
1	units_lstm1=64, units_lstm2=32, dropout_rate=0.0, feature_att=True, temporal_att=True, epochs=10, batch_size=32	410570.54	640.76	430.01	1.35	0.99	1.94	365.3	-0.000074	0.79
2	units_lstm1=64, units_lstm2=32, dropout_rate=0.0, feature_att=True, temporal_att=False, epochs=10, batch_size=32	589661.59	767.89	579.95	1.77	0.99	1.98	355.32	0.000058	1.17
3	units_lstm1=64, units_lstm2=32, dropout_rate=0.0, feature_att=False, temporal_att=True, epochs=10, batch_size=32	477187.93	690.79	492.32	1.6	0.99	2.23	346.99	0.000008	1.03
4	units_lstm1=64, units_lstm2=32, dropout_rate=0.0, feature_att=False, temporal_att=False, epochs=10, batch_size=32	547839.7	740.16	535.04	1.71	0.99	2.39	339.48	-0.000042	0.9
5	units_lstm1=64, units_lstm2=32, dropout_rate=0.0, feature_att=True, temporal_att=True, epochs=10, batch_size=64	1328547.62	1152.63	957.53	3.07	0.97	2.53	185.19	0.000461	2.04
6	units_lstm1=64, units_lstm2=32, dropout_rate=0.0, feature_att=True, temporal_att=False, epochs=10, batch_size=64	448395.44	669.62	477.49	1.5	0.99	2.08	180.36	-0.000067	0.82
7	units_lstm1=64, units_lstm2=32, dropout_rate=0.0, feature_att=False, temporal_att=True, epochs=10, batch_size=64	392220.73	626.28	443.74	1.44	0.99	1.98	177.28	-0.000088	0.75

Table 5.5: Results of the ablation study of the BiLSTM architecture.

ID	Config.	MSE	RMSE	MAE	MAPE	R^2	CVRMSE	Training Time (s)	Overfit Gap	Loss Ratio
8	units_lstm1=64, units_lstm2=32, dropout_rate=0.0, feature_att=False, temporal_att=False, epochs=10, batch_size=64	657870.71	811.09	603.66	1.93	0.99	2.61	174.05	-0.000101	0.82
9	units_lstm1=64, units_lstm2=32, dropout_rate=0.0, feature_att=True, temporal_att=True, epochs=20, batch_size=32	313162.62	559.61	387.64	1.24	0.99	1.8	732.84	-0.000061	0.78
10	units_lstm1=64, units_lstm2=32, dropout_rate=0.0, feature_att=True, temporal_att=False, epochs=20, batch_size=32	398778.35	631.49	465.23	1.47	0.99	1.81	710.16	0.000017	1.07
11	units_lstm1=64, units_lstm2=32, dropout_rate=0.0, feature_att=False, temporal_att=True, epochs=20, batch_size=32	323979.77	569.19	391.02	1.22	0.99	1.7	695.79	-0.000000	1
12	units_lstm1=64, units_lstm2=32, dropout_rate=0.0, feature_att=False, temporal_att=False, epochs=20, batch_size=32	301410.02	549.01	389.43	1.24	0.99	1.69	675.19	-0.000027	0.88
13	units_lstm1=64, units_lstm2=32, dropout_rate=0.0, feature_att=True, temporal_att=True, epochs=20, batch_size=64	365143.65	604.27	412.78	1.3	0.99	1.87	368.42	-0.000057	0.81
14	units_lstm1=64, units_lstm2=32, dropout_rate=0.0, feature_att=True, temporal_att=False, epochs=20, batch_size=64	327926.67	572.65	400.09	1.27	0.99	1.82	357.95	-0.000051	0.82
15	units_lstm1=64, units_lstm2=32, dropout_rate=0.0, feature_att=False, temporal_att=True, epochs=20, batch_size=64	350281.78	591.85	425.54	1.36	0.99	1.83	350.7	-0.000029	0.89
16	units_lstm1=64, units_lstm2=32, dropout_rate=0.0, feature_att=False, temporal_att=False, epochs=20, batch_size=64	448543.69	669.73	490.61	1.53	0.99	1.99	352.03	-0.000041	0.88
17	units_lstm1=64, units_lstm2=32, dropout_rate=0.0, feature_att=True, temporal_att=True, epochs=50, batch_size=32	262239.84	512.09	352.85	1.09	0.99	1.58	1842.44	-0.000008	0.96
18	units_lstm1=64, units_lstm2=32, dropout_rate=0.0, feature_att=True, temporal_att=False, epochs=50, batch_size=32	228250.23	477.76	326.87	1.04	1	1.52	1800.97	-0.000011	0.94
19	units_lstm1=64, units_lstm2=32, dropout_rate=0.0, feature_att=False, temporal_att=True, epochs=50, batch_size=32	220601.64	469.68	318.46	1.01	1	1.5	1765.21	-0.000019	0.89

Table 5.5: Results of the ablation study of the BiLSTM architecture.

ID	Config.	MSE	RMSE	MAE	MAPE	R^2	CVRMSE	Training Time (s)	Overfit Gap	Loss Ratio
20	units_lstm1=64, units_lstm2=32, dropout_rate=0.0, feature_att=False, temporal_att=False, epochs=50, batch_size=32	256910.12	506.86	350.53	1.1	0.99	1.57	1719.88	0.000013	1.08
21	units_lstm1=64, units_lstm2=32, dropout_rate=0.0, feature_att=True, temporal_att=True, epochs=50, batch_size=64	241977.29	491.91	323.44	1.03	1	1.59	924.98	-0.000043	0.79
22	units_lstm1=64, units_lstm2=32, dropout_rate=0.0, feature_att=True, temporal_att=False, epochs=50, batch_size=64	281701.17	530.76	381.52	1.22	0.99	1.56	921.71	0.000007	1.04
23	units_lstm1=64, units_lstm2=32, dropout_rate=0.0, feature_att=False, temporal_att=True, epochs=50, batch_size=64	244338.78	494.31	343.36	1.11	1	1.56	894.16	-0.000011	0.94
24	units_lstm1=64, units_lstm2=32, dropout_rate=0.0, feature_att=False, temporal_att=False, epochs=50, batch_size=64	245817.34	495.8	346.49	1.1	1	1.55	875.59	-0.000006	0.96
25	units_lstm1=64, units_lstm2=32, dropout_rate=0.2, feature_att=True, temporal_att=True, epochs=10, batch_size=32	494883.47	703.48	489.56	1.56	0.99	2.29	374.65	-0.000130	0.72
26	units_lstm1=64, units_lstm2=32, dropout_rate=0.2, feature_att=True, temporal_att=False, epochs=10, batch_size=32	652436.3	807.74	613.71	2.08	0.99	2.72	365.16	0.000023	1.05
27	units_lstm1=64, units_lstm2=32, dropout_rate=0.2, feature_att=False, temporal_att=True, epochs=10, batch_size=32	570904.92	755.58	552.37	1.7	0.99	1.99	359.4	-0.000001	1
28	units_lstm1=64, units_lstm2=32, dropout_rate=0.2, feature_att=False, temporal_att=False, epochs=10, batch_size=32	770666.91	877.88	659.01	2.06	0.98	2.35	355.57	-0.000038	0.93
29	units_lstm1=64, units_lstm2=32, dropout_rate=0.2, feature_att=True, temporal_att=True, epochs=10, batch_size=64	449889.51	670.74	467.18	1.48	0.99	2.15	192.6	-0.000159	0.66
30	units_lstm1=64, units_lstm2=32, dropout_rate=0.2, feature_att=True, temporal_att=False, epochs=10, batch_size=64	566649.29	752.76	524.97	1.64	0.99	2.33	189.09	-0.000067	0.85
31	units_lstm1=64, units_lstm2=32, dropout_rate=0.2, feature_att=False, temporal_att=True, epochs=10, batch_size=64	422375.57	649.9	463.32	1.47	0.99	2.06	184.55	-0.000146	0.66

Table 5.5: Results of the ablation study of the BiLSTM architecture.

ID	Config.	MSE	RMSE	MAE	MAPE	R^2	CVRMSE	Training Time (s)	Overfit Gap	Loss Ratio
32	units_lstm1=64, units_lstm2=32, dropout_rate=0.2, feature_att=False, temporal_att=False, epochs=10, batch_size=64	650977.61	806.83	584.55	1.87	0.99	2.52	180.87	-0.000237	0.65
33	units_lstm1=64, units_lstm2=32, dropout_rate=0.2, feature_att=True, temporal_att=True, epochs=20, batch_size=32	628605.92	792.85	511.46	1.59	0.99	2.39	746.23	0.000133	1.45
34	units_lstm1=64, units_lstm2=32, dropout_rate=0.2, feature_att=True, temporal_att=False, epochs=20, batch_size=32	511473.08	715.17	555.61	1.72	0.99	1.78	739.72	0.000050	1.17
35	units_lstm1=64, units_lstm2=32, dropout_rate=0.2, feature_att=False, temporal_att=True, epochs=20, batch_size=32	435628.99	660.02	513.57	1.61	0.99	1.64	725.5	0.000042	1.16
36	units_lstm1=64, units_lstm2=32, dropout_rate=0.2, feature_att=False, temporal_att=False, epochs=20, batch_size=32	377932.67	614.76	447.22	1.47	0.99	2.1	697.51	-0.000048	0.84
37	units_lstm1=64, units_lstm2=32, dropout_rate=0.2, feature_att=True, temporal_att=True, epochs=20, batch_size=64	327879.89	572.61	402.12	1.29	0.99	1.83	376.52	-0.000105	0.68
38	units_lstm1=64, units_lstm2=32, dropout_rate=0.2, feature_att=True, temporal_att=False, epochs=20, batch_size=64	557728.94	746.81	517.8	1.63	0.99	2.2	372.87	0.000075	1.24
39	units_lstm1=64, units_lstm2=32, dropout_rate=0.2, feature_att=False, temporal_att=True, epochs=20, batch_size=64	353208.87	594.31	417.93	1.32	0.99	1.89	364.21	-0.000086	0.74
40	units_lstm1=64, units_lstm2=32, dropout_rate=0.2, feature_att=False, temporal_att=False, epochs=20, batch_size=64	493191.61	702.28	506.23	1.63	0.99	2.28	359.2	-0.000134	0.72
41	units_lstm1=64, units_lstm2=32, dropout_rate=0.2, feature_att=True, temporal_att=True, epochs=50, batch_size=32	1089606.05	1043.84	730.88	2.03	0.98	2.21	1877.12	0.000533	3.54
42	units_lstm1=64, units_lstm2=32, dropout_rate=0.2, feature_att=True, temporal_att=False, epochs=50, batch_size=32	283462.61	532.41	376.12	1.15	0.99	1.61	1841.3	0.000000	1
43	units_lstm1=64, units_lstm2=32, dropout_rate=0.2, feature_att=False, temporal_att=True, epochs=50, batch_size=32	265332	515.1	364.17	1.12	0.99	1.55	1805.92	0.000004	1.02

Table 5.5: Results of the ablation study of the BiLSTM architecture.

ID	Config.	MSE	RMSE	MAE	MAPE	R^2	CVRMSE	Training Time (s)	Overfit Gap	Loss Ratio
44	units_lstm1=64, units_lstm2=32, dropout_rate=0.2, feature_att=False, temporal_att=False, epochs=50, batch_size=32	254523.96	504.5	350.94	1.1	1	1.58	1752.24	-0.000023	0.88
45	units_lstm1=64, units_lstm2=32, dropout_rate=0.2, feature_att=True, temporal_att=True, epochs=50, batch_size=64	285424.12	534.25	373.3	1.2	0.99	1.73	943.62	-0.000032	0.86
46	units_lstm1=64, units_lstm2=32, dropout_rate=0.2, feature_att=True, temporal_att=False, epochs=50, batch_size=64	249846.22	499.85	340.5	1.08	1	1.6	903.81	-0.000052	0.77
47	units_lstm1=64, units_lstm2=32, dropout_rate=0.2, feature_att=False, temporal_att=True, epochs=50, batch_size=64	314084.29	560.43	415.37	1.26	0.99	1.54	868.3	0.000024	1.13
48	units_lstm1=64, units_lstm2=32, dropout_rate=0.2, feature_att=False, temporal_att=False, epochs=50, batch_size=64	241639.66	491.57	342.76	1.09	1	1.57	844.81	-0.000053	0.76
49	units_lstm1=128, units_lstm2=64, dropout_rate=0.0, feature_att=True, temporal_att=True, epochs=10, batch_size=32	440429.67	663.65	452.19	1.44	0.99	2.11	369.47	-0.000085	0.78
50	units_lstm1=128, units_lstm2=64, dropout_rate=0.0, feature_att=True, temporal_att=False, epochs=10, batch_size=32	398772.05	631.48	458.06	1.5	0.99	2.06	375.73	-0.000069	0.8
51	units_lstm1=128, units_lstm2=64, dropout_rate=0.0, feature_att=False, temporal_att=True, epochs=10, batch_size=32	292122.23	540.48	372.14	1.2	0.99	1.78	358.22	-0.000095	0.68
52	units_lstm1=128, units_lstm2=64, dropout_rate=0.0, feature_att=False, temporal_att=False, epochs=10, batch_size=32	915443.88	956.79	761.24	2.4	0.98	2.43	356.94	0.000197	1.46
53	units_lstm1=128, units_lstm2=64, dropout_rate=0.0, feature_att=True, temporal_att=True, epochs=10, batch_size=64	385519.87	620.9	434.54	1.38	0.99	2.01	188.72	-0.000051	0.84
54	units_lstm1=128, units_lstm2=64, dropout_rate=0.0, feature_att=True, temporal_att=False, epochs=10, batch_size=64	502824.24	709.1	537.12	1.71	0.99	1.96	190.54	-0.000023	0.94
55	units_lstm1=128, units_lstm2=64, dropout_rate=0.0, feature_att=False, temporal_att=True, epochs=10, batch_size=64	381167.23	617.39	442.62	1.41	0.99	1.97	180.6	-0.000094	0.74

Table 5.5: Results of the ablation study of the BiLSTM architecture.

ID	Config.	MSE	RMSE	MAE	MAPE	R^2	CVRMSE	Training Time (s)	Overfit Gap	Loss Ratio
56	units_lstm1=128, units_lstm2=64, dropout_rate=0.0, feature_att=False, temporal_att=False, epochs=10, batch_size=64	629979.99	793.71	592.38	1.9	0.99	2.45	182.56	-0.000110	0.8
57	units_lstm1=128, units_lstm2=64, dropout_rate=0.0, feature_att=True, temporal_att=True, epochs=20, batch_size=32	415894.7	644.9	473.1	1.55	0.99	1.87	737.43	0.000011	1.04
58	units_lstm1=128, units_lstm2=64, dropout_rate=0.0, feature_att=True, temporal_att=False, epochs=20, batch_size=32	437778.37	661.65	503.64	1.55	0.99	1.72	741.08	0.000044	1.17
59	units_lstm1=128, units_lstm2=64, dropout_rate=0.0, feature_att=False, temporal_att=True, epochs=20, batch_size=32	283300.68	532.26	376.34	1.18	0.99	1.62	705.78	-0.000028	0.87
60	units_lstm1=128, units_lstm2=64, dropout_rate=0.0, feature_att=False, temporal_att=False, epochs=20, batch_size=32	288947.75	537.54	374.17	1.2	0.99	1.73	713.76	-0.000048	0.8
61	units_lstm1=128, units_lstm2=64, dropout_rate=0.0, feature_att=True, temporal_att=True, epochs=20, batch_size=64	395405.08	628.81	440.61	1.43	0.99	2.05	378.17	-0.000146	0.65
62	units_lstm1=128, units_lstm2=64, dropout_rate=0.0, feature_att=True, temporal_att=False, epochs=20, batch_size=64	332521.75	576.65	392.37	1.25	0.99	1.85	381.17	-0.000043	0.84
63	units_lstm1=128, units_lstm2=64, dropout_rate=0.0, feature_att=False, temporal_att=True, epochs=20, batch_size=64	267120.38	516.84	350.1	1.12	0.99	1.68	359.66	-0.000047	0.79
64	units_lstm1=128, units_lstm2=64, dropout_rate=0.0, feature_att=False, temporal_att=False, epochs=20, batch_size=64	512357.42	715.79	532	1.73	0.99	2.29	365.24	-0.000016	0.96
65	units_lstm1=128, units_lstm2=64, dropout_rate=0.0, feature_att=True, temporal_att=True, epochs=50, batch_size=32	254926.97	504.9	345.37	1.11	1	1.6	1832.69	-0.000014	0.93
66	units_lstm1=128, units_lstm2=64, dropout_rate=0.0, feature_att=True, temporal_att=False, epochs=50, batch_size=32	242340.53	492.28	337.8	1.08	1	1.57	1848.43	0.000001	1.01
67	units_lstm1=128, units_lstm2=64, dropout_rate=0.0, feature_att=False, temporal_att=True, epochs=50, batch_size=32	242823.68	492.77	336.88	1.07	1	1.58	1764.34	0.000001	1.01

Table 5.5: Results of the ablation study of the BiLSTM architecture.

ID	Config.	MSE	RMSE	MAE	MAPE	R^2	CVRMSE	Training Time (s)	Overfit Gap	Loss Ratio
68	units_lstm1=128, units_lstm2=64, dropout_rate=0.0, feature_att=False, temporal_att=False, epochs=50, batch_size=32	293252.49	541.53	374.28	1.19	0.99	1.73	1768.88	0.000037	1.22
69	units_lstm1=128, units_lstm2=64, dropout_rate=0.0, feature_att=True, temporal_att=True, epochs=50, batch_size=64	222351.09	471.54	310.01	0.99	1	1.53	936.92	-0.000036	0.81
70	units_lstm1=128, units_lstm2=64, dropout_rate=0.0, feature_att=True, temporal_att=False, epochs=50, batch_size=64	265852.18	515.61	354.86	1.11	0.99	1.6	949.27	-0.000006	0.97
71	units_lstm1=128, units_lstm2=64, dropout_rate=0.0, feature_att=False, temporal_att=True, epochs=50, batch_size=64	290072.02	538.58	395.67	1.24	0.99	1.54	893.97	0.000023	1.13
72	units_lstm1=128, units_lstm2=64, dropout_rate=0.0, feature_att=False, temporal_att=False, epochs=50, batch_size=64	242020.15	491.96	338.95	1.07	1	1.57	911.82	-0.000017	0.91
73	units_lstm1=128, units_lstm2=64, dropout_rate=0.2, feature_att=True, temporal_att=True, epochs=10, batch_size=32	1745438.98	1321.15	1096.37	3.25	0.97	2.57	379.69	0.000756	2.75
74	units_lstm1=128, units_lstm2=64, dropout_rate=0.2, feature_att=True, temporal_att=False, epochs=10, batch_size=32	670897.76	819.08	642.98	2.1	0.99	2.2	379.13	0.000044	1.11
75	units_lstm1=128, units_lstm2=64, dropout_rate=0.2, feature_att=False, temporal_att=True, epochs=10, batch_size=32	600971.81	775.22	586.44	1.78	0.99	1.91	374.71	0.000052	1.15
76	units_lstm1=128, units_lstm2=64, dropout_rate=0.2, feature_att=False, temporal_att=False, epochs=10, batch_size=32	799989.59	894.42	693.8	2.2	0.98	2.41	361.43	-0.000007	0.99
77	units_lstm1=128, units_lstm2=64, dropout_rate=0.2, feature_att=True, temporal_att=True, epochs=10, batch_size=64	461364.08	679.24	499.44	1.59	0.99	2	192.12	-0.000136	0.7
78	units_lstm1=128, units_lstm2=64, dropout_rate=0.2, feature_att=True, temporal_att=False, epochs=10, batch_size=64	525544.04	724.94	524.68	1.68	0.99	2.37	195.78	-0.000115	0.76
79	units_lstm1=128, units_lstm2=64, dropout_rate=0.2, feature_att=False, temporal_att=True, epochs=10, batch_size=64	615099.24	784.28	587.6	1.83	0.99	2.04	183.81	0.000025	1.06

Table 5.5: Results of the ablation study of the BiLSTM architecture.

ID	Config.	MSE	RMSE	MAE	MAPE	R^2	CVRMSE	Training Time (s)	Overfit Gap	Loss Ratio
80	units_lstm1=128, units_lstm2=64, dropout_rate=0.2, feature_att=False, temporal_att=False, epochs=10, batch_size=64	652643.09	807.86	569.06	1.82	0.99	2.66	186.85	-0.000192	0.7
81	units_lstm1=128, units_lstm2=64, dropout_rate=0.2, feature_att=True, temporal_att=True, epochs=20, batch_size=32	976126.58	987.99	761.96	2.36	0.98	2.97	818.9	0.000375	2.29
82	units_lstm1=128, units_lstm2=64, dropout_rate=0.2, feature_att=True, temporal_att=False, epochs=20, batch_size=32	319414.32	565.17	395.98	1.25	0.99	1.81	780.55	-0.000069	0.76
83	units_lstm1=128, units_lstm2=64, dropout_rate=0.2, feature_att=False, temporal_att=True, epochs=20, batch_size=32	312143.8	558.7	390.55	1.22	0.99	1.75	776.72	-0.000030	0.88
84	units_lstm1=128, units_lstm2=64, dropout_rate=0.2, feature_att=False, temporal_att=False, epochs=20, batch_size=32	432813.96	657.89	450.38	1.46	0.99	2.15	769.5	-0.000017	0.94
85	units_lstm1=128, units_lstm2=64, dropout_rate=0.2, feature_att=True, temporal_att=True, epochs=20, batch_size=64	322966.12	568.3	391.41	1.23	0.99	1.8	416.39	-0.000098	0.69
86	units_lstm1=128, units_lstm2=64, dropout_rate=0.2, feature_att=True, temporal_att=False, epochs=20, batch_size=64	437834.74	661.69	481.61	1.55	0.99	2.14	417.55	-0.000014	0.95
87	units_lstm1=128, units_lstm2=64, dropout_rate=0.2, feature_att=False, temporal_att=True, epochs=20, batch_size=64	330154.91	574.59	398.09	1.26	0.99	1.84	398.63	-0.000072	0.76
88	units_lstm1=128, units_lstm2=64, dropout_rate=0.2, feature_att=False, temporal_att=False, epochs=20, batch_size=64	491478	701.05	511.1	1.64	0.99	2.26	397.47	-0.000076	0.82
89	units_lstm1=128, units_lstm2=64, dropout_rate=0.2, feature_att=True, temporal_att=True, epochs=50, batch_size=32	2072971.03	1439.78	1078.28	2.96	0.96	2.95	2018.88	0.001216	7.2
90	units_lstm1=128, units_lstm2=64, dropout_rate=0.2, feature_att=True, temporal_att=False, epochs=50, batch_size=32	234363.1	484.11	333.47	1.05	1	1.53	2004.68	-0.000027	0.86
91	units_lstm1=128, units_lstm2=64, dropout_rate=0.2, feature_att=False, temporal_att=True, epochs=50, batch_size=32	243623.43	493.58	349.94	1.12	1	1.59	1948.21	0.000005	1.03

Table 5.5: Results of the ablation study of the BiLSTM architecture.

ID	Config.	MSE	RMSE	MAE	MAPE	R^2	CVRMSE	Training Time (s)	Overfit Gap	Loss Ratio
92	units_lstm1=128, units_lstm2=64, dropout_rate=0.2, feature_att=False, temporal_att=False, epochs=50, batch_size=32	244835.37	494.81	340.89	1.09	1	1.58	1929.81	-0.000027	0.86
93	units_lstm1=128, units_lstm2=64, dropout_rate=0.2, feature_att=True, temporal_att=True, epochs=50, batch_size=64	275939.04	525.3	363.81	1.16	0.99	1.71	1034.81	-0.000019	0.91
94	units_lstm1=128, units_lstm2=64, dropout_rate=0.2, feature_att=True, temporal_att=False, epochs=50, batch_size=64	256189.98	506.15	343.67	1.1	0.99	1.64	1041.89	-0.000029	0.86
95	units_lstm1=128, units_lstm2=64, dropout_rate=0.2, feature_att=False, temporal_att=True, epochs=50, batch_size=64	231014.92	480.64	330.13	1.05	1	1.56	987.04	-0.000024	0.87
96	units_lstm1=128, units_lstm2=64, dropout_rate=0.2, feature_att=False, temporal_att=False, epochs=50, batch_size=64	276832.85	526.15	370.22	1.18	0.99	1.7	992.06	-0.000038	0.83
97	units_lstm1=256, units_lstm2=128, dropout_rate=0.0, feature_att=True, temporal_att=True, epochs=10, batch_size=32	793111.11	890.57	704.73	2.23	0.98	2.15	489.82	0.000160	1.42
98	units_lstm1=256, units_lstm2=128, dropout_rate=0.0, feature_att=True, temporal_att=False, epochs=10, batch_size=32	417720.26	646.31	450.97	1.41	0.99	1.94	466.23	-0.000060	0.83
99	units_lstm1=256, units_lstm2=128, dropout_rate=0.0, feature_att=False, temporal_att=True, epochs=10, batch_size=32	389648.43	624.22	455.59	1.43	0.99	1.82	462.14	-0.000011	0.96
100	units_lstm1=256, units_lstm2=128, dropout_rate=0.0, feature_att=False, temporal_att=False, epochs=10, batch_size=32	2049194.08	1431.5	1065.99	3.58	0.96	4.6	441.17	-0.000471	0.75
101	units_lstm1=256, units_lstm2=128, dropout_rate=0.0, feature_att=True, temporal_att=True, epochs=10, batch_size=64	478762.72	691.93	496.35	1.59	0.99	2.12	261.39	-0.000071	0.82
102	units_lstm1=256, units_lstm2=128, dropout_rate=0.0, feature_att=True, temporal_att=False, epochs=10, batch_size=64	695249.75	833.82	653.77	2.1	0.99	2.06	250.38	0.000099	1.26
103	units_lstm1=256, units_lstm2=128, dropout_rate=0.0, feature_att=False, temporal_att=True, epochs=10, batch_size=64	421387.56	649.14	471.35	1.51	0.99	1.95	241.46	-0.000039	0.88

Table 5.5: Results of the ablation study of the BiLSTM architecture.

ID	Config.	MSE	RMSE	MAE	MAPE	R^2	CVRMSE	Training Time (s)	Overfit Gap	Loss Ratio
104	units_lstm1=256, units_lstm2=128, dropout_rate=0.0, feature_att=False, temporal_att=False, epochs=10, batch_size=64	740653.72	860.61	625.96	1.95	0.99	2.64	230.22	-0.000066	0.88
105	units_lstm1=256, units_lstm2=128, dropout_rate=0.0, feature_att=True, temporal_att=True, epochs=20, batch_size=32	333612.38	577.59	417.39	1.33	0.99	1.78	1009.85	-0.000045	0.83
106	units_lstm1=256, units_lstm2=128, dropout_rate=0.0, feature_att=True, temporal_att=False, epochs=20, batch_size=32	336834.98	580.37	415.36	1.29	0.99	1.73	931.63	-0.000025	0.9
107	units_lstm1=256, units_lstm2=128, dropout_rate=0.0, feature_att=False, temporal_att=True, epochs=20, batch_size=32	366972.5	605.78	455.73	1.43	0.99	1.64	921.71	0.000027	1.12
108	units_lstm1=256, units_lstm2=128, dropout_rate=0.0, feature_att=False, temporal_att=False, epochs=20, batch_size=32	638056.96	798.78	581.92	1.84	0.99	2.53	880.46	-0.000169	0.72
109	units_lstm1=256, units_lstm2=128, dropout_rate=0.0, feature_att=True, temporal_att=True, epochs=20, batch_size=64	386917.22	622.03	454.07	1.49	0.99	1.9	520.7	-0.000031	0.89
110	units_lstm1=256, units_lstm2=128, dropout_rate=0.0, feature_att=True, temporal_att=False, epochs=20, batch_size=64	574714.12	758.1	597.08	1.92	0.99	1.82	498.1	0.000129	1.49
111	units_lstm1=256, units_lstm2=128, dropout_rate=0.0, feature_att=False, temporal_att=True, epochs=20, batch_size=64	322726.56	568.09	404.31	1.27	0.99	1.79	477.29	-0.000007	0.97
112	units_lstm1=256, units_lstm2=128, dropout_rate=0.0, feature_att=False, temporal_att=False, epochs=20, batch_size=64	4532053.73	2128.86	1573.66	4.85	0.91	6.04	458.23	-0.003883	0.44
113	units_lstm1=256, units_lstm2=128, dropout_rate=0.0, feature_att=True, temporal_att=True, epochs=50, batch_size=32	284370.51	533.26	372.48	1.21	0.99	1.72	2438.48	-0.000010	0.95
114	units_lstm1=256, units_lstm2=128, dropout_rate=0.0, feature_att=True, temporal_att=False, epochs=50, batch_size=32	215745.18	464.48	311.73	0.99	1	1.48	2327.55	-0.000016	0.9
115	units_lstm1=256, units_lstm2=128, dropout_rate=0.0, feature_att=False, temporal_att=True, epochs=50, batch_size=32	237715.12	487.56	339.75	1.07	1	1.55	2351.72	0.000002	1.01

Table 5.5: Results of the ablation study of the BiLSTM architecture.

ID	Config.	MSE	RMSE	MAE	MAPE	R^2	CVRMSE	Training Time (s)	Overfit Gap	Loss Ratio
116	units_lstm1=256, units_lstm2=128, dropout_rate=0.0, feature_att=False, temporal_att=False, epochs=50, batch_size=32	229058.51	478.6	323.49	1.03	1	1.53	2201.54	-0.000016	0.91
117	units_lstm1=256, units_lstm2=128, dropout_rate=0.0, feature_att=True, temporal_att=True, epochs=50, batch_size=64	262292.8	512.15	353.84	1.12	0.99	1.64	1299.58	-0.000026	0.87
118	units_lstm1=256, units_lstm2=128, dropout_rate=0.0, feature_att=True, temporal_att=False, epochs=50, batch_size=64	256834.19	506.79	349.16	1.11	0.99	1.58	1261.94	-0.000024	0.88
119	units_lstm1=256, units_lstm2=128, dropout_rate=0.0, feature_att=False, temporal_att=True, epochs=50, batch_size=64	302831.11	550.3	405.24	1.26	0.99	1.56	1196.09	0.000032	1.19
120	units_lstm1=256, units_lstm2=128, dropout_rate=0.0, feature_att=False, temporal_att=False, epochs=50, batch_size=64	478288.25	691.58	522.51	1.66	0.99	1.87	1177.12	0.000002	1.01
121	units_lstm1=256, units_lstm2=128, dropout_rate=0.2, feature_att=True, temporal_att=True, epochs=10, batch_size=32	3708250.08	1925.68	1425.31	3.9	0.93	3.65	497.25	0.002085	5.71
122	units_lstm1=256, units_lstm2=128, dropout_rate=0.2, feature_att=True, temporal_att=False, epochs=10, batch_size=32	602692.68	776.33	576.69	1.76	0.99	2.1	479.78	-0.000008	0.98
123	units_lstm1=256, units_lstm2=128, dropout_rate=0.2, feature_att=False, temporal_att=True, epochs=10, batch_size=32	607483.44	779.41	588.69	1.77	0.99	1.92	475.11	0.000082	1.25
124	units_lstm1=256, units_lstm2=128, dropout_rate=0.2, feature_att=False, temporal_att=False, epochs=10, batch_size=32	8832993.66	2972.04	2304.06	6.88	0.83	8.18	435.62	-0.017074	0.26
125	units_lstm1=256, units_lstm2=128, dropout_rate=0.2, feature_att=True, temporal_att=True, epochs=10, batch_size=64	680101.61	824.68	598.41	1.82	0.99	2.45	257.49	-0.000008	0.98
126	units_lstm1=256, units_lstm2=128, dropout_rate=0.2, feature_att=True, temporal_att=False, epochs=10, batch_size=64	514347	717.18	513.56	1.62	0.99	2.21	246.42	-0.000083	0.81
127	units_lstm1=256, units_lstm2=128, dropout_rate=0.2, feature_att=False, temporal_att=True, epochs=10, batch_size=64	512250.75	715.72	516.84	1.66	0.99	2.3	237.49	-0.000040	0.9

Table 5.5: Results of the ablation study of the BiLSTM architecture.

ID	Config.	MSE	RMSE	MAE	MAPE	R^2	CVRMSE	Training Time (s)	Overfit Gap	Loss Ratio
128	units_lstm1=256, units_lstm2=128, dropout_rate=0.2, feature_att=False, temporal_att=False, epochs=10, batch_size=64	762581.39	873.26	639.09	2.08	0.99	2.78	225.5	-0.000205	0.72
129	units_lstm1=256, units_lstm2=128, dropout_rate=0.2, feature_att=True, temporal_att=True, epochs=20, batch_size=32	1720488.37	1311.67	971.17	2.69	0.97	2.72	950.54	0.000883	4.05
130	units_lstm1=256, units_lstm2=128, dropout_rate=0.2, feature_att=True, temporal_att=False, epochs=20, batch_size=32	304694.05	551.99	383.38	1.23	0.99	1.79	907.39	-0.000077	0.73
131	units_lstm1=256, units_lstm2=128, dropout_rate=0.2, feature_att=False, temporal_att=True, epochs=20, batch_size=32	264327.96	514.13	357.32	1.14	0.99	1.62	924.88	-0.000060	0.75
132	units_lstm1=256, units_lstm2=128, dropout_rate=0.2, feature_att=False, temporal_att=False, epochs=20, batch_size=32	762661.65	873.31	709.88	2.32	0.99	2.08	875.95	0.000179	1.53
133	units_lstm1=256, units_lstm2=128, dropout_rate=0.2, feature_att=True, temporal_att=True, epochs=20, batch_size=64	1347641.53	1160.88	839.24	2.4	0.97	2.98	514.65	0.000599	2.88
134	units_lstm1=256, units_lstm2=128, dropout_rate=0.2, feature_att=True, temporal_att=False, epochs=20, batch_size=64	501128.56	707.9	492.63	1.55	0.99	2.14	499.03	-0.000115	0.75
135	units_lstm1=256, units_lstm2=128, dropout_rate=0.2, feature_att=False, temporal_att=True, epochs=20, batch_size=64	404774.37	636.22	430.95	1.36	0.99	1.97	481.34	-0.000019	0.94
136	units_lstm1=256, units_lstm2=128, dropout_rate=0.2, feature_att=False, temporal_att=False, epochs=20, batch_size=64	642272.07	801.42	613.5	1.93	0.99	2.15	456.42	-0.000029	0.94
137	units_lstm1=256, units_lstm2=128, dropout_rate=0.2, feature_att=True, temporal_att=True, epochs=50, batch_size=32	285326.07	534.16	381.49	1.21	0.99	1.58	2429.79	-0.000005	0.97
138	units_lstm1=256, units_lstm2=128, dropout_rate=0.2, feature_att=True, temporal_att=False, epochs=50, batch_size=32	203213.47	450.79	303.37	0.96	1	1.45	2322.59	-0.000042	0.77
139	units_lstm1=256, units_lstm2=128, dropout_rate=0.2, feature_att=False, temporal_att=True, epochs=50, batch_size=32	227074.61	476.52	323.57	1.03	1	1.53	2306.54	-0.000010	0.94

Table 5.5: Results of the ablation study of the BiLSTM architecture.

ID	Config.	MSE	RMSE	MAE	MAPE	R^2	CVRMSE	Training Time (s)	Overfit Gap	Loss Ratio
140	units_lstm1=256, units_lstm2=128, dropout_rate=0.2, feature_att=False, temporal_att=False, epochs=50, batch_size=32	255076.78	505.05	347.23	1.11	1	1.64	2193.36	-0.000028	0.86
141	units_lstm1=256, units_lstm2=128, dropout_rate=0.2, feature_att=True, temporal_att=True, epochs=50, batch_size=64	1174778.48	1083.87	745.49	2.04	0.98	2.31	1291.21	0.000600	4
142	units_lstm1=256, units_lstm2=128, dropout_rate=0.2, feature_att=True, temporal_att=False, epochs=50, batch_size=64	397092.37	630.15	458.34	1.38	0.99	1.72	1249	0.000045	1.2
143	units_lstm1=256, units_lstm2=128, dropout_rate=0.2, feature_att=False, temporal_att=True, epochs=50, batch_size=64	263299.6	513.13	369.57	1.19	0.99	1.63	1469.69	0.000004	1.02
144	units_lstm1=256, units_lstm2=128, dropout_rate=0.2, feature_att=False, temporal_att=False, epochs=50, batch_size=64	373514.86	611.16	434.6	1.43	0.99	2	1228.77	-0.000060	0.81

Table 5.6: Results of the ablation study of the Transformers architecture.

ID	Config.	MSE	RMSE	MAE	MAPE	R^2	CVRMSE	Training Time (s)	Overfit Gap	Loss Ratio
1	architecture=encoder, d_model=32, num_heads=2, ff_dim=64, num_layers=1, dropout_rate=0.0, epochs=20, batch_size=64	414109.1	643.51	458.46	1.46	0.99	2.07	55.4	-0.000057	0.83
2	architecture=encoder, d_model=32, num_heads=2, ff_dim=64, num_layers=1, dropout_rate=0.0, epochs=20, batch_size=128	503719.28	709.73	489.33	1.58	0.99	2.29	31.8	-0.000070	0.83
3	architecture=encoder, d_model=32, num_heads=2, ff_dim=64, num_layers=1, dropout_rate=0.0, epochs=50, batch_size=64	349932.77	591.55	429.84	1.39	0.99	1.76	123.2	0.000003	1.01
4	architecture=encoder, d_model=32, num_heads=2, ff_dim=64, num_layers=1, dropout_rate=0.0, epochs=50, batch_size=128	354360.79	595.28	420.04	1.35	0.99	1.91	67.19	-0.000031	0.89

Table 5.6: Results of the ablation study of the Transformers architecture.

ID	Config.	MSE	RMSE	MAE	MAPE	R^2	CVRMSE	Training Time (s)	Overfit Gap	Loss Ratio
5	architecture=encoder, d_model=32, num_heads=2, ff_dim=64, num_layers=1, dropout_rate=0.2, epochs=20, batch_size=64	3441745.58	1855.19	1460.6	4.08	0.93	3.18	59.36	0.001866	4.9
6	architecture=encoder, d_model=32, num_heads=2, ff_dim=64, num_layers=1, dropout_rate=0.2, epochs=20, batch_size=128	6729769.3	2594.18	2349.69	7	0.87	2.52	34.59	0.004005	7.89
7	architecture=encoder, d_model=32, num_heads=2, ff_dim=64, num_layers=1, dropout_rate=0.2, epochs=50, batch_size=64	2412466.97	1553.21	1363.16	4	0.95	2.29	129.21	0.001345	5.51
8	architecture=encoder, d_model=32, num_heads=2, ff_dim=64, num_layers=1, dropout_rate=0.2, epochs=50, batch_size=128	6084200.99	2466.62	2131.3	6.1	0.88	2.94	76.4	0.003815	12.53
9	architecture=encoder, d_model=32, num_heads=2, ff_dim=64, num_layers=2, dropout_rate=0.0, epochs=20, batch_size=64	525853.99	725.16	526.21	1.68	0.99	2.18	69.4	0.000035	1.11
10	architecture=encoder, d_model=32, num_heads=2, ff_dim=64, num_layers=2, dropout_rate=0.0, epochs=20, batch_size=128	473477.89	688.1	496.96	1.58	0.99	2.16	40.07	-0.000106	0.75
11	architecture=encoder, d_model=32, num_heads=2, ff_dim=64, num_layers=2, dropout_rate=0.0, epochs=50, batch_size=64	270964.21	520.54	358.53	1.15	0.99	1.69	151.81	-0.000036	0.84
12	architecture=encoder, d_model=32, num_heads=2, ff_dim=64, num_layers=2, dropout_rate=0.0, epochs=50, batch_size=128	396799.87	629.92	467.88	1.5	0.99	1.81	84.2	0.000027	1.11
13	architecture=encoder, d_model=32, num_heads=2, ff_dim=64, num_layers=2, dropout_rate=0.2, epochs=20, batch_size=64	4682441.4	2163.89	1789.52	5.03	0.91	3.34	73.7	0.002792	8
14	architecture=encoder, d_model=32, num_heads=2, ff_dim=64, num_layers=2, dropout_rate=0.2, epochs=20, batch_size=128	7171632.86	2677.99	2391.45	7.01	0.86	2.88	43.85	0.004279	8.03

Table 5.6: Results of the ablation study of the Transformers architecture.

ID	Config.	MSE	RMSE	MAE	MAPE	R^2	CVRMSE	Training Time (s)	Overfit Gap	Loss Ratio
15	architecture=encoder, d_model=32, num_heads=2, ff_dim=64, num_layers=2, dropout_rate=0.2, epochs=50, batch_size=64	3313469.08	1820.29	1532.01	4.36	0.94	2.77	162.99	0.001951	7.37
16	architecture=encoder, d_model=32, num_heads=2, ff_dim=64, num_layers=2, dropout_rate=0.2, epochs=50, batch_size=128	5742742.03	2396.4	2022.02	5.71	0.89	3.37	93.97	0.003592	12.19
17	architecture=encoder, d_model=32, num_heads=2, ff_dim=128, num_layers=1, dropout_rate=0.0, epochs=20, batch_size=64	458992.77	677.49	485.48	1.55	0.99	2.12	59.2	-0.000002	0.99
18	architecture=encoder, d_model=32, num_heads=2, ff_dim=128, num_layers=1, dropout_rate=0.0, epochs=20, batch_size=128	448793.65	669.92	496.47	1.58	0.99	2.03	33.18	-0.000087	0.78
19	architecture=encoder, d_model=32, num_heads=2, ff_dim=128, num_layers=1, dropout_rate=0.0, epochs=50, batch_size=64	305925.68	553.11	377.27	1.2	0.99	1.77	126.65	-0.000023	0.9
20	architecture=encoder, d_model=32, num_heads=2, ff_dim=128, num_layers=1, dropout_rate=0.0, epochs=50, batch_size=128	301926.85	549.48	382.5	1.22	0.99	1.78	68.83	-0.000051	0.8
21	architecture=encoder, d_model=32, num_heads=2, ff_dim=128, num_layers=1, dropout_rate=0.2, epochs=20, batch_size=64	6139650.61	2477.83	2127.52	6.09	0.88	3.06	62.23	0.003701	8.68
22	architecture=encoder, d_model=32, num_heads=2, ff_dim=128, num_layers=1, dropout_rate=0.2, epochs=20, batch_size=128	5008082.53	2237.87	1879.64	5.34	0.9	3.14	34.33	0.002937	7.18
23	architecture=encoder, d_model=32, num_heads=2, ff_dim=128, num_layers=1, dropout_rate=0.2, epochs=50, batch_size=64	2859435.47	1690.99	1341.16	3.75	0.94	3.01	136.03	0.001663	6.82
24	architecture=encoder, d_model=32, num_heads=2, ff_dim=128, num_layers=1, dropout_rate=0.2, epochs=50, batch_size=128	4825924.04	2196.8	1815.41	5.1	0.91	3.31	75.59	0.002956	9.9

Table 5.6: Results of the ablation study of the Transformers architecture.

ID	Config.	MSE	RMSE	MAE	MAPE	R^2	CVRMSE	Training Time (s)	Overfit Gap	Loss Ratio
25	architecture=encoder, d_model=32, num_heads=2, ff_dim=128, num_layers=2, dropout_rate=0.0, epochs=20, batch_size=64	312672.47	559.17	395.32	1.26	0.99	1.77	69.01	-0.000080	0.73
26	architecture=encoder, d_model=32, num_heads=2, ff_dim=128, num_layers=2, dropout_rate=0.0, epochs=20, batch_size=128	479944.64	692.78	523.28	1.71	0.99	2.07	39.64	-0.000059	0.85
27	architecture=encoder, d_model=32, num_heads=2, ff_dim=128, num_layers=2, dropout_rate=0.0, epochs=50, batch_size=64	269971.33	519.59	372.35	1.23	0.99	1.67	159.97	-0.000018	0.91
28	architecture=encoder, d_model=32, num_heads=2, ff_dim=128, num_layers=2, dropout_rate=0.0, epochs=50, batch_size=128	386773.65	621.91	466.04	1.54	0.99	1.88	84.9	0.000021	1.09
29	architecture=encoder, d_model=32, num_heads=2, ff_dim=128, num_layers=2, dropout_rate=0.2, epochs=20, batch_size=64	5179565.15	2275.87	1890.63	5.33	0.9	3.15	73.1	0.003119	8.61
30	architecture=encoder, d_model=32, num_heads=2, ff_dim=128, num_layers=2, dropout_rate=0.2, epochs=20, batch_size=128	8854808.85	2975.7	2659.63	7.8	0.83	2.84	43.82	0.005463	10.58
31	architecture=encoder, d_model=32, num_heads=2, ff_dim=128, num_layers=2, dropout_rate=0.2, epochs=50, batch_size=64	5314382.6	2305.29	1871.68	5.21	0.9	3.35	165.29	0.003336	12.68
32	architecture=encoder, d_model=32, num_heads=2, ff_dim=128, num_layers=2, dropout_rate=0.2, epochs=50, batch_size=128	3219818.56	1794.39	1439.51	4.02	0.94	3.07	96.56	0.001908	7.65
33	architecture=encoder, d_model=32, num_heads=4, ff_dim=64, num_layers=1, dropout_rate=0.0, epochs=20, batch_size=64	389512.46	624.11	437.53	1.4	0.99	1.98	60.99	-0.000060	0.81
34	architecture=encoder, d_model=32, num_heads=4, ff_dim=64, num_layers=1, dropout_rate=0.0, epochs=20, batch_size=128	434194.56	658.93	467.79	1.49	0.99	2.11	35.51	-0.000117	0.72

Table 5.6: Results of the ablation study of the Transformers architecture.

ID	Config.	MSE	RMSE	MAE	MAPE	R^2	CVRMSE	Training Time (s)	Overfit Gap	Loss Ratio
35	architecture=encoder, d_model=32, num_heads=4, ff_dim=64, num_layers=1, dropout_rate=0.0, epochs=50, batch_size=64	393542.9	627.33	464.83	1.47	0.99	1.72	131.2	0.000038	1.17
36	architecture=encoder, d_model=32, num_heads=4, ff_dim=64, num_layers=1, dropout_rate=0.0, epochs=50, batch_size=128	353643.21	594.68	407.95	1.3	0.99	1.89	71.54	-0.000012	0.95
37	architecture=encoder, d_model=32, num_heads=4, ff_dim=64, num_layers=1, dropout_rate=0.2, epochs=20, batch_size=64	5399175.01	2323.61	1963.43	5.58	0.89	3.17	62.56	0.003225	8.11
38	architecture=encoder, d_model=32, num_heads=4, ff_dim=64, num_layers=1, dropout_rate=0.2, epochs=20, batch_size=128	4529804.77	2128.33	1816.12	5.23	0.91	2.91	36.53	0.002537	5.62
39	architecture=encoder, d_model=32, num_heads=4, ff_dim=64, num_layers=1, dropout_rate=0.2, epochs=50, batch_size=64	2728943.08	1651.95	1349.96	3.83	0.95	2.65	136.12	0.001536	5.75
40	architecture=encoder, d_model=32, num_heads=4, ff_dim=64, num_layers=1, dropout_rate=0.2, epochs=50, batch_size=128	4642549.95	2154.66	1774.31	4.98	0.91	3.29	76.17	0.002815	9.09
41	architecture=encoder, d_model=32, num_heads=4, ff_dim=64, num_layers=2, dropout_rate=0.0, epochs=20, batch_size=64	420185.01	648.22	476.3	1.49	0.99	1.85	73.41	-0.000031	0.9
42	architecture=encoder, d_model=32, num_heads=4, ff_dim=64, num_layers=2, dropout_rate=0.0, epochs=20, batch_size=128	452296.8	672.53	497.31	1.62	0.99	2.05	42.43	-0.000044	0.87
43	architecture=encoder, d_model=32, num_heads=4, ff_dim=64, num_layers=2, dropout_rate=0.0, epochs=50, batch_size=64	262604.53	512.45	355.05	1.12	0.99	1.6	156.75	-0.000036	0.83
44	architecture=encoder, d_model=32, num_heads=4, ff_dim=64, num_layers=2, dropout_rate=0.0, epochs=50, batch_size=128	298825.93	546.65	384.99	1.22	0.99	1.76	86	-0.000045	0.82

Table 5.6: Results of the ablation study of the Transformers architecture.

ID	Config.	MSE	RMSE	MAE	MAPE	R^2	CVRMSE	Training Time (s)	Overfit Gap	Loss Ratio
45	architecture=encoder, d_model=32, num_heads=4, ff_dim=64, num_layers=2, dropout_rate=0.2, epochs=20, batch_size=64	9133469.93	3022.16	2621.68	7.49	0.82	3.38	73.15	0.005825	15.62
46	architecture=encoder, d_model=32, num_heads=4, ff_dim=64, num_layers=2, dropout_rate=0.2, epochs=20, batch_size=128	8917477.69	2986.21	2659.86	7.75	0.83	3	44.47	0.005546	11.45
47	architecture=encoder, d_model=32, num_heads=4, ff_dim=64, num_layers=2, dropout_rate=0.2, epochs=50, batch_size=64	3857032.9	1963.93	1600.66	4.49	0.92	2.86	173.92	0.002355	9.62
48	architecture=encoder, d_model=32, num_heads=4, ff_dim=64, num_layers=2, dropout_rate=0.2, epochs=50, batch_size=128	4074733.28	2018.6	1644.62	4.61	0.92	3.04	97.11	0.002438	8.2
49	architecture=encoder, d_model=32, num_heads=4, ff_dim=128, num_layers=1, dropout_rate=0.0, epochs=20, batch_size=64	436749.5	660.87	493.99	1.52	0.99	1.85	59.32	-0.000015	0.95
50	architecture=encoder, d_model=32, num_heads=4, ff_dim=128, num_layers=1, dropout_rate=0.0, epochs=20, batch_size=128	422404.94	649.93	488.13	1.61	0.99	2.03	32.88	-0.000047	0.86
51	architecture=encoder, d_model=32, num_heads=4, ff_dim=128, num_layers=1, dropout_rate=0.0, epochs=50, batch_size=64	271292.1	520.86	358.92	1.15	0.99	1.67	128.73	-0.000033	0.85
52	architecture=encoder, d_model=32, num_heads=4, ff_dim=128, num_layers=1, dropout_rate=0.0, epochs=50, batch_size=128	279851.77	529.01	366.11	1.16	0.99	1.7	70.1	-0.000045	0.81
53	architecture=encoder, d_model=32, num_heads=4, ff_dim=128, num_layers=1, dropout_rate=0.2, epochs=20, batch_size=64	4417044.73	2101.68	1825.97	5.27	0.91	2.66	59.36	0.002600	7.35
54	architecture=encoder, d_model=32, num_heads=4, ff_dim=128, num_layers=1, dropout_rate=0.2, epochs=20, batch_size=128	6016407.27	2452.84	2114.87	6.08	0.88	2.92	34.56	0.003600	8.19

Table 5.6: Results of the ablation study of the Transformers architecture.

ID	Config.	MSE	RMSE	MAE	MAPE	R^2	CVRMSE	Training Time (s)	Overfit Gap	Loss Ratio
55	architecture=encoder, d_model=32, num_heads=4, ff_dim=128, num_layers=1, dropout_rate=0.2, epochs=50, batch_size=64	6257213.26	2501.44	2102.94	5.92	0.88	3.33	132.52	0.003982	15.16
56	architecture=encoder, d_model=32, num_heads=4, ff_dim=128, num_layers=1, dropout_rate=0.2, epochs=50, batch_size=128	3896633.26	1973.99	1661.53	4.72	0.92	3.32	74.42	0.002351	8.71
57	architecture=encoder, d_model=32, num_heads=4, ff_dim=128, num_layers=2, dropout_rate=0.0, epochs=20, batch_size=64	403447.45	635.18	473.24	1.51	0.99	1.87	70.81	-0.000019	0.93
58	architecture=encoder, d_model=32, num_heads=4, ff_dim=128, num_layers=2, dropout_rate=0.0, epochs=20, batch_size=128	463722.45	680.97	494.25	1.57	0.99	2.07	41.07	-0.000062	0.84
59	architecture=encoder, d_model=32, num_heads=4, ff_dim=128, num_layers=2, dropout_rate=0.0, epochs=50, batch_size=64	253679.63	503.67	347.78	1.11	1	1.58	164.42	-0.000021	0.89
60	architecture=encoder, d_model=32, num_heads=4, ff_dim=128, num_layers=2, dropout_rate=0.0, epochs=50, batch_size=128	305792.88	552.99	395.69	1.26	0.99	1.73	87.53	-0.000009	0.96
61	architecture=encoder, d_model=32, num_heads=4, ff_dim=128, num_layers=2, dropout_rate=0.2, epochs=20, batch_size=64	4020714.17	2005.17	1644.58	4.64	0.92	3.64	74.93	0.002363	7.28
62	architecture=encoder, d_model=32, num_heads=4, ff_dim=128, num_layers=2, dropout_rate=0.2, epochs=20, batch_size=128	8661933	2943.12	2577.88	7.43	0.83	3.3	44.96	0.005368	11.07
63	architecture=encoder, d_model=32, num_heads=4, ff_dim=128, num_layers=2, dropout_rate=0.2, epochs=50, batch_size=64	5122220.06	2263.23	1861.52	5.2	0.9	3.43	169.97	0.003226	13.18
64	architecture=encoder, d_model=32, num_heads=4, ff_dim=128, num_layers=2, dropout_rate=0.2, epochs=50, batch_size=128	4888774.31	2211.06	1885.92	5.37	0.9	3.04	95.78	0.003029	11

Table 5.6: Results of the ablation study of the Transformers architecture.

ID	Config.	MSE	RMSE	MAE	MAPE	R^2	CVRMSE	Training Time (s)	Overfit Gap	Loss Ratio
65	architecture=encoder, d_model=64, num_heads=2, ff_dim=64, num_layers=1, dropout_rate=0.0, epochs=20, batch_size=64	442077.23	664.89	486.7	1.58	0.99	1.98	64.61	-0.000019	0.94
66	architecture=encoder, d_model=64, num_heads=2, ff_dim=64, num_layers=1, dropout_rate=0.0, epochs=20, batch_size=128	601075.78	775.29	591.38	1.89	0.99	2.16	37.29	-0.000055	0.88
67	architecture=encoder, d_model=64, num_heads=2, ff_dim=64, num_layers=1, dropout_rate=0.0, epochs=50, batch_size=64	421154.77	648.96	504.18	1.61	0.99	1.66	130.69	0.000073	1.34
68	architecture=encoder, d_model=64, num_heads=2, ff_dim=64, num_layers=1, dropout_rate=0.0, epochs=50, batch_size=128	359080.6	599.23	429.74	1.35	0.99	1.75	70.89	-0.000025	0.91
69	architecture=encoder, d_model=64, num_heads=2, ff_dim=64, num_layers=1, dropout_rate=0.2, epochs=20, batch_size=64	3330193.92	1824.88	1622.09	4.81	0.93	2.23	64.86	0.001804	4.88
70	architecture=encoder, d_model=64, num_heads=2, ff_dim=64, num_layers=1, dropout_rate=0.2, epochs=20, batch_size=128	1168239.94	1080.85	883.24	2.95	0.98	2.8	39.95	0.000030	1.04
71	architecture=encoder, d_model=64, num_heads=2, ff_dim=64, num_layers=1, dropout_rate=0.2, epochs=50, batch_size=64	2031856.26	1425.43	1165.95	3.37	0.96	2.53	141.8	0.001030	3.91
72	architecture=encoder, d_model=64, num_heads=2, ff_dim=64, num_layers=1, dropout_rate=0.2, epochs=50, batch_size=128	1879207.27	1370.84	1108.87	3.19	0.96	2.42	77.97	0.000912	3.48
73	architecture=encoder, d_model=64, num_heads=2, ff_dim=64, num_layers=2, dropout_rate=0.0, epochs=20, batch_size=64	309881.51	556.67	387.97	1.25	0.99	1.82	73.22	-0.000098	0.68
74	architecture=encoder, d_model=64, num_heads=2, ff_dim=64, num_layers=2, dropout_rate=0.0, epochs=20, batch_size=128	533014.51	730.08	556.93	1.74	0.99	2.09	40.25	0.000025	1.07

Table 5.6: Results of the ablation study of the Transformers architecture.

ID	Config.	MSE	RMSE	MAE	MAPE	R^2	CVRMSE	Training Time (s)	Overfit Gap	Loss Ratio
75	architecture=encoder, d_model=64, num_heads=2, ff_dim=64, num_layers=2, dropout_rate=0.0, epochs=50, batch_size=64	381409.23	617.58	480.08	1.57	0.99	1.66	156.04	0.000049	1.23
76	architecture=encoder, d_model=64, num_heads=2, ff_dim=64, num_layers=2, dropout_rate=0.0, epochs=50, batch_size=128	607918.94	779.69	630.01	1.99	0.99	1.82	87.22	0.000141	1.52
77	architecture=encoder, d_model=64, num_heads=2, ff_dim=64, num_layers=2, dropout_rate=0.2, epochs=20, batch_size=64	3197345.62	1788.11	1421.83	3.98	0.94	2.79	76.13	0.001763	5.24
78	architecture=encoder, d_model=64, num_heads=2, ff_dim=64, num_layers=2, dropout_rate=0.2, epochs=20, batch_size=128	2325625.59	1525	1287.67	3.77	0.95	2.27	47.09	0.001037	2.89
79	architecture=encoder, d_model=64, num_heads=2, ff_dim=64, num_layers=2, dropout_rate=0.2, epochs=50, batch_size=64	1772781.44	1331.46	1026.08	2.89	0.97	2.9	162.09	0.000916	4.13
80	architecture=encoder, d_model=64, num_heads=2, ff_dim=64, num_layers=2, dropout_rate=0.2, epochs=50, batch_size=128	4760496.29	2181.86	1834.18	5.2	0.91	2.81	91.47	0.002906	9.59
81	architecture=encoder, d_model=64, num_heads=2, ff_dim=128, num_layers=1, dropout_rate=0.0, epochs=20, batch_size=64	650258.86	806.39	615.19	1.87	0.99	2.04	60.43	0.000093	1.26
82	architecture=encoder, d_model=64, num_heads=2, ff_dim=128, num_layers=1, dropout_rate=0.0, epochs=20, batch_size=128	552621.97	743.39	562.69	1.81	0.99	2.42	38.11	-0.000099	0.79
83	architecture=encoder, d_model=64, num_heads=2, ff_dim=128, num_layers=1, dropout_rate=0.0, epochs=50, batch_size=64	350935.1	592.4	436.25	1.37	0.99	1.65	124.98	0.000025	1.12
84	architecture=encoder, d_model=64, num_heads=2, ff_dim=128, num_layers=1, dropout_rate=0.0, epochs=50, batch_size=128	326408.31	571.32	401.28	1.28	0.99	1.86	69.37	-0.000042	0.84

Table 5.6: Results of the ablation study of the Transformers architecture.

ID	Config.	MSE	RMSE	MAE	MAPE	R^2	CVRMSE	Training Time (s)	Overfit Gap	Loss Ratio
85	architecture=encoder, d_model=64, num_heads=2, ff_dim=128, num_layers=1, dropout_rate=0.2, epochs=20, batch_size=64	3079735.37	1754.92	1467.02	4.2	0.94	2.45	65.48	0.001684	5.07
86	architecture=encoder, d_model=64, num_heads=2, ff_dim=128, num_layers=1, dropout_rate=0.2, epochs=20, batch_size=128	4583852.5	2140.99	1709.54	4.78	0.91	3.44	37.2	0.002490	4.96
87	architecture=encoder, d_model=64, num_heads=2, ff_dim=128, num_layers=1, dropout_rate=0.2, epochs=50, batch_size=64	3452600.02	1858.12	1503.96	4.21	0.93	2.76	135.48	0.002061	8.06
88	architecture=encoder, d_model=64, num_heads=2, ff_dim=128, num_layers=1, dropout_rate=0.2, epochs=50, batch_size=128	1350120.85	1161.95	903.79	2.61	0.97	2.68	99.62	0.000583	2.72
89	architecture=encoder, d_model=64, num_heads=2, ff_dim=128, num_layers=2, dropout_rate=0.0, epochs=20, batch_size=64	380441.46	616.8	454.69	1.49	0.99	1.87	76.95	-0.000056	0.82
90	architecture=encoder, d_model=64, num_heads=2, ff_dim=128, num_layers=2, dropout_rate=0.0, epochs=20, batch_size=128	484104.56	695.78	513.55	1.65	0.99	2.16	46.35	-0.000125	0.72
91	architecture=encoder, d_model=64, num_heads=2, ff_dim=128, num_layers=2, dropout_rate=0.0, epochs=50, batch_size=64	251032.93	501.03	347.43	1.12	1	1.62	152.94	-0.000029	0.86
92	architecture=encoder, d_model=64, num_heads=2, ff_dim=128, num_layers=2, dropout_rate=0.0, epochs=50, batch_size=128	427018	653.47	487.27	1.58	0.99	1.84	83.48	0.000034	1.13
93	architecture=encoder, d_model=64, num_heads=2, ff_dim=128, num_layers=2, dropout_rate=0.2, epochs=20, batch_size=64	2941443.53	1715.06	1411.08	4.01	0.94	3.08	79.58	0.001586	4.79
94	architecture=encoder, d_model=64, num_heads=2, ff_dim=128, num_layers=2, dropout_rate=0.2, epochs=20, batch_size=128	6302557.66	2510.49	2049.63	5.72	0.88	3.76	48.06	0.003727	7.53

Table 5.6: Results of the ablation study of the Transformers architecture.

ID	Config.	MSE	RMSE	MAE	MAPE	R^2	CVRMSE	Training Time (s)	Overfit Gap	Loss Ratio
95	architecture=encoder, d_model=64, num_heads=2, ff_dim=128, num_layers=2, dropout_rate=0.2, epochs=50, batch_size=64	4505598.55	2122.64	1768.96	4.96	0.91	2.97	166.29	0.002816	12.08
96	architecture=encoder, d_model=64, num_heads=2, ff_dim=128, num_layers=2, dropout_rate=0.2, epochs=50, batch_size=128	4747382.92	2178.85	1836.96	5.2	0.91	3.02	96.14	0.002910	10
97	architecture=encoder, d_model=64, num_heads=4, ff_dim=64, num_layers=1, dropout_rate=0.0, epochs=20, batch_size=64	404028.07	635.63	457.11	1.47	0.99	2.1	65.35	-0.000074	0.79
98	architecture=encoder, d_model=64, num_heads=4, ff_dim=64, num_layers=1, dropout_rate=0.0, epochs=20, batch_size=128	713448.78	844.66	681.07	2.27	0.99	2.22	38.42	0.000029	1.06
99	architecture=encoder, d_model=64, num_heads=4, ff_dim=64, num_layers=1, dropout_rate=0.0, epochs=50, batch_size=64	373744.83	611.35	433.17	1.39	0.99	1.77	131.49	0.000015	1.06
100	architecture=encoder, d_model=64, num_heads=4, ff_dim=64, num_layers=1, dropout_rate=0.0, epochs=50, batch_size=128	452850.82	672.94	522.34	1.64	0.99	1.73	71.5	0.000053	1.21
101	architecture=encoder, d_model=64, num_heads=4, ff_dim=64, num_layers=1, dropout_rate=0.2, epochs=20, batch_size=64	1234913.5	1111.27	963.3	3.02	0.98	1.97	64.69	0.000381	1.83
102	architecture=encoder, d_model=64, num_heads=4, ff_dim=64, num_layers=1, dropout_rate=0.2, epochs=20, batch_size=128	1658617.85	1287.87	1138.38	3.61	0.97	2.12	38.6	0.000520	1.85
103	architecture=encoder, d_model=64, num_heads=4, ff_dim=64, num_layers=1, dropout_rate=0.2, epochs=50, batch_size=64	1813258.1	1346.57	1062.19	3	0.96	2.7	140.35	0.000910	3.79
104	architecture=encoder, d_model=64, num_heads=4, ff_dim=64, num_layers=1, dropout_rate=0.2, epochs=50, batch_size=128	2360241.79	1536.31	1217.45	3.42	0.95	2.73	76.29	0.001285	4.97

Table 5.6: Results of the ablation study of the Transformers architecture.

ID	Config.	MSE	RMSE	MAE	MAPE	R^2	CVRMSE	Training Time (s)	Overfit Gap	Loss Ratio
105	architecture=encoder, d_model=64, num_heads=4, ff_dim=64, num_layers=2, dropout_rate=0.0, epochs=20, batch_size=64	391504.81	625.7	442.33	1.42	0.99	1.97	69.29	-0.000054	0.83
106	architecture=encoder, d_model=64, num_heads=4, ff_dim=64, num_layers=2, dropout_rate=0.0, epochs=20, batch_size=128	471533.28	686.68	510.62	1.7	0.99	2.1	43.15	-0.000104	0.76
107	architecture=encoder, d_model=64, num_heads=4, ff_dim=64, num_layers=2, dropout_rate=0.0, epochs=50, batch_size=64	305757.06	552.95	389.1	1.24	0.99	1.72	162.21	0.000005	1.02
108	architecture=encoder, d_model=64, num_heads=4, ff_dim=64, num_layers=2, dropout_rate=0.0, epochs=50, batch_size=128	353582.18	594.63	423.07	1.32	0.99	1.87	92.24	-0.000002	0.99
109	architecture=encoder, d_model=64, num_heads=4, ff_dim=64, num_layers=2, dropout_rate=0.2, epochs=20, batch_size=64	1927915.61	1388.49	1054.5	2.96	0.96	2.76	74.88	0.000845	2.8
110	architecture=encoder, d_model=64, num_heads=4, ff_dim=64, num_layers=2, dropout_rate=0.2, epochs=20, batch_size=128	4095135.79	2023.64	1829.9	5.45	0.92	2.47	49.14	0.002189	4.62
111	architecture=encoder, d_model=64, num_heads=4, ff_dim=64, num_layers=2, dropout_rate=0.2, epochs=50, batch_size=64	5492807.82	2343.67	1987.72	5.63	0.89	2.95	174.48	0.003483	14.38
112	architecture=encoder, d_model=64, num_heads=4, ff_dim=64, num_layers=2, dropout_rate=0.2, epochs=50, batch_size=128	4391923.89	2095.69	1727.24	4.84	0.91	3.03	104.95	0.002703	10.36
113	architecture=encoder, d_model=64, num_heads=4, ff_dim=128, num_layers=1, dropout_rate=0.0, epochs=20, batch_size=64	708310.05	841.61	692.59	2.17	0.99	1.91	58.8	0.000158	1.49
114	architecture=encoder, d_model=64, num_heads=4, ff_dim=128, num_layers=1, dropout_rate=0.0, epochs=20, batch_size=128	436693.74	660.83	496.87	1.56	0.99	1.9	32.92	-0.000048	0.86

Table 5.6: Results of the ablation study of the Transformers architecture.

ID	Config.	MSE	RMSE	MAE	MAPE	R^2	CVRMSE	Training Time (s)	Overfit Gap	Loss Ratio
115	architecture=encoder, d_model=64, num_heads=4, ff_dim=128, num_layers=1, dropout_rate=0.0, epochs=50, batch_size=64	265547.37	515.31	351.49	1.13	0.99	1.67	131.51	-0.000036	0.83
116	architecture=encoder, d_model=64, num_heads=4, ff_dim=128, num_layers=1, dropout_rate=0.0, epochs=50, batch_size=128	289063.71	537.65	377.11	1.21	0.99	1.73	71.23	-0.000058	0.77
117	architecture=encoder, d_model=64, num_heads=4, ff_dim=128, num_layers=1, dropout_rate=0.2, epochs=20, batch_size=64	1805722.59	1343.77	1023.81	2.91	0.96	3.03	59.98	0.000753	2.58
118	architecture=encoder, d_model=64, num_heads=4, ff_dim=128, num_layers=1, dropout_rate=0.2, epochs=20, batch_size=128	1497040.92	1223.54	937.65	2.7	0.97	2.69	34.79	0.000415	1.69
119	architecture=encoder, d_model=64, num_heads=4, ff_dim=128, num_layers=1, dropout_rate=0.2, epochs=50, batch_size=64	3683003.24	1919.12	1508.86	4.16	0.93	3.25	140.99	0.002248	9.58
120	architecture=encoder, d_model=64, num_heads=4, ff_dim=128, num_layers=1, dropout_rate=0.2, epochs=50, batch_size=128	2451143.52	1565.61	1238.99	3.45	0.95	2.79	76.9	0.001377	5.65
121	architecture=encoder, d_model=64, num_heads=4, ff_dim=128, num_layers=2, dropout_rate=0.0, epochs=20, batch_size=64	458065.26	676.81	507.15	1.75	0.99	2.19	69.82	-0.000000	1
122	architecture=encoder, d_model=64, num_heads=4, ff_dim=128, num_layers=2, dropout_rate=0.0, epochs=20, batch_size=128	596449.62	772.3	584.5	1.79	0.99	1.99	43.85	0.000023	1.06
123	architecture=encoder, d_model=64, num_heads=4, ff_dim=128, num_layers=2, dropout_rate=0.0, epochs=50, batch_size=64	237236.6	487.07	332.33	1.06	1	1.57	155.75	-0.000037	0.82
124	architecture=encoder, d_model=64, num_heads=4, ff_dim=128, num_layers=2, dropout_rate=0.0, epochs=50, batch_size=128	343066.2	585.72	414.58	1.32	0.99	1.9	94.25	0.000004	1.02

Table 5.6: Results of the ablation study of the Transformers architecture.

ID	Config.	MSE	RMSE	MAE	MAPE	R^2	CVRMSE	Training Time (s)	Overfit Gap	Loss Ratio
125	architecture=encoder, d_model=64, num_heads=4, ff_dim=128, num_layers=2, dropout_rate=0.2, epochs=20, batch_size=64	2166288.52	1471.83	1218.25	3.51	0.96	2.37	75.41	0.001052	3.48
126	architecture=encoder, d_model=64, num_heads=4, ff_dim=128, num_layers=2, dropout_rate=0.2, epochs=20, batch_size=128	3212496.87	1792.34	1454.86	4.11	0.94	2.86	49.98	0.001692	4.4
127	architecture=encoder, d_model=64, num_heads=4, ff_dim=128, num_layers=2, dropout_rate=0.2, epochs=50, batch_size=64	4059712.02	2014.87	1736.44	4.99	0.92	2.34	175.58	0.002501	10.44
128	architecture=encoder, d_model=64, num_heads=4, ff_dim=128, num_layers=2, dropout_rate=0.2, epochs=50, batch_size=128	1903228.77	1379.58	1088.72	3.06	0.96	2.72	105.12	0.000990	4.23
129	architecture=encdec, d_model=32, num_heads=2, ff_dim=64, num_layers=1, dropout_rate=0.0, epochs=20, batch_size=64	392867.9	626.79	449.61	1.45	0.99	2.03	60.83	-0.000085	0.76
130	architecture=encdec, d_model=32, num_heads=2, ff_dim=64, num_layers=1, dropout_rate=0.0, epochs=20, batch_size=128	486019.34	697.15	484.34	1.54	0.99	2.25	33.8	-0.000117	0.74
131	architecture=encdec, d_model=32, num_heads=2, ff_dim=64, num_layers=1, dropout_rate=0.0, epochs=50, batch_size=64	297591.19	545.52	382.77	1.23	0.99	1.77	129.3	-0.000029	0.87
132	architecture=encdec, d_model=32, num_heads=2, ff_dim=64, num_layers=1, dropout_rate=0.0, epochs=50, batch_size=128	353400.76	594.48	435.67	1.41	0.99	1.8	70.01	-0.000024	0.91
133	architecture=encdec, d_model=32, num_heads=2, ff_dim=64, num_layers=1, dropout_rate=0.2, epochs=20, batch_size=64	690408.45	830.91	615.07	1.84	0.99	2.2	60.66	0.000105	1.29
134	architecture=encdec, d_model=32, num_heads=2, ff_dim=64, num_layers=1, dropout_rate=0.2, epochs=20, batch_size=128	1264343.06	1124.43	906.74	2.71	0.98	2.45	35.11	0.000396	1.85

Table 5.6: Results of the ablation study of the Transformers architecture.

ID	Config.	MSE	RMSE	MAE	MAPE	R^2	CVRMSE	Training Time (s)	Overfit Gap	Loss Ratio
135	architecture=encdec, d_model=32, num_heads=2, ff_dim=64, num_layers=1, dropout_rate=0.2, epochs=50, batch_size=64	612503.51	782.63	621.44	2.03	0.99	1.96	141.33	0.000169	1.68
136	architecture=encdec, d_model=32, num_heads=2, ff_dim=64, num_layers=1, dropout_rate=0.2, epochs=50, batch_size=128	566333.35	752.55	597.68	1.98	0.99	2.08	74.45	0.000094	1.32
137	architecture=encdec, d_model=32, num_heads=2, ff_dim=64, num_layers=2, dropout_rate=0.0, epochs=20, batch_size=64	361359.6	601.13	438.16	1.42	0.99	1.83	69.79	-0.000067	0.79
138	architecture=encdec, d_model=32, num_heads=2, ff_dim=64, num_layers=2, dropout_rate=0.0, epochs=20, batch_size=128	405105.02	636.48	455.13	1.47	0.99	2.08	42.65	-0.000113	0.71
139	architecture=encdec, d_model=32, num_heads=2, ff_dim=64, num_layers=2, dropout_rate=0.0, epochs=50, batch_size=64	283045.89	532.02	367.01	1.18	0.99	1.69	162.56	-0.000022	0.9
140	architecture=encdec, d_model=32, num_heads=2, ff_dim=64, num_layers=2, dropout_rate=0.0, epochs=50, batch_size=128	345214.97	587.55	427.01	1.36	0.99	1.76	86.51	-0.000010	0.96
141	architecture=encdec, d_model=32, num_heads=2, ff_dim=64, num_layers=2, dropout_rate=0.2, epochs=20, batch_size=64	467147.03	683.48	471.96	1.46	0.99	2.05	74.38	-0.000032	0.91
142	architecture=encdec, d_model=32, num_heads=2, ff_dim=64, num_layers=2, dropout_rate=0.2, epochs=20, batch_size=128	1114508.39	1055.7	813.05	2.41	0.98	2.58	44.11	0.000292	1.63
143	architecture=encdec, d_model=32, num_heads=2, ff_dim=64, num_layers=2, dropout_rate=0.2, epochs=50, batch_size=64	716232.41	846.31	706.55	2.21	0.99	1.71	172.8	0.000246	2.02
144	architecture=encdec, d_model=32, num_heads=2, ff_dim=64, num_layers=2, dropout_rate=0.2, epochs=50, batch_size=128	372053.14	609.96	430.09	1.35	0.99	1.87	97.72	-0.000007	0.97

Table 5.6: Results of the ablation study of the Transformers architecture.

ID	Config.	MSE	RMSE	MAE	MAPE	R^2	CVRMSE	Training Time (s)	Overfit Gap	Loss Ratio
145	architecture=encdec, d_model=32, num_heads=2, ff_dim=128, num_layers=1, dropout_rate=0.0, epochs=20, batch_size=64	465987.42	682.63	488.31	1.55	0.99	2.19	59.07	-0.000036	0.9
146	architecture=encdec, d_model=32, num_heads=2, ff_dim=128, num_layers=1, dropout_rate=0.0, epochs=20, batch_size=128	627958.01	792.44	610.88	1.99	0.99	2.23	34.63	0.000024	1.06
147	architecture=encdec, d_model=32, num_heads=2, ff_dim=128, num_layers=1, dropout_rate=0.0, epochs=50, batch_size=64	279716.71	528.88	370.77	1.18	0.99	1.69	128.07	-0.000030	0.86
148	architecture=encdec, d_model=32, num_heads=2, ff_dim=128, num_layers=1, dropout_rate=0.0, epochs=50, batch_size=128	363894.19	603.24	449.15	1.49	0.99	1.84	69.36	-0.000006	0.98
149	architecture=encdec, d_model=32, num_heads=2, ff_dim=128, num_layers=1, dropout_rate=0.2, epochs=20, batch_size=64	412106.83	641.96	464.77	1.47	0.99	1.94	60.03	-0.000085	0.77
150	architecture=encdec, d_model=32, num_heads=2, ff_dim=128, num_layers=1, dropout_rate=0.2, epochs=20, batch_size=128	795719.8	892.03	705.28	2.25	0.98	2.31	34.39	0.000109	1.25
151	architecture=encdec, d_model=32, num_heads=2, ff_dim=128, num_layers=1, dropout_rate=0.2, epochs=50, batch_size=64	334274.85	578.17	411.25	1.3	0.99	1.74	131.83	-0.000013	0.94
152	architecture=encdec, d_model=32, num_heads=2, ff_dim=128, num_layers=1, dropout_rate=0.2, epochs=50, batch_size=128	445856.13	667.72	511.96	1.64	0.99	1.81	72.09	0.000028	1.1
153	architecture=encdec, d_model=32, num_heads=2, ff_dim=128, num_layers=2, dropout_rate=0.0, epochs=20, batch_size=64	439506.01	662.95	496.24	1.61	0.99	1.96	67.99	0.000017	1.06
154	architecture=encdec, d_model=32, num_heads=2, ff_dim=128, num_layers=2, dropout_rate=0.0, epochs=20, batch_size=128	802043.25	895.57	730.75	2.44	0.98	2.33	40.84	0.000136	1.33

Table 5.6: Results of the ablation study of the Transformers architecture.

ID	Config.	MSE	RMSE	MAE	MAPE	R^2	CVRMSE	Training Time (s)	Overfit Gap	Loss Ratio
155	architecture=encdec, d_model=32, num_heads=2, ff_dim=128, num_layers=2, dropout_rate=0.0, epochs=50, batch_size=64	300999.89	548.63	401.62	1.27	0.99	1.6	161.07	-0.000022	0.9
156	architecture=encdec, d_model=32, num_heads=2, ff_dim=128, num_layers=2, dropout_rate=0.0, epochs=50, batch_size=128	283445.52	532.4	377.94	1.22	0.99	1.74	95.91	-0.000051	0.79
157	architecture=encdec, d_model=32, num_heads=2, ff_dim=128, num_layers=2, dropout_rate=0.2, epochs=20, batch_size=64	481134.66	693.64	483.55	1.48	0.99	1.99	77.86	-0.000013	0.96
158	architecture=encdec, d_model=32, num_heads=2, ff_dim=128, num_layers=2, dropout_rate=0.2, epochs=20, batch_size=128	579708.15	761.39	556.52	1.73	0.99	2.35	44.83	-0.000038	0.91
159	architecture=encdec, d_model=32, num_heads=2, ff_dim=128, num_layers=2, dropout_rate=0.2, epochs=50, batch_size=64	326170.28	571.11	410.15	1.27	0.99	1.69	169.43	-0.000010	0.96
160	architecture=encdec, d_model=32, num_heads=2, ff_dim=128, num_layers=2, dropout_rate=0.2, epochs=50, batch_size=128	472349.18	687.28	486.81	1.46	0.99	1.88	95.3	0.000049	1.18
161	architecture=encdec, d_model=32, num_heads=4, ff_dim=64, num_layers=1, dropout_rate=0.0, epochs=20, batch_size=64	453348.52	673.31	488.23	1.56	0.99	2.07	61.69	-0.000087	0.78
162	architecture=encdec, d_model=32, num_heads=4, ff_dim=64, num_layers=1, dropout_rate=0.0, epochs=20, batch_size=128	406647.36	637.69	449.42	1.45	0.99	2.11	34.26	-0.000094	0.75
163	architecture=encdec, d_model=32, num_heads=4, ff_dim=64, num_layers=1, dropout_rate=0.0, epochs=50, batch_size=64	320377.07	566.02	397.69	1.28	0.99	1.76	130.67	-0.000013	0.94
164	architecture=encdec, d_model=32, num_heads=4, ff_dim=64, num_layers=1, dropout_rate=0.0, epochs=50, batch_size=128	325984.47	570.95	395.58	1.27	0.99	1.85	70.42	-0.000044	0.83

Table 5.6: Results of the ablation study of the Transformers architecture.

ID	Config.	MSE	RMSE	MAE	MAPE	R^2	CVRMSE	Training Time (s)	Overfit Gap	Loss Ratio
165	architecture=encdec, d_model=32, num_heads=4, ff_dim=64, num_layers=1, dropout_rate=0.2, epochs=20, batch_size=64	760850.34	872.27	624.88	1.86	0.99	2.3	61.15	0.000132	1.34
166	architecture=encdec, d_model=32, num_heads=4, ff_dim=64, num_layers=1, dropout_rate=0.2, epochs=20, batch_size=128	653871.06	808.62	627.08	1.99	0.99	2.12	34.56	0.000030	1.07
167	architecture=encdec, d_model=32, num_heads=4, ff_dim=64, num_layers=1, dropout_rate=0.2, epochs=50, batch_size=64	351302.36	592.71	411.62	1.27	0.99	1.77	132.94	-0.000009	0.96
168	architecture=encdec, d_model=32, num_heads=4, ff_dim=64, num_layers=1, dropout_rate=0.2, epochs=50, batch_size=128	790292.56	888.98	736.6	2.34	0.98	1.89	73.27	0.000256	1.91
169	architecture=encdec, d_model=32, num_heads=4, ff_dim=64, num_layers=2, dropout_rate=0.0, epochs=20, batch_size=64	680910.41	825.17	663.42	2.09	0.99	1.97	71.98	0.000163	1.54
170	architecture=encdec, d_model=32, num_heads=4, ff_dim=64, num_layers=2, dropout_rate=0.0, epochs=20, batch_size=128	863263.99	929.12	781.59	2.6	0.98	2.25	47.46	0.000211	1.56
171	architecture=encdec, d_model=32, num_heads=4, ff_dim=64, num_layers=2, dropout_rate=0.0, epochs=50, batch_size=64	275703.21	525.07	364.95	1.15	0.99	1.67	172.03	-0.000018	0.91
172	architecture=encdec, d_model=32, num_heads=4, ff_dim=64, num_layers=2, dropout_rate=0.0, epochs=50, batch_size=128	382130.28	618.17	440.28	1.39	0.99	1.76	90.34	0.000016	1.06
173	architecture=encdec, d_model=32, num_heads=4, ff_dim=64, num_layers=2, dropout_rate=0.2, epochs=20, batch_size=64	727010.31	852.65	662.25	2.12	0.99	2.76	75.71	0.000187	1.61
174	architecture=encdec, d_model=32, num_heads=4, ff_dim=64, num_layers=2, dropout_rate=0.2, epochs=20, batch_size=128	2610860.61	1615.82	1462.24	4.58	0.95	2.38	44.57	0.001306	3.76

Table 5.6: Results of the ablation study of the Transformers architecture.

ID	Config.	MSE	RMSE	MAE	MAPE	R^2	CVRMSE	Training Time (s)	Overfit Gap	Loss Ratio
175	architecture=encdec, d_model=32, num_heads=4, ff_dim=64, num_layers=2, dropout_rate=0.2, epochs=50, batch_size=64	285124.65	533.97	374.78	1.2	0.99	1.71	169.5	-0.000059	0.77
176	architecture=encdec, d_model=32, num_heads=4, ff_dim=64, num_layers=2, dropout_rate=0.2, epochs=50, batch_size=128	773926.29	879.73	747.37	2.38	0.98	1.76	93.54	0.000273	2.08
177	architecture=encdec, d_model=32, num_heads=4, ff_dim=128, num_layers=1, dropout_rate=0.0, epochs=20, batch_size=64	363578.02	602.97	434.69	1.4	0.99	1.94	57.54	-0.000041	0.86
178	architecture=encdec, d_model=32, num_heads=4, ff_dim=128, num_layers=1, dropout_rate=0.0, epochs=20, batch_size=128	867621.74	931.46	776.38	2.53	0.98	2.12	32.49	0.000242	1.7
179	architecture=encdec, d_model=32, num_heads=4, ff_dim=128, num_layers=1, dropout_rate=0.0, epochs=50, batch_size=64	295003.12	543.14	394.77	1.24	0.99	1.63	129.44	-0.000009	0.95
180	architecture=encdec, d_model=32, num_heads=4, ff_dim=128, num_layers=1, dropout_rate=0.0, epochs=50, batch_size=128	346561.68	588.69	423.01	1.35	0.99	1.79	69.92	-0.000015	0.94
181	architecture=encdec, d_model=32, num_heads=4, ff_dim=128, num_layers=1, dropout_rate=0.2, epochs=20, batch_size=64	576241.29	759.11	562.22	1.72	0.99	1.99	58.42	0.000046	1.13
182	architecture=encdec, d_model=32, num_heads=4, ff_dim=128, num_layers=1, dropout_rate=0.2, epochs=20, batch_size=128	554178.98	744.43	550.86	1.72	0.99	2.25	33.32	-0.000029	0.93
183	architecture=encdec, d_model=32, num_heads=4, ff_dim=128, num_layers=1, dropout_rate=0.2, epochs=50, batch_size=64	489780.77	699.84	532.24	1.62	0.99	1.75	130.4	0.000097	1.41
184	architecture=encdec, d_model=32, num_heads=4, ff_dim=128, num_layers=1, dropout_rate=0.2, epochs=50, batch_size=128	628367.04	792.7	643.26	2.05	0.99	1.78	72.7	0.000164	1.62

Table 5.6: Results of the ablation study of the Transformers architecture.

ID	Config.	MSE	RMSE	MAE	MAPE	R^2	CVRMSE	Training Time (s)	Overfit Gap	Loss Ratio
185	architecture=encdec, d_model=32, num_heads=4, ff_dim=128, num_layers=2, dropout_rate=0.0, epochs=20, batch_size=64	376820.99	613.86	456.1	1.51	0.99	1.88	71.22	-0.000026	0.91
186	architecture=encdec, d_model=32, num_heads=4, ff_dim=128, num_layers=2, dropout_rate=0.0, epochs=20, batch_size=128	397334.6	630.34	448.23	1.43	0.99	2.03	61.38	-0.000068	0.8
187	architecture=encdec, d_model=32, num_heads=4, ff_dim=128, num_layers=2, dropout_rate=0.0, epochs=50, batch_size=64	262107.77	511.96	358.17	1.13	0.99	1.63	181.14	-0.000022	0.89
188	architecture=encdec, d_model=32, num_heads=4, ff_dim=128, num_layers=2, dropout_rate=0.0, epochs=50, batch_size=128	527939.82	726.59	582.15	1.79	0.99	1.66	104	0.000116	1.47
189	architecture=encdec, d_model=32, num_heads=4, ff_dim=128, num_layers=2, dropout_rate=0.2, epochs=20, batch_size=64	1850771.6	1360.43	1185	3.69	0.96	2.16	91.16	0.000923	3.73
190	architecture=encdec, d_model=32, num_heads=4, ff_dim=128, num_layers=2, dropout_rate=0.2, epochs=20, batch_size=128	436458.93	660.65	479.12	1.52	0.99	2.02	52.63	-0.000115	0.72
191	architecture=encdec, d_model=32, num_heads=4, ff_dim=128, num_layers=2, dropout_rate=0.2, epochs=50, batch_size=64	306310.1	553.45	390.78	1.21	0.99	1.7	197.21	-0.000026	0.89
192	architecture=encdec, d_model=32, num_heads=4, ff_dim=128, num_layers=2, dropout_rate=0.2, epochs=50, batch_size=128	353024.22	594.16	428.22	1.36	0.99	1.75	108.71	-0.000039	0.86
193	architecture=encdec, d_model=64, num_heads=2, ff_dim=64, num_layers=1, dropout_rate=0.0, epochs=20, batch_size=64	618743.21	786.6	611.69	1.9	0.99	1.9	75.53	0.000118	1.39
194	architecture=encdec, d_model=64, num_heads=2, ff_dim=64, num_layers=1, dropout_rate=0.0, epochs=20, batch_size=128	695108.93	833.73	660.27	2.18	0.99	2.26	42.99	0.000003	1.01

Table 5.6: Results of the ablation study of the Transformers architecture.

ID	Config.	MSE	RMSE	MAE	MAPE	R^2	CVRMSE	Training Time (s)	Overfit Gap	Loss Ratio
195	architecture=encdec, d_model=64, num_heads=2, ff_dim=64, num_layers=1, dropout_rate=0.0, epochs=50, batch_size=64	268710.07	518.37	356.59	1.13	0.99	1.65	153.95	-0.000046	0.8
196	architecture=encdec, d_model=64, num_heads=2, ff_dim=64, num_layers=1, dropout_rate=0.0, epochs=50, batch_size=128	326294.94	571.22	395.9	1.28	0.99	1.89	73.71	-0.000045	0.83
197	architecture=encdec, d_model=64, num_heads=2, ff_dim=64, num_layers=1, dropout_rate=0.2, epochs=20, batch_size=64	512623.24	715.98	531.03	1.63	0.99	2.02	66.54	0.000010	1.03
198	architecture=encdec, d_model=64, num_heads=2, ff_dim=64, num_layers=1, dropout_rate=0.2, epochs=20, batch_size=128	1345593.53	1160	964.54	2.92	0.97	2.26	37.18	0.000465	2.03
199	architecture=encdec, d_model=64, num_heads=2, ff_dim=64, num_layers=1, dropout_rate=0.2, epochs=50, batch_size=64	709754.96	842.47	672.36	2.08	0.99	1.82	139.25	0.000242	2
200	architecture=encdec, d_model=64, num_heads=2, ff_dim=64, num_layers=1, dropout_rate=0.2, epochs=50, batch_size=128	813668.34	902.04	742.75	2.34	0.98	1.96	78.01	0.000243	1.78
201	architecture=encdec, d_model=64, num_heads=2, ff_dim=64, num_layers=2, dropout_rate=0.0, epochs=20, batch_size=64	305143.61	552.4	383.5	1.24	0.99	1.82	74.48	-0.000085	0.71
202	architecture=encdec, d_model=64, num_heads=2, ff_dim=64, num_layers=2, dropout_rate=0.0, epochs=20, batch_size=128	1273157.91	1128.34	983.75	3.15	0.98	2.09	43.15	0.000465	2.15
203	architecture=encdec, d_model=64, num_heads=2, ff_dim=64, num_layers=2, dropout_rate=0.0, epochs=50, batch_size=64	329278.12	573.83	425.03	1.39	0.99	1.68	162.8	0.000012	1.06
204	architecture=encdec, d_model=64, num_heads=2, ff_dim=64, num_layers=2, dropout_rate=0.0, epochs=50, batch_size=128	347821.85	589.76	430.28	1.35	0.99	1.66	90.73	-0.000014	0.95

Table 5.6: Results of the ablation study of the Transformers architecture.

ID	Config.	MSE	RMSE	MAE	MAPE	R^2	CVRMSE	Training Time (s)	Overfit Gap	Loss Ratio
205	architecture=encdec, d_model=64, num_heads=2, ff_dim=64, num_layers=2, dropout_rate=0.2, epochs=20, batch_size=64	1046856.86	1023.16	825.66	2.46	0.98	2.1	80.33	0.000378	2.12
206	architecture=encdec, d_model=64, num_heads=2, ff_dim=64, num_layers=2, dropout_rate=0.2, epochs=20, batch_size=128	1123154.76	1059.79	761.22	2.24	0.98	2.76	47.83	0.000309	1.68
207	architecture=encdec, d_model=64, num_heads=2, ff_dim=64, num_layers=2, dropout_rate=0.2, epochs=50, batch_size=64	678687.19	823.82	623.75	1.85	0.99	1.87	173.38	0.000223	1.93
208	architecture=encdec, d_model=64, num_heads=2, ff_dim=64, num_layers=2, dropout_rate=0.2, epochs=50, batch_size=128	662090.13	813.69	641.16	2.02	0.99	1.85	96.9	0.000174	1.63
209	architecture=encdec, d_model=64, num_heads=2, ff_dim=128, num_layers=1, dropout_rate=0.0, epochs=20, batch_size=64	361820.23	601.51	427.96	1.38	0.99	1.98	67.41	-0.000083	0.75
210	architecture=encdec, d_model=64, num_heads=2, ff_dim=128, num_layers=1, dropout_rate=0.0, epochs=20, batch_size=128	451270.2	671.77	495.92	1.63	0.99	2.13	38.59	-0.000067	0.82
211	architecture=encdec, d_model=64, num_heads=2, ff_dim=128, num_layers=1, dropout_rate=0.0, epochs=50, batch_size=64	516965.32	719	567.3	1.79	0.99	1.77	138.21	0.000127	1.56
212	architecture=encdec, d_model=64, num_heads=2, ff_dim=128, num_layers=1, dropout_rate=0.0, epochs=50, batch_size=128	290640.26	539.11	369.54	1.19	0.99	1.77	76.54	-0.000040	0.83
213	architecture=encdec, d_model=64, num_heads=2, ff_dim=128, num_layers=1, dropout_rate=0.2, epochs=20, batch_size=64	886459.27	941.52	758.33	2.26	0.98	2.14	65.07	0.000270	1.81
214	architecture=encdec, d_model=64, num_heads=2, ff_dim=128, num_layers=1, dropout_rate=0.2, epochs=20, batch_size=128	821741.93	906.5	714.03	2.19	0.98	2.19	37.76	0.000109	1.24

Table 5.6: Results of the ablation study of the Transformers architecture.

ID	Config.	MSE	RMSE	MAE	MAPE	R^2	CVRMSE	Training Time (s)	Overfit Gap	Loss Ratio
215	architecture=encdec, d_model=64, num_heads=2, ff_dim=128, num_layers=1, dropout_rate=0.2, epochs=50, batch_size=64	841949.29	917.58	739.9	2.23	0.98	1.9	138.16	0.000328	2.34
216	architecture=encdec, d_model=64, num_heads=2, ff_dim=128, num_layers=1, dropout_rate=0.2, epochs=50, batch_size=128	897492.52	947.36	789.7	2.45	0.98	1.84	75.88	0.000329	2.18
217	architecture=encdec, d_model=64, num_heads=2, ff_dim=128, num_layers=2, dropout_rate=0.0, epochs=20, batch_size=64	557046.97	746.36	560.17	1.7	0.99	1.92	76.58	0.000014	1.04
218	architecture=encdec, d_model=64, num_heads=2, ff_dim=128, num_layers=2, dropout_rate=0.0, epochs=20, batch_size=128	363639.41	603.03	421.16	1.36	0.99	1.97	46.03	-0.000142	0.63
219	architecture=encdec, d_model=64, num_heads=2, ff_dim=128, num_layers=2, dropout_rate=0.0, epochs=50, batch_size=64	293592.55	541.84	387.58	1.22	0.99	1.61	168.06	0.000002	1.01
220	architecture=encdec, d_model=64, num_heads=2, ff_dim=128, num_layers=2, dropout_rate=0.0, epochs=50, batch_size=128	486155.32	697.25	550.75	1.69	0.99	1.78	90.66	0.000086	1.35
221	architecture=encdec, d_model=64, num_heads=2, ff_dim=128, num_layers=2, dropout_rate=0.2, epochs=20, batch_size=64	1667551.22	1291.34	1095.68	3.27	0.97	2.07	77.13	0.000776	3.15
222	architecture=encdec, d_model=64, num_heads=2, ff_dim=128, num_layers=2, dropout_rate=0.2, epochs=20, batch_size=128	1027618.17	1013.72	852.25	2.71	0.98	2.1	50.18	0.000238	1.51
223	architecture=encdec, d_model=64, num_heads=2, ff_dim=128, num_layers=2, dropout_rate=0.2, epochs=50, batch_size=64	416133.29	645.08	485.23	1.49	0.99	1.74	182.21	0.000048	1.2
224	architecture=encdec, d_model=64, num_heads=2, ff_dim=128, num_layers=2, dropout_rate=0.2, epochs=50, batch_size=128	453565.98	673.47	492.22	1.5	0.99	1.83	99.66	0.000035	1.13

Table 5.6: Results of the ablation study of the Transformers architecture.

ID	Config.	MSE	RMSE	MAE	MAPE	R^2	CVRMSE	Training Time (s)	Overfit Gap	Loss Ratio
225	architecture=encdec, d_model=64, num_heads=4, ff_dim=64, num_layers=1, dropout_rate=0.0, epochs=20, batch_size=64	383474.61	619.25	444.78	1.4	0.99	1.92	66.35	-0.000049	0.84
226	architecture=encdec, d_model=64, num_heads=4, ff_dim=64, num_layers=1, dropout_rate=0.0, epochs=20, batch_size=128	598833	773.84	607.58	1.93	0.99	2.04	38.62	0.000028	1.07
227	architecture=encdec, d_model=64, num_heads=4, ff_dim=64, num_layers=1, dropout_rate=0.0, epochs=50, batch_size=64	267255.23	516.97	353.02	1.13	0.99	1.68	138.25	-0.000034	0.84
228	architecture=encdec, d_model=64, num_heads=4, ff_dim=64, num_layers=1, dropout_rate=0.0, epochs=50, batch_size=128	349280.53	591	425.4	1.4	0.99	1.92	74.4	-0.000028	0.89
229	architecture=encdec, d_model=64, num_heads=4, ff_dim=64, num_layers=1, dropout_rate=0.2, epochs=20, batch_size=64	1076357.55	1037.48	865.36	2.63	0.98	1.98	64.55	0.000382	2.09
230	architecture=encdec, d_model=64, num_heads=4, ff_dim=64, num_layers=1, dropout_rate=0.2, epochs=20, batch_size=128	1349117.45	1161.52	999.78	3.11	0.97	2.14	37.99	0.000499	2.18
231	architecture=encdec, d_model=64, num_heads=4, ff_dim=64, num_layers=1, dropout_rate=0.2, epochs=50, batch_size=64	781237.04	883.88	677.06	2.03	0.98	1.95	140.63	0.000288	2.18
232	architecture=encdec, d_model=64, num_heads=4, ff_dim=64, num_layers=1, dropout_rate=0.2, epochs=50, batch_size=128	730305.57	854.58	715.71	2.35	0.99	1.87	82.51	0.000249	2
233	architecture=encdec, d_model=64, num_heads=4, ff_dim=64, num_layers=2, dropout_rate=0.0, epochs=20, batch_size=64	400385.81	632.76	451.73	1.43	0.99	1.89	76.84	-0.000012	0.96
234	architecture=encdec, d_model=64, num_heads=4, ff_dim=64, num_layers=2, dropout_rate=0.0, epochs=20, batch_size=128	380961.65	617.22	438.02	1.41	0.99	1.93	46.62	-0.000072	0.78

Table 5.6: Results of the ablation study of the Transformers architecture.

ID	Config.	MSE	RMSE	MAE	MAPE	R^2	CVRMSE	Training Time (s)	Overfit Gap	Loss Ratio
235	architecture=encdec, d_model=64, num_heads=4, ff_dim=64, num_layers=2, dropout_rate=0.0, epochs=50, batch_size=64	282194.61	531.22	368.2	1.18	0.99	1.69	168.68	-0.000010	0.95
236	architecture=encdec, d_model=64, num_heads=4, ff_dim=64, num_layers=2, dropout_rate=0.0, epochs=50, batch_size=128	277445.58	526.73	367.75	1.18	0.99	1.72	96.28	-0.000054	0.78
237	architecture=encdec, d_model=64, num_heads=4, ff_dim=64, num_layers=2, dropout_rate=0.2, epochs=20, batch_size=64	1050854.37	1025.11	778.54	2.32	0.98	2.24	78.21	0.000344	1.92
238	architecture=encdec, d_model=64, num_heads=4, ff_dim=64, num_layers=2, dropout_rate=0.2, epochs=20, batch_size=128	2937457.01	1713.9	1487.15	4.43	0.94	2.39	48.05	0.001525	4.19
239	architecture=encdec, d_model=64, num_heads=4, ff_dim=64, num_layers=2, dropout_rate=0.2, epochs=50, batch_size=64	352611.74	593.81	412.31	1.27	0.99	1.8	181.93	-0.000006	0.98
240	architecture=encdec, d_model=64, num_heads=4, ff_dim=64, num_layers=2, dropout_rate=0.2, epochs=50, batch_size=128	852213.63	923.15	635.25	1.83	0.98	2.16	106.64	0.000298	2.05
241	architecture=encdec, d_model=64, num_heads=4, ff_dim=128, num_layers=1, dropout_rate=0.0, epochs=20, batch_size=64	343403.56	586.01	414.94	1.32	0.99	1.83	62.88	-0.000091	0.72
242	architecture=encdec, d_model=64, num_heads=4, ff_dim=128, num_layers=1, dropout_rate=0.0, epochs=20, batch_size=128	371587.66	609.58	440.76	1.44	0.99	1.92	35.68	-0.000088	0.74
243	architecture=encdec, d_model=64, num_heads=4, ff_dim=128, num_layers=1, dropout_rate=0.0, epochs=50, batch_size=64	300674.17	548.34	402.02	1.29	0.99	1.65	141.77	-0.000023	0.9
244	architecture=encdec, d_model=64, num_heads=4, ff_dim=128, num_layers=1, dropout_rate=0.0, epochs=50, batch_size=128	296569.61	544.58	380.86	1.23	0.99	1.76	77.89	-0.000050	0.8

Table 5.6: Results of the ablation study of the Transformers architecture.

ID	Config.	MSE	RMSE	MAE	MAPE	R^2	CVRMSE	Training Time (s)	Overfit Gap	Loss Ratio
245	architecture=encdec, d_model=64, num_heads=4, ff_dim=128, num_layers=1, dropout_rate=0.2, epochs=20, batch_size=64	526203.04	725.4	533.99	1.67	0.99	2.02	64.13	0.000007	1.02
246	architecture=encdec, d_model=64, num_heads=4, ff_dim=128, num_layers=1, dropout_rate=0.2, epochs=20, batch_size=128	4320708.02	2078.63	1924.34	5.92	0.92	2.12	35.95	0.002499	6.64
247	architecture=encdec, d_model=64, num_heads=4, ff_dim=128, num_layers=1, dropout_rate=0.2, epochs=50, batch_size=64	539082.25	734.22	526.31	1.56	0.99	1.95	141.05	0.000134	1.58
248	architecture=encdec, d_model=64, num_heads=4, ff_dim=128, num_layers=1, dropout_rate=0.2, epochs=50, batch_size=128	1114332.02	1055.62	919	2.98	0.98	1.96	77.35	0.000486	2.76
249	architecture=encdec, d_model=64, num_heads=4, ff_dim=128, num_layers=2, dropout_rate=0.0, epochs=20, batch_size=64	340888.08	583.86	422.8	1.38	0.99	1.88	74.29	-0.000060	0.79
250	architecture=encdec, d_model=64, num_heads=4, ff_dim=128, num_layers=2, dropout_rate=0.0, epochs=20, batch_size=128	1288543.56	1135.14	990.04	3.17	0.97	2.16	49.11	0.000519	2.44
251	architecture=encdec, d_model=64, num_heads=4, ff_dim=128, num_layers=2, dropout_rate=0.0, epochs=50, batch_size=64	309046.68	555.92	389.33	1.24	0.99	1.62	170.99	0.000012	1.06
252	architecture=encdec, d_model=64, num_heads=4, ff_dim=128, num_layers=2, dropout_rate=0.0, epochs=50, batch_size=128	636628.53	797.89	656.13	2.13	0.99	1.89	106.41	0.000151	1.53
253	architecture=encdec, d_model=64, num_heads=4, ff_dim=128, num_layers=2, dropout_rate=0.2, epochs=20, batch_size=64	1021467.77	1010.68	809.35	2.42	0.98	2.52	82.26	0.000353	2.03
254	architecture=encdec, d_model=64, num_heads=4, ff_dim=128, num_layers=2, dropout_rate=0.2, epochs=20, batch_size=128	661192.42	813.14	575.04	1.76	0.99	2.41	51.35	0.000014	1.03

Table 5.6: Results of the ablation study of the Transformers architecture.

ID	Config.	MSE	RMSE	MAE	MAPE	R^2	CVRMSE	Training Time (s)	Overfit Gap	Loss Ratio
255	architecture=encdec, d_model=64, num_heads=4, ff_dim=128, num_layers=2, dropout_rate=0.2, epochs=50, batch_size=64	341383.52	584.28	379.8	1.2	0.99	1.77	180.8	0.000010	1.05
256	architecture=encdec, d_model=64, num_heads=4, ff_dim=128, num_layers=2, dropout_rate=0.2, epochs=50, batch_size=128	673621.17	820.74	629.08	1.89	0.99	2.2	103.91	0.000190	1.71