✅ Saved new best params with score 284.9240417480469 to best\_params.json

🔍 DEBUG: Dataövergripande information:

▶️ train: 2021-01-02 00:00:00 → 2023-12-31 23:00:00 (26256 timsteg)

✅ val: 2024-01-01 00:00:00 → 2024-12-31 23:00:00 (8784 timsteg)

⏪ past\_cov\_train: 2021-01-02 00:00:00 → 2023-12-31 23:00:00

⏪ past\_cov\_val: 2023-12-28 00:00:00 → 2024-12-31 23:00:00

⏩ future\_cov\_train: 2021-01-02 00:00:00 → 2023-12-31 23:00:00

GPU available: True (cuda), used: True

TPU available: False, using: 0 TPU cores

HPU available: False, using: 0 HPUs

LOCAL\_RANK: 0 - CUDA\_VISIBLE\_DEVICES: [0]

| Name | Type | Params | Mode

------------------------------------------------------------------------------------------------

0 | train\_metrics | MetricCollection | 0 | train

1 | val\_metrics | MetricCollection | 0 | train

2 | input\_embeddings | \_MultiEmbedding | 0 | train

3 | static\_covariates\_vsn | \_VariableSelectionNetwork | 0 | train

4 | encoder\_vsn | \_VariableSelectionNetwork | 201 K | train

5 | decoder\_vsn | \_VariableSelectionNetwork | 101 K | train

6 | static\_context\_grn | \_GatedResidualNetwork | 6.6 K | train

7 | static\_context\_hidden\_encoder\_grn | \_GatedResidualNetwork | 6.6 K | train

8 | static\_context\_cell\_encoder\_grn | \_GatedResidualNetwork | 6.6 K | train

9 | static\_context\_enrichment | \_GatedResidualNetwork | 6.6 K | train

10 | lstm\_encoder | LSTM | 39.4 K | train

11 | lstm\_decoder | LSTM | 39.4 K | train

12 | post\_lstm\_gan | \_GateAddNorm | 3.4 K | train

13 | static\_enrichment\_grn | \_GatedResidualNetwork | 8.2 K | train

14 | multihead\_attn | \_InterpretableMultiHeadAttention | 4.1 K | train

15 | post\_attn\_gan | \_GateAddNorm | 3.4 K | train

16 | feed\_forward\_block | \_GatedResidualNetwork | 6.6 K | train

17 | pre\_output\_gan | \_GateAddNorm | 3.4 K | train

18 | output\_layer | Linear | 123 | train

------------------------------------------------------------------------------------------------

435 K Trainable params

0 Non-trainable params

435 K Total params

1.743 Total estimated model params size (MB)

3062 Modules in train mode

0 Modules in eval mode

🚀 Startar ny gridsearch-körning med hyperparametrar:

input\_chunk\_length: 72

output\_chunk\_length: 48

hidden\_size: 40

lstm\_layers: 3

num\_attention\_heads: 4

dropout: 0.4

optimizer\_kwargs: {'lr': 0.001}

Epoch 0: 100%|██████████| 205/205 [05:16<00:00, 0.65it/s, train\_loss=0.725, val\_loss=0.757]📉 Epoch 1: train\_loss=0.7250 | val\_loss=0.7568

Epoch 0: 100%|██████████| 205/205 [05:16<00:00, 0.65it/s, train\_loss=0.725, val\_loss=0.757]Metric val\_loss improved. New best score: 0.757

Epoch 1: 100%|██████████| 205/205 [05:16<00:00, 0.65it/s, train\_loss=0.418, val\_loss=0.543]📉 Epoch 2: train\_loss=0.4179 | val\_loss=0.5429

Epoch 1: 100%|██████████| 205/205 [05:16<00:00, 0.65it/s, train\_loss=0.418, val\_loss=0.543]Metric val\_loss improved by 0.214 >= min\_delta = 0.01. New best score: 0.543

Epoch 2: 100%|██████████| 205/205 [05:16<00:00, 0.65it/s, train\_loss=0.541, val\_loss=0.585]📉 Epoch 3: train\_loss=0.5412 | val\_loss=0.5849

Epoch 3: 100%|██████████| 205/205 [05:15<00:00, 0.65it/s, train\_loss=0.346, val\_loss=0.510]📉 Epoch 4: train\_loss=0.3462 | val\_loss=0.5103

Epoch 3: 100%|██████████| 205/205 [05:15<00:00, 0.65it/s, train\_loss=0.346, val\_loss=0.510]Metric val\_loss improved by 0.033 >= min\_delta = 0.01. New best score: 0.510

Epoch 4: 100%|██████████| 205/205 [05:15<00:00, 0.65it/s, train\_loss=0.279, val\_loss=0.401]📉 Epoch 5: train\_loss=0.2794 | val\_loss=0.4011

Epoch 4: 100%|██████████| 205/205 [05:15<00:00, 0.65it/s, train\_loss=0.279, val\_loss=0.401]Metric val\_loss improved by 0.109 >= min\_delta = 0.01. New best score: 0.401

Epoch 5: 100%|██████████| 205/205 [02:34<00:00, 1.33it/s, train\_loss=0.380, val\_loss=0.404]📉 Epoch 6: train\_loss=0.3797 | val\_loss=0.4042

Epoch 6: 100%|██████████| 205/205 [02:24<00:00, 1.42it/s, train\_loss=0.305, val\_loss=0.392]📉 Epoch 7: train\_loss=0.3051 | val\_loss=0.3918

Epoch 7: 100%|██████████| 205/205 [02:24<00:00, 1.42it/s, train\_loss=0.340, val\_loss=0.393]📉 Epoch 8: train\_loss=0.3403 | val\_loss=0.3930

Epoch 7: 100%|██████████| 205/205 [02:24<00:00, 1.42it/s, train\_loss=0.340, val\_loss=0.393]Monitored metric val\_loss did not improve in the last 3 records. Best score: 0.401. Signaling Trainer to stop.

Epoch 7: 100%|██████████| 205/205 [02:24<00:00, 1.42it/s, train\_loss=0.340, val\_loss=0.393]

`predict()` was called with `n > output\_chunk\_length`: using auto-regression to forecast the values after `output\_chunk\_length` points. The model will access `(n - output\_chunk\_length)` future values of your `past\_covariates` (relative to the first predicted time step). To hide this warning, set `show\_warnings=False`.

GPU available: True (cuda), used: True

TPU available: False, using: 0 TPU cores

HPU available: False, using: 0 HPUs

LOCAL\_RANK: 0 - CUDA\_VISIBLE\_DEVICES: [0]

Predicting DataLoader 0: 100%|██████████| 1/1 [00:12<00:00, 0.08it/s][I 2025-05-10 12:41:01,214] Trial 0 finished with value: 284.9240417480469 and parameters: {'input\_chunk\_length': 72, 'output\_chunk\_length': 48, 'hidden\_size': 40, 'lstm\_layers': 3, 'num\_attention\_heads': 4, 'dropout': 0.4, 'n\_epochs': 30, 'lr': 0.001, 'batch\_size': 128}. Best is trial 0 with value: 284.9240417480469.

✅ Saved new best params with score 284.9240417480469 to best\_params.json

🔍 DEBUG: Dataövergripande information:

▶️ train: 2021-01-02 00:00:00 → 2023-12-31 23:00:00 (26256 timsteg)

✅ val: 2024-01-01 00:00:00 → 2024-12-31 23:00:00 (8784 timsteg)

⏪ past\_cov\_train: 2021-01-02 00:00:00 → 2023-12-31 23:00:00

⏪ past\_cov\_val: 2023-12-28 00:00:00 → 2024-12-31 23:00:00

⏩ future\_cov\_train: 2021-01-02 00:00:00 → 2023-12-31 23:00:00

⏩ future\_cov\_val: 2023-12-28 00:00:00 → 2024-12-31 23:00:00

GPU available: True (cuda), used: True

TPU available: False, using: 0 TPU cores

HPU available: False, using: 0 HPUs

LOCAL\_RANK: 0 - CUDA\_VISIBLE\_DEVICES: [0]

| Name | Type | Params | Mode

------------------------------------------------------------------------------------------------

0 | train\_metrics | MetricCollection | 0 | train

1 | val\_metrics | MetricCollection | 0 | train

2 | input\_embeddings | \_MultiEmbedding | 0 | train

3 | static\_covariates\_vsn | \_VariableSelectionNetwork | 0 | train

4 | encoder\_vsn | \_VariableSelectionNetwork | 110 K | train

5 | decoder\_vsn | \_VariableSelectionNetwork | 55.4 K | train

6 | static\_context\_grn | \_GatedResidualNetwork | 1.7 K | train

7 | static\_context\_hidden\_encoder\_grn | \_GatedResidualNetwork | 1.7 K | train

8 | static\_context\_cell\_encoder\_grn | \_GatedResidualNetwork | 1.7 K | train

9 | static\_context\_enrichment | \_GatedResidualNetwork | 1.7 K | train

10 | lstm\_encoder | LSTM | 3.4 K | train

11 | lstm\_decoder | LSTM | 3.4 K | train

12 | post\_lstm\_gan | \_GateAddNorm | 880 | train

13 | static\_enrichment\_grn | \_GatedResidualNetwork | 2.1 K | train

14 | multihead\_attn | \_InterpretableMultiHeadAttention | 1.7 K | train

15 | post\_attn\_gan | \_GateAddNorm | 880 | train

16 | feed\_forward\_block | \_GatedResidualNetwork | 1.7 K | train

17 | pre\_output\_gan | \_GateAddNorm | 880 | train

18 | output\_layer | Linear | 63 | train

------------------------------------------------------------------------------------------------

187 K Trainable params

0 Non-trainable params

187 K Total params

0.748 Total estimated model params size (MB)

3056 Modules in train mode

0 Modules in eval mode

🚀 Startar ny gridsearch-körning med hyperparametrar:

input\_chunk\_length: 72

output\_chunk\_length: 48

hidden\_size: 20

lstm\_layers: 1

num\_attention\_heads: 1

dropout: 0.4

optimizer\_kwargs: {'lr': 0.01}

Epoch 0: 100%|██████████| 205/205 [02:22<00:00, 1.43it/s, train\_loss=0.466, val\_loss=0.490]📉 Epoch 1: train\_loss=0.4661 | val\_loss=0.4903

Epoch 0: 100%|██████████| 205/205 [02:22<00:00, 1.43it/s, train\_loss=0.466, val\_loss=0.490]Metric val\_loss improved. New best score: 0.490

Epoch 1: 100%|██████████| 205/205 [02:22<00:00, 1.44it/s, train\_loss=0.383, val\_loss=0.544]📉 Epoch 2: train\_loss=0.3830 | val\_loss=0.5439

Epoch 2: 100%|██████████| 205/205 [02:22<00:00, 1.44it/s, train\_loss=0.387, val\_loss=0.434]📉 Epoch 3: train\_loss=0.3872 | val\_loss=0.4343

Epoch 2: 100%|██████████| 205/205 [02:22<00:00, 1.44it/s, train\_loss=0.387, val\_loss=0.434]Metric val\_loss improved by 0.056 >= min\_delta = 0.01. New best score: 0.434

Epoch 3: 100%|██████████| 205/205 [02:22<00:00, 1.44it/s, train\_loss=0.319, val\_loss=0.407]📉 Epoch 4: train\_loss=0.3189 | val\_loss=0.4074

Epoch 3: 100%|██████████| 205/205 [02:22<00:00, 1.44it/s, train\_loss=0.319, val\_loss=0.407]Metric val\_loss improved by 0.027 >= min\_delta = 0.01. New best score: 0.407

Epoch 4: 100%|██████████| 205/205 [02:22<00:00, 1.44it/s, train\_loss=0.327, val\_loss=0.472]📉 Epoch 5: train\_loss=0.3270 | val\_loss=0.4722

Epoch 5: 100%|██████████| 205/205 [02:22<00:00, 1.44it/s, train\_loss=0.283, val\_loss=0.458]📉 Epoch 6: train\_loss=0.2826 | val\_loss=0.4583

Epoch 6: 100%|██████████| 205/205 [02:22<00:00, 1.44it/s, train\_loss=0.256, val\_loss=0.456]📉 Epoch 7: train\_loss=0.2555 | val\_loss=0.4564

Epoch 6: 100%|██████████| 205/205 [02:22<00:00, 1.44it/s, train\_loss=0.256, val\_loss=0.456]Monitored metric val\_loss did not improve in the last 3 records. Best score: 0.407. Signaling Trainer to stop.

Epoch 6: 100%|██████████| 205/205 [02:22<00:00, 1.44it/s, train\_loss=0.256, val\_loss=0.456]

`predict()` was called with `n > output\_chunk\_length`: using auto-regression to forecast the values after `output\_chunk\_length` points. The model will access `(n - output\_chunk\_length)` future values of your `past\_covariates` (relative to the first predicted time step). To hide this warning, set `show\_warnings=False`.

GPU available: True (cuda), used: True

TPU available: False, using: 0 TPU cores

HPU available: False, using: 0 HPUs

LOCAL\_RANK: 0 - CUDA\_VISIBLE\_DEVICES: [0]

Predicting DataLoader 0: 100%|██████████| 1/1 [00:11<00:00, 0.08it/s][I 2025-05-10 12:57:51,672] Trial 1 finished with value: 309.8942565917969 and parameters: {'input\_chunk\_length': 72, 'output\_chunk\_length': 48, 'hidden\_size': 20, 'lstm\_layers': 1, 'num\_attention\_heads': 1, 'dropout': 0.4, 'n\_epochs': 30, 'lr': 0.01, 'batch\_size': 128}. Best is trial 0 with value: 284.9240417480469.

🔍 DEBUG: Dataövergripande information:

▶️ train: 2021-01-02 00:00:00 → 2023-12-31 23:00:00 (26256 timsteg)

✅ val: 2024-01-01 00:00:00 → 2024-12-31 23:00:00 (8784 timsteg)

⏪ past\_cov\_train: 2021-01-02 00:00:00 → 2023-12-31 23:00:00

⏪ past\_cov\_val: 2023-12-28 00:00:00 → 2024-12-31 23:00:00

⏩ future\_cov\_train: 2021-01-02 00:00:00 → 2023-12-31 23:00:00

⏩ future\_cov\_val: 2023-12-28 00:00:00 → 2024-12-31 23:00:00

GPU available: True (cuda), used: True

TPU available: False, using: 0 TPU cores

HPU available: False, using: 0 HPUs

LOCAL\_RANK: 0 - CUDA\_VISIBLE\_DEVICES: [0]

| Name | Type | Params | Mode

------------------------------------------------------------------------------------------------

0 | train\_metrics | MetricCollection | 0 | train

1 | val\_metrics | MetricCollection | 0 | train

2 | input\_embeddings | \_MultiEmbedding | 0 | train

3 | static\_covariates\_vsn | \_VariableSelectionNetwork | 0 | train

4 | encoder\_vsn | \_VariableSelectionNetwork | 983 K | train

5 | decoder\_vsn | \_VariableSelectionNetwork | 439 K | train

6 | static\_context\_grn | \_GatedResidualNetwork | 231 K | train

7 | static\_context\_hidden\_encoder\_grn | \_GatedResidualNetwork | 231 K | train

8 | static\_context\_cell\_encoder\_grn | \_GatedResidualNetwork | 231 K | train

9 | static\_context\_enrichment | \_GatedResidualNetwork | 231 K | train

10 | lstm\_encoder | LSTM | 1.4 M | train

11 | lstm\_decoder | LSTM | 1.4 M | train

12 | post\_lstm\_gan | \_GateAddNorm | 116 K | train

13 | static\_enrichment\_grn | \_GatedResidualNetwork | 289 K | train

14 | multihead\_attn | \_InterpretableMultiHeadAttention | 231 K | train

15 | post\_attn\_gan | \_GateAddNorm | 116 K | train

16 | feed\_forward\_block | \_GatedResidualNetwork | 231 K | train

17 | pre\_output\_gan | \_GateAddNorm | 116 K | train

18 | output\_layer | Linear | 723 | train

------------------------------------------------------------------------------------------------

6.2 M Trainable params

0 Non-trainable params

6.2 M Total params

24.906 Total estimated model params size (MB)

3056 Modules in train mode

0 Modules in eval mode

🚀 Startar ny gridsearch-körning med hyperparametrar:

input\_chunk\_length: 96

output\_chunk\_length: 48

hidden\_size: 240

lstm\_layers: 3

num\_attention\_heads: 1

dropout: 0.4

optimizer\_kwargs: {'lr': 0.001}

Epoch 0: 100%|██████████| 205/205 [04:36<00:00, 0.74it/s, train\_loss=0.413, val\_loss=0.538]📉 Epoch 1: train\_loss=0.4125 | val\_loss=0.5378

Epoch 0: 100%|██████████| 205/205 [04:36<00:00, 0.74it/s, train\_loss=0.413, val\_loss=0.538]Metric val\_loss improved. New best score: 0.538

Epoch 1: 100%|██████████| 205/205 [04:37<00:00, 0.74it/s, train\_loss=0.352, val\_loss=0.459]📉 Epoch 2: train\_loss=0.3524 | val\_loss=0.4585

Epoch 1: 100%|██████████| 205/205 [04:37<00:00, 0.74it/s, train\_loss=0.352, val\_loss=0.459]Metric val\_loss improved by 0.079 >= min\_delta = 0.01. New best score: 0.459

Epoch 2: 100%|██████████| 205/205 [04:37<00:00, 0.74it/s, train\_loss=0.827, val\_loss=0.503]📉 Epoch 3: train\_loss=0.8272 | val\_loss=0.5026

Epoch 3: 100%|██████████| 205/205 [04:37<00:00, 0.74it/s, train\_loss=0.468, val\_loss=0.483]📉 Epoch 4: train\_loss=0.4680 | val\_loss=0.4829

Epoch 4: 100%|██████████| 205/205 [04:37<00:00, 0.74it/s, train\_loss=0.731, val\_loss=0.506]📉 Epoch 5: train\_loss=0.7309 | val\_loss=0.5060

Epoch 4: 100%|██████████| 205/205 [04:37<00:00, 0.74it/s, train\_loss=0.731, val\_loss=0.506]Monitored metric val\_loss did not improve in the last 3 records. Best score: 0.459. Signaling Trainer to stop.

Epoch 4: 100%|██████████| 205/205 [04:37<00:00, 0.74it/s, train\_loss=0.731, val\_loss=0.506]

`predict()` was called with `n > output\_chunk\_length`: using auto-regression to forecast the values after `output\_chunk\_length` points. The model will access `(n - output\_chunk\_length)` future values of your `past\_covariates` (relative to the first predicted time step). To hide this warning, set `show\_warnings=False`.

GPU available: True (cuda), used: True

TPU available: False, using: 0 TPU cores

HPU available: False, using: 0 HPUs

LOCAL\_RANK: 0 - CUDA\_VISIBLE\_DEVICES: [0]

Predicting DataLoader 0: 100%|██████████| 1/1 [00:12<00:00, 0.08it/s][I 2025-05-10 13:21:13,320] Trial 2 finished with value: 402.2064208984375 and parameters: {'input\_chunk\_length': 96, 'output\_chunk\_length': 48, 'hidden\_size': 240, 'lstm\_layers': 3, 'num\_attention\_heads': 1, 'dropout': 0.4, 'n\_epochs': 30, 'lr': 0.001, 'batch\_size': 128}. Best is trial 0 with value: 284.9240417480469.

🔍 DEBUG: Dataövergripande information:

▶️ train: 2021-01-02 00:00:00 → 2023-12-31 23:00:00 (26256 timsteg)

✅ val: 2024-01-01 00:00:00 → 2024-12-31 23:00:00 (8784 timsteg)

⏪ past\_cov\_train: 2021-01-02 00:00:00 → 2023-12-31 23:00:00

⏪ past\_cov\_val: 2023-12-28 00:00:00 → 2024-12-31 23:00:00

⏩ future\_cov\_train: 2021-01-02 00:00:00 → 2023-12-31 23:00:00

⏩ future\_cov\_val: 2023-12-28 00:00:00 → 2024-12-31 23:00:00

GPU available: True (cuda), used: True

TPU available: False, using: 0 TPU cores

HPU available: False, using: 0 HPUs

LOCAL\_RANK: 0 - CUDA\_VISIBLE\_DEVICES: [0]

| Name | Type | Params | Mode

------------------------------------------------------------------------------------------------

0 | train\_metrics | MetricCollection | 0 | train

1 | val\_metrics | MetricCollection | 0 | train

2 | input\_embeddings | \_MultiEmbedding | 0 | train

3 | static\_covariates\_vsn | \_VariableSelectionNetwork | 0 | train

4 | encoder\_vsn | \_VariableSelectionNetwork | 983 K | train

5 | decoder\_vsn | \_VariableSelectionNetwork | 439 K | train

6 | static\_context\_grn | \_GatedResidualNetwork | 231 K | train

7 | static\_context\_hidden\_encoder\_grn | \_GatedResidualNetwork | 231 K | train

8 | static\_context\_cell\_encoder\_grn | \_GatedResidualNetwork | 231 K | train

9 | static\_context\_enrichment | \_GatedResidualNetwork | 231 K | train

10 | lstm\_encoder | LSTM | 1.4 M | train

11 | lstm\_decoder | LSTM | 1.4 M | train

12 | post\_lstm\_gan | \_GateAddNorm | 116 K | train

13 | static\_enrichment\_grn | \_GatedResidualNetwork | 289 K | train

14 | multihead\_attn | \_InterpretableMultiHeadAttention | 231 K | train

15 | post\_attn\_gan | \_GateAddNorm | 116 K | train

16 | feed\_forward\_block | \_GatedResidualNetwork | 231 K | train

17 | pre\_output\_gan | \_GateAddNorm | 116 K | train

18 | output\_layer | Linear | 723 | train

------------------------------------------------------------------------------------------------

6.2 M Trainable params

0 Non-trainable params

6.2 M Total params

24.906 Total estimated model params size (MB)

3056 Modules in train mode

0 Modules in eval mode

🚀 Startar ny gridsearch-körning med hyperparametrar:

input\_chunk\_length: 96

output\_chunk\_length: 48

hidden\_size: 240

lstm\_layers: 3

num\_attention\_heads: 1

dropout: 0.5

optimizer\_kwargs: {'lr': 0.0001}

Epoch 0: 100%|██████████| 205/205 [04:37<00:00, 0.74it/s, train\_loss=0.583, val\_loss=0.756]📉 Epoch 1: train\_loss=0.5834 | val\_loss=0.7560

Epoch 0: 100%|██████████| 205/205 [04:37<00:00, 0.74it/s, train\_loss=0.583, val\_loss=0.756]Metric val\_loss improved. New best score: 0.756

Epoch 1: 100%|██████████| 205/205 [04:37<00:00, 0.74it/s, train\_loss=0.277, val\_loss=0.543]📉 Epoch 2: train\_loss=0.2770 | val\_loss=0.5433

Epoch 1: 100%|██████████| 205/205 [04:37<00:00, 0.74it/s, train\_loss=0.277, val\_loss=0.543]Metric val\_loss improved by 0.213 >= min\_delta = 0.01. New best score: 0.543

Epoch 2: 100%|██████████| 205/205 [04:37<00:00, 0.74it/s, train\_loss=0.749, val\_loss=0.650]📉 Epoch 3: train\_loss=0.7489 | val\_loss=0.6497

Epoch 3: 100%|██████████| 205/205 [04:37<00:00, 0.74it/s, train\_loss=0.383, val\_loss=0.563]📉 Epoch 4: train\_loss=0.3829 | val\_loss=0.5628

Epoch 4: 100%|██████████| 205/205 [04:37<00:00, 0.74it/s, train\_loss=0.733, val\_loss=0.424]📉 Epoch 5: train\_loss=0.7327 | val\_loss=0.4243

Epoch 4: 100%|██████████| 205/205 [04:37<00:00, 0.74it/s, train\_loss=0.733, val\_loss=0.424]Metric val\_loss improved by 0.119 >= min\_delta = 0.01. New best score: 0.424

Epoch 5: 100%|██████████| 205/205 [04:37<00:00, 0.74it/s, train\_loss=0.109, val\_loss=0.416]📉 Epoch 6: train\_loss=0.1089 | val\_loss=0.4155

Epoch 6: 100%|██████████| 205/205 [04:37<00:00, 0.74it/s, train\_loss=0.116, val\_loss=0.407]📉 Epoch 7: train\_loss=0.1157 | val\_loss=0.4069

Epoch 6: 100%|██████████| 205/205 [04:37<00:00, 0.74it/s, train\_loss=0.116, val\_loss=0.407]Metric val\_loss improved by 0.017 >= min\_delta = 0.01. New best score: 0.407

Epoch 7: 100%|██████████| 205/205 [04:37<00:00, 0.74it/s, train\_loss=0.321, val\_loss=0.393]📉 Epoch 8: train\_loss=0.3210 | val\_loss=0.3935

Epoch 7: 100%|██████████| 205/205 [04:37<00:00, 0.74it/s, train\_loss=0.321, val\_loss=0.393]Metric val\_loss improved by 0.013 >= min\_delta = 0.01. New best score: 0.393

Epoch 8: 100%|██████████| 205/205 [04:37<00:00, 0.74it/s, train\_loss=0.504, val\_loss=0.425]📉 Epoch 9: train\_loss=0.5041 | val\_loss=0.4248

Epoch 9: 100%|██████████| 205/205 [04:37<00:00, 0.74it/s, train\_loss=0.421, val\_loss=0.407]📉 Epoch 10: train\_loss=0.4210 | val\_loss=0.4069

Epoch 10: 100%|██████████| 205/205 [04:37<00:00, 0.74it/s, train\_loss=0.548, val\_loss=0.666]📉 Epoch 11: train\_loss=0.5484 | val\_loss=0.6660

Epoch 10: 100%|██████████| 205/205 [04:37<00:00, 0.74it/s, train\_loss=0.548, val\_loss=0.666]Monitored metric val\_loss did not improve in the last 3 records. Best score: 0.393. Signaling Trainer to stop.

Epoch 10: 100%|██████████| 205/205 [04:37<00:00, 0.74it/s, train\_loss=0.548, val\_loss=0.666]

`predict()` was called with `n > output\_chunk\_length`: using auto-regression to forecast the values after `output\_chunk\_length` points. The model will access `(n - output\_chunk\_length)` future values of your `past\_covariates` (relative to the first predicted time step). To hide this warning, set `show\_warnings=False`.

GPU available: True (cuda), used: True

TPU available: False, using: 0 TPU cores

HPU available: False, using: 0 HPUs

LOCAL\_RANK: 0 - CUDA\_VISIBLE\_DEVICES: [0]

Predicting DataLoader 0: 100%|██████████| 1/1 [00:12<00:00, 0.08it/s][I 2025-05-10 14:12:22,382] Trial 3 finished with value: 384.7181396484375 and parameters: {'input\_chunk\_length': 96, 'output\_chunk\_length': 48, 'hidden\_size': 240, 'lstm\_layers': 3, 'num\_attention\_heads': 1, 'dropout': 0.5, 'n\_epochs': 30, 'lr': 0.0001, 'batch\_size': 128}. Best is trial 0 with value: 284.9240417480469.

🔍 DEBUG: Dataövergripande information:

▶️ train: 2021-01-02 00:00:00 → 2023-12-31 23:00:00 (26256 timsteg)

✅ val: 2024-01-01 00:00:00 → 2024-12-31 23:00:00 (8784 timsteg)

⏪ past\_cov\_train: 2021-01-02 00:00:00 → 2023-12-31 23:00:00

⏪ past\_cov\_val: 2023-12-28 00:00:00 → 2024-12-31 23:00:00

⏩ future\_cov\_train: 2021-01-02 00:00:00 → 2023-12-31 23:00:00

⏩ future\_cov\_val: 2023-12-28 00:00:00 → 2024-12-31 23:00:00

GPU available: True (cuda), used: True

TPU available: False, using: 0 TPU cores

HPU available: False, using: 0 HPUs

LOCAL\_RANK: 0 - CUDA\_VISIBLE\_DEVICES: [0]

| Name | Type | Params | Mode

------------------------------------------------------------------------------------------------

0 | train\_metrics | MetricCollection | 0 | train

1 | val\_metrics | MetricCollection | 0 | train

2 | input\_embeddings | \_MultiEmbedding | 0 | train

3 | static\_covariates\_vsn | \_VariableSelectionNetwork | 0 | train

4 | encoder\_vsn | \_VariableSelectionNetwork | 66.3 K | train

5 | decoder\_vsn | \_VariableSelectionNetwork | 33.0 K | train

6 | static\_context\_grn | \_GatedResidualNetwork | 460 | train

7 | static\_context\_hidden\_encoder\_grn | \_GatedResidualNetwork | 460 | train

8 | static\_context\_cell\_encoder\_grn | \_GatedResidualNetwork | 460 | train

9 | static\_context\_enrichment | \_GatedResidualNetwork | 460 | train

10 | lstm\_encoder | LSTM | 880 | train

11 | lstm\_decoder | LSTM | 880 | train

12 | post\_lstm\_gan | \_GateAddNorm | 240 | train

13 | static\_enrichment\_grn | \_GatedResidualNetwork | 560 | train

14 | multihead\_attn | \_InterpretableMultiHeadAttention | 430 | train

15 | post\_attn\_gan | \_GateAddNorm | 240 | train

16 | feed\_forward\_block | \_GatedResidualNetwork | 460 | train

17 | pre\_output\_gan | \_GateAddNorm | 240 | train

18 | output\_layer | Linear | 33 | train

------------------------------------------------------------------------------------------------

104 K Trainable params

0 Non-trainable params

104 K Total params

0.416 Total estimated model params size (MB)

3056 Modules in train mode

0 Modules in eval mode

🚀 Startar ny gridsearch-körning med hyperparametrar:

input\_chunk\_length: 72

output\_chunk\_length: 48

hidden\_size: 10

lstm\_layers: 1

num\_attention\_heads: 1

dropout: 0.2

optimizer\_kwargs: {'lr': 0.0001}

Epoch 0: 100%|██████████| 205/205 [02:25<00:00, 1.40it/s, train\_loss=7.720, val\_loss=7.550]📉 Epoch 1: train\_loss=7.7166 | val\_loss=7.5523

Epoch 0: 100%|██████████| 205/205 [02:25<00:00, 1.40it/s, train\_loss=7.720, val\_loss=7.550]Metric val\_loss improved. New best score: 7.552

Epoch 1: 100%|██████████| 205/205 [02:25<00:00, 1.41it/s, train\_loss=7.170, val\_loss=7.150]📉 Epoch 2: train\_loss=7.1687 | val\_loss=7.1471

Epoch 1: 100%|██████████| 205/205 [02:25<00:00, 1.41it/s, train\_loss=7.170, val\_loss=7.150]Metric val\_loss improved by 0.405 >= min\_delta = 0.01. New best score: 7.147

Epoch 2: 100%|██████████| 205/205 [02:25<00:00, 1.41it/s, train\_loss=5.870, val\_loss=6.700]📉 Epoch 3: train\_loss=5.8661 | val\_loss=6.6963

Epoch 2: 100%|██████████| 205/205 [02:25<00:00, 1.41it/s, train\_loss=5.870, val\_loss=6.700]Metric val\_loss improved by 0.451 >= min\_delta = 0.01. New best score: 6.696

Epoch 3: 100%|██████████| 205/205 [02:26<00:00, 1.40it/s, train\_loss=5.630, val\_loss=6.180]📉 Epoch 4: train\_loss=5.6333 | val\_loss=6.1780

Epoch 3: 100%|██████████| 205/205 [02:26<00:00, 1.40it/s, train\_loss=5.630, val\_loss=6.180]Metric val\_loss improved by 0.518 >= min\_delta = 0.01. New best score: 6.178

Epoch 4: 100%|██████████| 205/205 [02:25<00:00, 1.41it/s, train\_loss=5.040, val\_loss=5.600]📉 Epoch 5: train\_loss=5.0433 | val\_loss=5.6001

Epoch 4: 100%|██████████| 205/205 [02:25<00:00, 1.41it/s, train\_loss=5.040, val\_loss=5.600]Metric val\_loss improved by 0.578 >= min\_delta = 0.01. New best score: 5.600

Epoch 5: 100%|██████████| 205/205 [02:25<00:00, 1.41it/s, train\_loss=4.140, val\_loss=4.980]📉 Epoch 6: train\_loss=4.1395 | val\_loss=4.9820

Epoch 5: 100%|██████████| 205/205 [02:25<00:00, 1.41it/s, train\_loss=4.140, val\_loss=4.980]Metric val\_loss improved by 0.618 >= min\_delta = 0.01. New best score: 4.982

Epoch 6: 100%|██████████| 205/205 [02:25<00:00, 1.41it/s, train\_loss=3.590, val\_loss=4.340]📉 Epoch 7: train\_loss=3.5875 | val\_loss=4.3434

Epoch 6: 100%|██████████| 205/205 [02:25<00:00, 1.41it/s, train\_loss=3.590, val\_loss=4.340]Metric val\_loss improved by 0.639 >= min\_delta = 0.01. New best score: 4.343

Epoch 7: 100%|██████████| 205/205 [02:25<00:00, 1.40it/s, train\_loss=3.240, val\_loss=3.710]📉 Epoch 8: train\_loss=3.2445 | val\_loss=3.7052

Epoch 7: 100%|██████████| 205/205 [02:25<00:00, 1.40it/s, train\_loss=3.240, val\_loss=3.710]Metric val\_loss improved by 0.638 >= min\_delta = 0.01. New best score: 3.705

Epoch 8: 100%|██████████| 205/205 [02:26<00:00, 1.40it/s, train\_loss=2.260, val\_loss=3.090]📉 Epoch 9: train\_loss=2.2588 | val\_loss=3.0899

Epoch 8: 100%|██████████| 205/205 [02:26<00:00, 1.40it/s, train\_loss=2.260, val\_loss=3.090]Metric val\_loss improved by 0.615 >= min\_delta = 0.01. New best score: 3.090

Epoch 9: 100%|██████████| 205/205 [02:25<00:00, 1.40it/s, train\_loss=2.010, val\_loss=2.520]📉 Epoch 10: train\_loss=2.0066 | val\_loss=2.5229

Epoch 9: 100%|██████████| 205/205 [02:25<00:00, 1.40it/s, train\_loss=2.010, val\_loss=2.520]Metric val\_loss improved by 0.567 >= min\_delta = 0.01. New best score: 2.523

Epoch 10: 100%|██████████| 205/205 [02:25<00:00, 1.40it/s, train\_loss=1.580, val\_loss=2.030]📉 Epoch 11: train\_loss=1.5773 | val\_loss=2.0286

Epoch 10: 100%|██████████| 205/205 [02:25<00:00, 1.40it/s, train\_loss=1.580, val\_loss=2.030]Metric val\_loss improved by 0.494 >= min\_delta = 0.01. New best score: 2.029

Epoch 11: 100%|██████████| 205/205 [02:25<00:00, 1.41it/s, train\_loss=1.130, val\_loss=1.630]📉 Epoch 12: train\_loss=1.1283 | val\_loss=1.6279

Epoch 11: 100%|██████████| 205/205 [02:25<00:00, 1.41it/s, train\_loss=1.130, val\_loss=1.630]Metric val\_loss improved by 0.401 >= min\_delta = 0.01. New best score: 1.628

Epoch 12: 100%|██████████| 205/205 [02:25<00:00, 1.40it/s, train\_loss=1.080, val\_loss=1.330]📉 Epoch 13: train\_loss=1.0758 | val\_loss=1.3268

Epoch 12: 100%|██████████| 205/205 [02:25<00:00, 1.40it/s, train\_loss=1.080, val\_loss=1.330]Metric val\_loss improved by 0.301 >= min\_delta = 0.01. New best score: 1.327

Epoch 13: 100%|██████████| 205/205 [02:25<00:00, 1.40it/s, train\_loss=0.779, val\_loss=1.120]📉 Epoch 14: train\_loss=0.7793 | val\_loss=1.1157

Epoch 13: 100%|██████████| 205/205 [02:25<00:00, 1.40it/s, train\_loss=0.779, val\_loss=1.120]Metric val\_loss improved by 0.211 >= min\_delta = 0.01. New best score: 1.116

Epoch 14: 100%|██████████| 205/205 [02:25<00:00, 1.40it/s, train\_loss=0.809, val\_loss=0.980]📉 Epoch 15: train\_loss=0.8085 | val\_loss=0.9804

Epoch 14: 100%|██████████| 205/205 [02:25<00:00, 1.40it/s, train\_loss=0.809, val\_loss=0.980]Metric val\_loss improved by 0.135 >= min\_delta = 0.01. New best score: 0.980

Epoch 15: 100%|██████████| 205/205 [02:25<00:00, 1.40it/s, train\_loss=0.776, val\_loss=0.901]📉 Epoch 16: train\_loss=0.7758 | val\_loss=0.9012

Epoch 15: 100%|██████████| 205/205 [02:25<00:00, 1.40it/s, train\_loss=0.776, val\_loss=0.901]Metric val\_loss improved by 0.079 >= min\_delta = 0.01. New best score: 0.901

Epoch 16: 100%|██████████| 205/205 [02:25<00:00, 1.40it/s, train\_loss=0.673, val\_loss=0.805]📉 Epoch 17: train\_loss=0.6727 | val\_loss=0.8053

Epoch 16: 100%|██████████| 205/205 [02:25<00:00, 1.40it/s, train\_loss=0.673, val\_loss=0.805]Metric val\_loss improved by 0.096 >= min\_delta = 0.01. New best score: 0.805

Epoch 17: 100%|██████████| 205/205 [02:26<00:00, 1.40it/s, train\_loss=0.591, val\_loss=0.689]📉 Epoch 18: train\_loss=0.5909 | val\_loss=0.6890

Epoch 17: 100%|██████████| 205/205 [02:26<00:00, 1.40it/s, train\_loss=0.591, val\_loss=0.689]Metric val\_loss improved by 0.116 >= min\_delta = 0.01. New best score: 0.689

Epoch 18: 100%|██████████| 205/205 [02:25<00:00, 1.41it/s, train\_loss=0.508, val\_loss=0.632]📉 Epoch 19: train\_loss=0.5084 | val\_loss=0.6316

Epoch 18: 100%|██████████| 205/205 [02:25<00:00, 1.41it/s, train\_loss=0.508, val\_loss=0.632]Metric val\_loss improved by 0.057 >= min\_delta = 0.01. New best score: 0.632

Epoch 19: 100%|██████████| 205/205 [02:25<00:00, 1.41it/s, train\_loss=0.535, val\_loss=0.634]📉 Epoch 20: train\_loss=0.5348 | val\_loss=0.6338

Epoch 20: 100%|██████████| 205/205 [02:25<00:00, 1.41it/s, train\_loss=0.398, val\_loss=0.616]📉 Epoch 21: train\_loss=0.3984 | val\_loss=0.6158

Epoch 20: 100%|██████████| 205/205 [02:25<00:00, 1.41it/s, train\_loss=0.398, val\_loss=0.616]Metric val\_loss improved by 0.016 >= min\_delta = 0.01. New best score: 0.616

Epoch 21: 100%|██████████| 205/205 [02:26<00:00, 1.40it/s, train\_loss=0.508, val\_loss=0.558]📉 Epoch 22: train\_loss=0.5078 | val\_loss=0.5579

Epoch 21: 100%|██████████| 205/205 [02:26<00:00, 1.40it/s, train\_loss=0.508, val\_loss=0.558]Metric val\_loss improved by 0.058 >= min\_delta = 0.01. New best score: 0.558

Epoch 22: 100%|██████████| 205/205 [02:26<00:00, 1.40it/s, train\_loss=0.579, val\_loss=0.547]📉 Epoch 23: train\_loss=0.5785 | val\_loss=0.5473

Epoch 22: 100%|██████████| 205/205 [02:26<00:00, 1.40it/s, train\_loss=0.579, val\_loss=0.547]Metric val\_loss improved by 0.011 >= min\_delta = 0.01. New best score: 0.547

Epoch 23: 100%|██████████| 205/205 [02:25<00:00, 1.41it/s, train\_loss=0.471, val\_loss=0.519]📉 Epoch 24: train\_loss=0.4706 | val\_loss=0.5190

Epoch 23: 100%|██████████| 205/205 [02:25<00:00, 1.41it/s, train\_loss=0.471, val\_loss=0.519]Metric val\_loss improved by 0.028 >= min\_delta = 0.01. New best score: 0.519

Epoch 24: 100%|██████████| 205/205 [02:25<00:00, 1.41it/s, train\_loss=0.455, val\_loss=0.513]📉 Epoch 25: train\_loss=0.4552 | val\_loss=0.5125

Epoch 25: 100%|██████████| 205/205 [02:25<00:00, 1.41it/s, train\_loss=0.382, val\_loss=0.546]📉 Epoch 26: train\_loss=0.3823 | val\_loss=0.5463

Epoch 26: 100%|██████████| 205/205 [02:26<00:00, 1.40it/s, train\_loss=0.417, val\_loss=0.522]📉 Epoch 27: train\_loss=0.4165 | val\_loss=0.5220

Epoch 26: 100%|██████████| 205/205 [02:26<00:00, 1.40it/s, train\_loss=0.417, val\_loss=0.522]Monitored metric val\_loss did not improve in the last 3 records. Best score: 0.519. Signaling Trainer to stop.

Epoch 26: 100%|██████████| 205/205 [02:26<00:00, 1.40it/s, train\_loss=0.417, val\_loss=0.522]

`predict()` was called with `n > output\_chunk\_length`: using auto-regression to forecast the values after `output\_chunk\_length` points. The model will access `(n - output\_chunk\_length)` future values of your `past\_covariates` (relative to the first predicted time step). To hide this warning, set `show\_warnings=False`.

GPU available: True (cuda), used: True

TPU available: False, using: 0 TPU cores

HPU available: False, using: 0 HPUs

LOCAL\_RANK: 0 - CUDA\_VISIBLE\_DEVICES: [0]

Predicting DataLoader 0: 100%|██████████| 1/1 [00:12<00:00, 0.08it/s][I 2025-05-10 15:18:17,092] Trial 4 finished with value: 399.2142639160156 and parameters: {'input\_chunk\_length': 72, 'output\_chunk\_length': 48, 'hidden\_size': 10, 'lstm\_layers': 1, 'num\_attention\_heads': 1, 'dropout': 0.2, 'n\_epochs': 30, 'lr': 0.0001, 'batch\_size': 128}. Best is trial 0 with value: 284.9240417480469.

🔍 DEBUG: Dataövergripande information:

▶️ train: 2021-01-02 00:00:00 → 2023-12-31 23:00:00 (26256 timsteg)

✅ val: 2024-01-01 00:00:00 → 2024-12-31 23:00:00 (8784 timsteg)

⏪ past\_cov\_train: 2021-01-02 00:00:00 → 2023-12-31 23:00:00

⏪ past\_cov\_val: 2023-12-28 00:00:00 → 2024-12-31 23:00:00

⏩ future\_cov\_train: 2021-01-02 00:00:00 → 2023-12-31 23:00:00

⏩ future\_cov\_val: 2023-12-28 00:00:00 → 2024-12-31 23:00:00

GPU available: True (cuda), used: True

TPU available: False, using: 0 TPU cores

HPU available: False, using: 0 HPUs

LOCAL\_RANK: 0 - CUDA\_VISIBLE\_DEVICES: [0]

| Name | Type | Params | Mode

------------------------------------------------------------------------------------------------

0 | train\_metrics | MetricCollection | 0 | train

1 | val\_metrics | MetricCollection | 0 | train

2 | input\_embeddings | \_MultiEmbedding | 0 | train

3 | static\_covariates\_vsn | \_VariableSelectionNetwork | 0 | train

4 | encoder\_vsn | \_VariableSelectionNetwork | 110 K | train

5 | decoder\_vsn | \_VariableSelectionNetwork | 55.4 K | train

6 | static\_context\_grn | \_GatedResidualNetwork | 1.7 K | train

7 | static\_context\_hidden\_encoder\_grn | \_GatedResidualNetwork | 1.7 K | train

8 | static\_context\_cell\_encoder\_grn | \_GatedResidualNetwork | 1.7 K | train

9 | static\_context\_enrichment | \_GatedResidualNetwork | 1.7 K | train

10 | lstm\_encoder | LSTM | 10.1 K | train

11 | lstm\_decoder | LSTM | 10.1 K | train

12 | post\_lstm\_gan | \_GateAddNorm | 880 | train

13 | static\_enrichment\_grn | \_GatedResidualNetwork | 2.1 K | train

14 | multihead\_attn | \_InterpretableMultiHeadAttention | 1.7 K | train

15 | post\_attn\_gan | \_GateAddNorm | 880 | train

16 | feed\_forward\_block | \_GatedResidualNetwork | 1.7 K | train

17 | pre\_output\_gan | \_GateAddNorm | 880 | train

18 | output\_layer | Linear | 63 | train

------------------------------------------------------------------------------------------------

200 K Trainable params

0 Non-trainable params

200 K Total params

0.802 Total estimated model params size (MB)

3056 Modules in train mode

0 Modules in eval mode

🚀 Startar ny gridsearch-körning med hyperparametrar:

input\_chunk\_length: 96

output\_chunk\_length: 48

hidden\_size: 20

lstm\_layers: 3

num\_attention\_heads: 1

dropout: 0.1

optimizer\_kwargs: {'lr': 0.0001}

Epoch 0: 100%|██████████| 205/205 [02:54<00:00, 1.17it/s, train\_loss=5.940, val\_loss=6.310]📉 Epoch 1: train\_loss=5.9389 | val\_loss=6.3146

Epoch 0: 100%|██████████| 205/205 [02:54<00:00, 1.17it/s, train\_loss=5.940, val\_loss=6.310]Metric val\_loss improved. New best score: 6.315

Epoch 1: 100%|██████████| 205/205 [02:54<00:00, 1.17it/s, train\_loss=5.020, val\_loss=5.530]📉 Epoch 2: train\_loss=5.0165 | val\_loss=5.5318

Epoch 1: 100%|██████████| 205/205 [02:54<00:00, 1.17it/s, train\_loss=5.020, val\_loss=5.530]Metric val\_loss improved by 0.783 >= min\_delta = 0.01. New best score: 5.532

Epoch 2: 100%|██████████| 205/205 [02:55<00:00, 1.17it/s, train\_loss=2.020, val\_loss=4.660]📉 Epoch 3: train\_loss=2.0233 | val\_loss=4.6611

Epoch 2: 100%|██████████| 205/205 [02:55<00:00, 1.17it/s, train\_loss=2.020, val\_loss=4.660]Metric val\_loss improved by 0.871 >= min\_delta = 0.01. New best score: 4.661

Epoch 3: 100%|██████████| 205/205 [02:54<00:00, 1.17it/s, train\_loss=4.180, val\_loss=3.740]📉 Epoch 4: train\_loss=4.1796 | val\_loss=3.7381

Epoch 3: 100%|██████████| 205/205 [02:54<00:00, 1.17it/s, train\_loss=4.180, val\_loss=3.740]Metric val\_loss improved by 0.923 >= min\_delta = 0.01. New best score: 3.738

Epoch 4: 100%|██████████| 205/205 [02:54<00:00, 1.17it/s, train\_loss=1.560, val\_loss=2.840]📉 Epoch 5: train\_loss=1.5608 | val\_loss=2.8417

Epoch 4: 100%|██████████| 205/205 [02:54<00:00, 1.17it/s, train\_loss=1.560, val\_loss=2.840]Metric val\_loss improved by 0.896 >= min\_delta = 0.01. New best score: 2.842

Epoch 5: 100%|██████████| 205/205 [02:54<00:00, 1.17it/s, train\_loss=1.880, val\_loss=2.070]📉 Epoch 6: train\_loss=1.8798 | val\_loss=2.0666

Epoch 5: 100%|██████████| 205/205 [02:54<00:00, 1.17it/s, train\_loss=1.880, val\_loss=2.070]Metric val\_loss improved by 0.775 >= min\_delta = 0.01. New best score: 2.067

Epoch 6: 100%|██████████| 205/205 [02:54<00:00, 1.17it/s, train\_loss=1.790, val\_loss=1.490]📉 Epoch 7: train\_loss=1.7930 | val\_loss=1.4932

Epoch 6: 100%|██████████| 205/205 [02:54<00:00, 1.17it/s, train\_loss=1.790, val\_loss=1.490]Metric val\_loss improved by 0.573 >= min\_delta = 0.01. New best score: 1.493

Epoch 7: 100%|██████████| 205/205 [02:55<00:00, 1.17it/s, train\_loss=0.728, val\_loss=1.140]📉 Epoch 8: train\_loss=0.7278 | val\_loss=1.1361

Epoch 7: 100%|██████████| 205/205 [02:55<00:00, 1.17it/s, train\_loss=0.728, val\_loss=1.140]Metric val\_loss improved by 0.357 >= min\_delta = 0.01. New best score: 1.136

Epoch 8: 100%|██████████| 205/205 [02:54<00:00, 1.17it/s, train\_loss=0.455, val\_loss=0.949]📉 Epoch 9: train\_loss=0.4553 | val\_loss=0.9493

Epoch 8: 100%|██████████| 205/205 [02:54<00:00, 1.17it/s, train\_loss=0.455, val\_loss=0.949]Metric val\_loss improved by 0.187 >= min\_delta = 0.01. New best score: 0.949

Epoch 9: 100%|██████████| 205/205 [02:54<00:00, 1.17it/s, train\_loss=0.721, val\_loss=0.856]📉 Epoch 10: train\_loss=0.7208 | val\_loss=0.8556

Epoch 9: 100%|██████████| 205/205 [02:54<00:00, 1.17it/s, train\_loss=0.721, val\_loss=0.856]Metric val\_loss improved by 0.094 >= min\_delta = 0.01. New best score: 0.856

Epoch 10: 100%|██████████| 205/205 [02:54<00:00, 1.17it/s, train\_loss=0.458, val\_loss=0.820]📉 Epoch 11: train\_loss=0.4577 | val\_loss=0.8201

Epoch 10: 100%|██████████| 205/205 [02:54<00:00, 1.17it/s, train\_loss=0.458, val\_loss=0.820]Metric val\_loss improved by 0.036 >= min\_delta = 0.01. New best score: 0.820

Epoch 11: 100%|██████████| 205/205 [02:54<00:00, 1.17it/s, train\_loss=0.344, val\_loss=0.593]📉 Epoch 12: train\_loss=0.3438 | val\_loss=0.5934

Epoch 11: 100%|██████████| 205/205 [02:54<00:00, 1.17it/s, train\_loss=0.344, val\_loss=0.593]Metric val\_loss improved by 0.227 >= min\_delta = 0.01. New best score: 0.593

Epoch 12: 100%|██████████| 205/205 [02:54<00:00, 1.17it/s, train\_loss=0.303, val\_loss=0.658]📉 Epoch 13: train\_loss=0.3026 | val\_loss=0.6583

Epoch 13: 100%|██████████| 205/205 [02:54<00:00, 1.17it/s, train\_loss=0.697, val\_loss=0.495]📉 Epoch 14: train\_loss=0.6966 | val\_loss=0.4951

Epoch 13: 100%|██████████| 205/205 [02:54<00:00, 1.17it/s, train\_loss=0.697, val\_loss=0.495]Metric val\_loss improved by 0.098 >= min\_delta = 0.01. New best score: 0.495

Epoch 14: 100%|██████████| 205/205 [02:54<00:00, 1.17it/s, train\_loss=0.361, val\_loss=0.507]📉 Epoch 15: train\_loss=0.3607 | val\_loss=0.5074

Epoch 15: 100%|██████████| 205/205 [02:54<00:00, 1.17it/s, train\_loss=0.280, val\_loss=0.521]📉 Epoch 16: train\_loss=0.2798 | val\_loss=0.5209

Epoch 16: 100%|██████████| 205/205 [02:54<00:00, 1.17it/s, train\_loss=0.269, val\_loss=0.558]📉 Epoch 17: train\_loss=0.2689 | val\_loss=0.5581

Epoch 16: 100%|██████████| 205/205 [02:54<00:00, 1.17it/s, train\_loss=0.269, val\_loss=0.558]Monitored metric val\_loss did not improve in the last 3 records. Best score: 0.495. Signaling Trainer to stop.

Epoch 16: 100%|██████████| 205/205 [02:54<00:00, 1.17it/s, train\_loss=0.269, val\_loss=0.558]

`predict()` was called with `n > output\_chunk\_length`: using auto-regression to forecast the values after `output\_chunk\_length` points. The model will access `(n - output\_chunk\_length)` future values of your `past\_covariates` (relative to the first predicted time step). To hide this warning, set `show\_warnings=False`.

GPU available: True (cuda), used: True

TPU available: False, using: 0 TPU cores

HPU available: False, using: 0 HPUs

LOCAL\_RANK: 0 - CUDA\_VISIBLE\_DEVICES: [0]

Predicting DataLoader 0: 100%|██████████| 1/1 [00:11<00:00, 0.08it/s][I 2025-05-10 16:08:01,063] Trial 5 finished with value: 407.4541931152344 and parameters: {'input\_chunk\_length': 96, 'output\_chunk\_length': 48, 'hidden\_size': 20, 'lstm\_layers': 3, 'num\_attention\_heads': 1, 'dropout': 0.1, 'n\_epochs': 30, 'lr': 0.0001, 'batch\_size': 128}. Best is trial 0 with value: 284.9240417480469.

🔍 DEBUG: Dataövergripande information:

▶️ train: 2021-01-02 00:00:00 → 2023-12-31 23:00:00 (26256 timsteg)

✅ val: 2024-01-01 00:00:00 → 2024-12-31 23:00:00 (8784 timsteg)

⏪ past\_cov\_train: 2021-01-02 00:00:00 → 2023-12-31 23:00:00

⏪ past\_cov\_val: 2023-12-28 00:00:00 → 2024-12-31 23:00:00

⏩ future\_cov\_train: 2021-01-02 00:00:00 → 2023-12-31 23:00:00

⏩ future\_cov\_val: 2023-12-28 00:00:00 → 2024-12-31 23:00:00

GPU available: True (cuda), used: True

TPU available: False, using: 0 TPU cores

HPU available: False, using: 0 HPUs

LOCAL\_RANK: 0 - CUDA\_VISIBLE\_DEVICES: [0]

| Name | Type | Params | Mode

------------------------------------------------------------------------------------------------

0 | train\_metrics | MetricCollection | 0 | train

1 | val\_metrics | MetricCollection | 0 | train

2 | input\_embeddings | \_MultiEmbedding | 0 | train

3 | static\_covariates\_vsn | \_VariableSelectionNetwork | 0 | train

4 | encoder\_vsn | \_VariableSelectionNetwork | 1.2 M | train

5 | decoder\_vsn | \_VariableSelectionNetwork | 565 K | train

6 | static\_context\_grn | \_GatedResidualNetwork | 411 K | train

7 | static\_context\_hidden\_encoder\_grn | \_GatedResidualNetwork | 411 K | train

8 | static\_context\_cell\_encoder\_grn | \_GatedResidualNetwork | 411 K | train

9 | static\_context\_enrichment | \_GatedResidualNetwork | 411 K | train

10 | lstm\_encoder | LSTM | 2.5 M | train

11 | lstm\_decoder | LSTM | 2.5 M | train

12 | post\_lstm\_gan | \_GateAddNorm | 206 K | train

13 | static\_enrichment\_grn | \_GatedResidualNetwork | 513 K | train

14 | multihead\_attn | \_InterpretableMultiHeadAttention | 410 K | train

15 | post\_attn\_gan | \_GateAddNorm | 206 K | train

16 | feed\_forward\_block | \_GatedResidualNetwork | 411 K | train

17 | pre\_output\_gan | \_GateAddNorm | 206 K | train

18 | output\_layer | Linear | 963 | train

------------------------------------------------------------------------------------------------

10.3 M Trainable params

0 Non-trainable params

10.3 M Total params

41.341 Total estimated model params size (MB)

3056 Modules in train mode

0 Modules in eval mode

🚀 Startar ny gridsearch-körning med hyperparametrar:

input\_chunk\_length: 96

output\_chunk\_length: 48

hidden\_size: 320

lstm\_layers: 3

num\_attention\_heads: 1

dropout: 0.4

optimizer\_kwargs: {'lr': 0.01}

Epoch 0: 100%|██████████| 205/205 [05:09<00:00, 0.66it/s, train\_loss=0.857, val\_loss=0.645]📉 Epoch 1: train\_loss=0.8572 | val\_loss=0.6451

Epoch 0: 100%|██████████| 205/205 [05:09<00:00, 0.66it/s, train\_loss=0.857, val\_loss=0.645]Metric val\_loss improved. New best score: 0.645

Epoch 1: 100%|██████████| 205/205 [05:09<00:00, 0.66it/s, train\_loss=0.607, val\_loss=1.320]📉 Epoch 2: train\_loss=0.6069 | val\_loss=1.3190

Epoch 2: 100%|██████████| 205/205 [05:09<00:00, 0.66it/s, train\_loss=0.465, val\_loss=0.641]📉 Epoch 3: train\_loss=0.4645 | val\_loss=0.6415

Epoch 3: 100%|██████████| 205/205 [05:09<00:00, 0.66it/s, train\_loss=0.828, val\_loss=0.749]📉 Epoch 4: train\_loss=0.8284 | val\_loss=0.7491

Epoch 3: 100%|██████████| 205/205 [05:09<00:00, 0.66it/s, train\_loss=0.828, val\_loss=0.749]Monitored metric val\_loss did not improve in the last 3 records. Best score: 0.645. Signaling Trainer to stop.

Epoch 3: 100%|██████████| 205/205 [05:09<00:00, 0.66it/s, train\_loss=0.828, val\_loss=0.749]

`predict()` was called with `n > output\_chunk\_length`: using auto-regression to forecast the values after `output\_chunk\_length` points. The model will access `(n - output\_chunk\_length)` future values of your `past\_covariates` (relative to the first predicted time step). To hide this warning, set `show\_warnings=False`.

GPU available: True (cuda), used: True

TPU available: False, using: 0 TPU cores

HPU available: False, using: 0 HPUs

LOCAL\_RANK: 0 - CUDA\_VISIBLE\_DEVICES: [0]

Predicting DataLoader 0: 100%|██████████| 1/1 [00:12<00:00, 0.08it/s][I 2025-05-10 16:28:54,780] Trial 6 finished with value: 562.4006958007812 and parameters: {'input\_chunk\_length': 96, 'output\_chunk\_length': 48, 'hidden\_size': 320, 'lstm\_layers': 3, 'num\_attention\_heads': 1, 'dropout': 0.4, 'n\_epochs': 30, 'lr': 0.01, 'batch\_size': 128}. Best is trial 0 with value: 284.9240417480469.

🔍 DEBUG: Dataövergripande information:

▶️ train: 2021-01-02 00:00:00 → 2023-12-31 23:00:00 (26256 timsteg)

✅ val: 2024-01-01 00:00:00 → 2024-12-31 23:00:00 (8784 timsteg)

⏪ past\_cov\_train: 2021-01-02 00:00:00 → 2023-12-31 23:00:00

⏪ past\_cov\_val: 2023-12-28 00:00:00 → 2024-12-31 23:00:00

⏩ future\_cov\_train: 2021-01-02 00:00:00 → 2023-12-31 23:00:00

⏩ future\_cov\_val: 2023-12-28 00:00:00 → 2024-12-31 23:00:00

GPU available: True (cuda), used: True

TPU available: False, using: 0 TPU cores

HPU available: False, using: 0 HPUs

LOCAL\_RANK: 0 - CUDA\_VISIBLE\_DEVICES: [0]

| Name | Type | Params | Mode

------------------------------------------------------------------------------------------------

0 | train\_metrics | MetricCollection | 0 | train

1 | val\_metrics | MetricCollection | 0 | train

2 | input\_embeddings | \_MultiEmbedding | 0 | train

3 | static\_covariates\_vsn | \_VariableSelectionNetwork | 0 | train

4 | encoder\_vsn | \_VariableSelectionNetwork | 727 K | train

5 | decoder\_vsn | \_VariableSelectionNetwork | 312 K | train

6 | static\_context\_grn | \_GatedResidualNetwork | 103 K | train

7 | static\_context\_hidden\_encoder\_grn | \_GatedResidualNetwork | 103 K | train

8 | static\_context\_cell\_encoder\_grn | \_GatedResidualNetwork | 103 K | train

9 | static\_context\_enrichment | \_GatedResidualNetwork | 103 K | train

10 | lstm\_encoder | LSTM | 412 K | train

11 | lstm\_decoder | LSTM | 412 K | train

12 | post\_lstm\_gan | \_GateAddNorm | 51.8 K | train

13 | static\_enrichment\_grn | \_GatedResidualNetwork | 128 K | train

14 | multihead\_attn | \_InterpretableMultiHeadAttention | 64.4 K | train

15 | post\_attn\_gan | \_GateAddNorm | 51.8 K | train

16 | feed\_forward\_block | \_GatedResidualNetwork | 103 K | train

17 | pre\_output\_gan | \_GateAddNorm | 51.8 K | train

18 | output\_layer | Linear | 483 | train

------------------------------------------------------------------------------------------------

2.7 M Trainable params

0 Non-trainable params

2.7 M Total params

10.919 Total estimated model params size (MB)

3062 Modules in train mode

0 Modules in eval mode

🚀 Startar ny gridsearch-körning med hyperparametrar:

input\_chunk\_length: 96

output\_chunk\_length: 48

hidden\_size: 160

lstm\_layers: 2

num\_attention\_heads: 4

dropout: 0.2

optimizer\_kwargs: {'lr': 0.001}

Epoch 0: 100%|██████████| 205/205 [03:50<00:00, 0.89it/s, train\_loss=0.681, val\_loss=0.652]📉 Epoch 1: train\_loss=0.6811 | val\_loss=0.6520

Epoch 0: 100%|██████████| 205/205 [03:50<00:00, 0.89it/s, train\_loss=0.681, val\_loss=0.652]Metric val\_loss improved. New best score: 0.652

Epoch 1: 100%|██████████| 205/205 [03:49<00:00, 0.89it/s, train\_loss=0.227, val\_loss=0.708]📉 Epoch 2: train\_loss=0.2266 | val\_loss=0.7078

Epoch 2: 100%|██████████| 205/205 [03:50<00:00, 0.89it/s, train\_loss=0.267, val\_loss=0.465]📉 Epoch 3: train\_loss=0.2667 | val\_loss=0.4653

Epoch 2: 100%|██████████| 205/205 [03:50<00:00, 0.89it/s, train\_loss=0.267, val\_loss=0.465]Metric val\_loss improved by 0.187 >= min\_delta = 0.01. New best score: 0.465

Epoch 3: 100%|██████████| 205/205 [03:49<00:00, 0.89it/s, train\_loss=0.471, val\_loss=0.590]📉 Epoch 4: train\_loss=0.4714 | val\_loss=0.5897

Epoch 4: 100%|██████████| 205/205 [03:49<00:00, 0.89it/s, train\_loss=0.192, val\_loss=0.506]📉 Epoch 5: train\_loss=0.1915 | val\_loss=0.5065

Epoch 5: 100%|██████████| 205/205 [03:49<00:00, 0.89it/s, train\_loss=0.226, val\_loss=0.465]📉 Epoch 6: train\_loss=0.2255 | val\_loss=0.4645

Epoch 5: 100%|██████████| 205/205 [03:49<00:00, 0.89it/s, train\_loss=0.226, val\_loss=0.465]Monitored metric val\_loss did not improve in the last 3 records. Best score: 0.465. Signaling Trainer to stop.

Epoch 5: 100%|██████████| 205/205 [03:49<00:00, 0.89it/s, train\_loss=0.226, val\_loss=0.465]

`predict()` was called with `n > output\_chunk\_length`: using auto-regression to forecast the values after `output\_chunk\_length` points. The model will access `(n - output\_chunk\_length)` future values of your `past\_covariates` (relative to the first predicted time step). To hide this warning, set `show\_warnings=False`.

GPU available: True (cuda), used: True

TPU available: False, using: 0 TPU cores

HPU available: False, using: 0 HPUs

LOCAL\_RANK: 0 - CUDA\_VISIBLE\_DEVICES: [0]

Predicting DataLoader 0: 100%|██████████| 1/1 [00:12<00:00, 0.08it/s][I 2025-05-10 16:52:08,352] Trial 7 finished with value: 333.6116943359375 and parameters: {'input\_chunk\_length': 96, 'output\_chunk\_length': 48, 'hidden\_size': 160, 'lstm\_layers': 2, 'num\_attention\_heads': 4, 'dropout': 0.2, 'n\_epochs': 30, 'lr': 0.001, 'batch\_size': 128}. Best is trial 0 with value: 284.9240417480469.

🔍 DEBUG: Dataövergripande information:

▶️ train: 2021-01-02 00:00:00 → 2023-12-31 23:00:00 (26256 timsteg)

✅ val: 2024-01-01 00:00:00 → 2024-12-31 23:00:00 (8784 timsteg)

⏪ past\_cov\_train: 2021-01-02 00:00:00 → 2023-12-31 23:00:00

⏪ past\_cov\_val: 2023-12-28 00:00:00 → 2024-12-31 23:00:00

⏩ future\_cov\_train: 2021-01-02 00:00:00 → 2023-12-31 23:00:00

⏩ future\_cov\_val: 2023-12-28 00:00:00 → 2024-12-31 23:00:00

GPU available: True (cuda), used: True

TPU available: False, using: 0 TPU cores

HPU available: False, using: 0 HPUs

LOCAL\_RANK: 0 - CUDA\_VISIBLE\_DEVICES: [0]

| Name | Type | Params | Mode

------------------------------------------------------------------------------------------------

0 | train\_metrics | MetricCollection | 0 | train

1 | val\_metrics | MetricCollection | 0 | train

2 | input\_embeddings | \_MultiEmbedding | 0 | train

3 | static\_covariates\_vsn | \_VariableSelectionNetwork | 0 | train

4 | encoder\_vsn | \_VariableSelectionNetwork | 983 K | train

5 | decoder\_vsn | \_VariableSelectionNetwork | 439 K | train

6 | static\_context\_grn | \_GatedResidualNetwork | 231 K | train

7 | static\_context\_hidden\_encoder\_grn | \_GatedResidualNetwork | 231 K | train

8 | static\_context\_cell\_encoder\_grn | \_GatedResidualNetwork | 231 K | train

9 | static\_context\_enrichment | \_GatedResidualNetwork | 231 K | train

10 | lstm\_encoder | LSTM | 1.4 M | train

11 | lstm\_decoder | LSTM | 1.4 M | train

12 | post\_lstm\_gan | \_GateAddNorm | 116 K | train

13 | static\_enrichment\_grn | \_GatedResidualNetwork | 289 K | train

14 | multihead\_attn | \_InterpretableMultiHeadAttention | 144 K | train

15 | post\_attn\_gan | \_GateAddNorm | 116 K | train

16 | feed\_forward\_block | \_GatedResidualNetwork | 231 K | train

17 | pre\_output\_gan | \_GateAddNorm | 116 K | train

18 | output\_layer | Linear | 723 | train

------------------------------------------------------------------------------------------------

6.1 M Trainable params

0 Non-trainable params

6.1 M Total params

24.560 Total estimated model params size (MB)

3062 Modules in train mode

0 Modules in eval mode

🚀 Startar ny gridsearch-körning med hyperparametrar:

input\_chunk\_length: 72

output\_chunk\_length: 48

hidden\_size: 240

lstm\_layers: 3

num\_attention\_heads: 4

dropout: 0.1

optimizer\_kwargs: {'lr': 0.0001}

Epoch 0: 100%|██████████| 205/205 [03:44<00:00, 0.91it/s, train\_loss=0.573, val\_loss=0.662]📉 Epoch 1: train\_loss=0.5732 | val\_loss=0.6622

Epoch 0: 100%|██████████| 205/205 [03:44<00:00, 0.91it/s, train\_loss=0.573, val\_loss=0.662]Metric val\_loss improved. New best score: 0.662

Epoch 1: 100%|██████████| 205/205 [03:44<00:00, 0.91it/s, train\_loss=0.424, val\_loss=0.573]📉 Epoch 2: train\_loss=0.4242 | val\_loss=0.5729

Epoch 1: 100%|██████████| 205/205 [03:44<00:00, 0.91it/s, train\_loss=0.424, val\_loss=0.573]Metric val\_loss improved by 0.089 >= min\_delta = 0.01. New best score: 0.573

Epoch 2: 100%|██████████| 205/205 [03:44<00:00, 0.91it/s, train\_loss=0.407, val\_loss=0.461]📉 Epoch 3: train\_loss=0.4070 | val\_loss=0.4607

Epoch 2: 100%|██████████| 205/205 [03:44<00:00, 0.91it/s, train\_loss=0.407, val\_loss=0.461]Metric val\_loss improved by 0.112 >= min\_delta = 0.01. New best score: 0.461

Epoch 3: 100%|██████████| 205/205 [03:44<00:00, 0.91it/s, train\_loss=0.425, val\_loss=0.535]📉 Epoch 4: train\_loss=0.4248 | val\_loss=0.5352

Epoch 4: 100%|██████████| 205/205 [03:44<00:00, 0.91it/s, train\_loss=0.406, val\_loss=0.472]📉 Epoch 5: train\_loss=0.4057 | val\_loss=0.4725

Epoch 5: 100%|██████████| 205/205 [03:44<00:00, 0.91it/s, train\_loss=0.369, val\_loss=0.406]📉 Epoch 6: train\_loss=0.3689 | val\_loss=0.4064

Epoch 5: 100%|██████████| 205/205 [03:44<00:00, 0.91it/s, train\_loss=0.369, val\_loss=0.406]Metric val\_loss improved by 0.054 >= min\_delta = 0.01. New best score: 0.406

Epoch 6: 100%|██████████| 205/205 [03:44<00:00, 0.91it/s, train\_loss=0.271, val\_loss=0.402]📉 Epoch 7: train\_loss=0.2709 | val\_loss=0.4023

Epoch 7: 100%|██████████| 205/205 [03:44<00:00, 0.91it/s, train\_loss=0.373, val\_loss=0.442]📉 Epoch 8: train\_loss=0.3731 | val\_loss=0.4419

Epoch 8: 100%|██████████| 205/205 [03:44<00:00, 0.91it/s, train\_loss=0.340, val\_loss=0.422]📉 Epoch 9: train\_loss=0.3398 | val\_loss=0.4220

Epoch 8: 100%|██████████| 205/205 [03:44<00:00, 0.91it/s, train\_loss=0.340, val\_loss=0.422]Monitored metric val\_loss did not improve in the last 3 records. Best score: 0.406. Signaling Trainer to stop.

Epoch 8: 100%|██████████| 205/205 [03:44<00:00, 0.91it/s, train\_loss=0.340, val\_loss=0.422]

`predict()` was called with `n > output\_chunk\_length`: using auto-regression to forecast the values after `output\_chunk\_length` points. The model will access `(n - output\_chunk\_length)` future values of your `past\_covariates` (relative to the first predicted time step). To hide this warning, set `show\_warnings=False`.

GPU available: True (cuda), used: True

TPU available: False, using: 0 TPU cores

HPU available: False, using: 0 HPUs

LOCAL\_RANK: 0 - CUDA\_VISIBLE\_DEVICES: [0]

Predicting DataLoader 0: 100%|██████████| 1/1 [00:12<00:00, 0.08it/s][I 2025-05-10 17:26:06,450] Trial 8 finished with value: 314.4120788574219 and parameters: {'input\_chunk\_length': 72, 'output\_chunk\_length': 48, 'hidden\_size': 240, 'lstm\_layers': 3, 'num\_attention\_heads': 4, 'dropout': 0.1, 'n\_epochs': 30, 'lr': 0.0001, 'batch\_size': 128}. Best is trial 0 with value: 284.9240417480469.

🔍 DEBUG: Dataövergripande information:

▶️ train: 2021-01-02 00:00:00 → 2023-12-31 23:00:00 (26256 timsteg)

✅ val: 2024-01-01 00:00:00 → 2024-12-31 23:00:00 (8784 timsteg)

⏪ past\_cov\_train: 2021-01-02 00:00:00 → 2023-12-31 23:00:00

⏪ past\_cov\_val: 2023-12-28 00:00:00 → 2024-12-31 23:00:00

⏩ future\_cov\_train: 2021-01-02 00:00:00 → 2023-12-31 23:00:00

⏩ future\_cov\_val: 2023-12-28 00:00:00 → 2024-12-31 23:00:00

GPU available: True (cuda), used: True

TPU available: False, using: 0 TPU cores

HPU available: False, using: 0 HPUs

LOCAL\_RANK: 0 - CUDA\_VISIBLE\_DEVICES: [0]

| Name | Type | Params | Mode

------------------------------------------------------------------------------------------------

0 | train\_metrics | MetricCollection | 0 | train

1 | val\_metrics | MetricCollection | 0 | train

2 | input\_embeddings | \_MultiEmbedding | 0 | train

3 | static\_covariates\_vsn | \_VariableSelectionNetwork | 0 | train

4 | encoder\_vsn | \_VariableSelectionNetwork | 386 K | train

5 | decoder\_vsn | \_VariableSelectionNetwork | 185 K | train

6 | static\_context\_grn | \_GatedResidualNetwork | 26.1 K | train

7 | static\_context\_hidden\_encoder\_grn | \_GatedResidualNetwork | 26.1 K | train

8 | static\_context\_cell\_encoder\_grn | \_GatedResidualNetwork | 26.1 K | train

9 | static\_context\_enrichment | \_GatedResidualNetwork | 26.1 K | train

10 | lstm\_encoder | LSTM | 155 K | train

11 | lstm\_decoder | LSTM | 155 K | train

12 | post\_lstm\_gan | \_GateAddNorm | 13.1 K | train

13 | static\_enrichment\_grn | \_GatedResidualNetwork | 32.5 K | train

14 | multihead\_attn | \_InterpretableMultiHeadAttention | 25.8 K | train

15 | post\_attn\_gan | \_GateAddNorm | 13.1 K | train

16 | feed\_forward\_block | \_GatedResidualNetwork | 26.1 K | train

17 | pre\_output\_gan | \_GateAddNorm | 13.1 K | train

18 | output\_layer | Linear | 243 | train

------------------------------------------------------------------------------------------------

1.1 M Trainable params

0 Non-trainable params

1.1 M Total params

4.442 Total estimated model params size (MB)

3056 Modules in train mode

0 Modules in eval mode

🚀 Startar ny gridsearch-körning med hyperparametrar:

input\_chunk\_length: 72

output\_chunk\_length: 48

hidden\_size: 80

lstm\_layers: 3

num\_attention\_heads: 1

dropout: 0.3

optimizer\_kwargs: {'lr': 0.001}

Epoch 0: 100%|██████████| 205/205 [02:32<00:00, 1.35it/s, train\_loss=0.711, val\_loss=0.662]📉 Epoch 1: train\_loss=0.7109 | val\_loss=0.6624

Epoch 0: 100%|██████████| 205/205 [02:32<00:00, 1.35it/s, train\_loss=0.711, val\_loss=0.662]Metric val\_loss improved. New best score: 0.662

Epoch 1: 100%|██████████| 205/205 [02:31<00:00, 1.35it/s, train\_loss=0.444, val\_loss=0.535]📉 Epoch 2: train\_loss=0.4443 | val\_loss=0.5349

Epoch 1: 100%|██████████| 205/205 [02:31<00:00, 1.35it/s, train\_loss=0.444, val\_loss=0.535]Metric val\_loss improved by 0.128 >= min\_delta = 0.01. New best score: 0.535

Epoch 2: 100%|██████████| 205/205 [02:31<00:00, 1.35it/s, train\_loss=0.420, val\_loss=0.458]📉 Epoch 3: train\_loss=0.4203 | val\_loss=0.4582

Epoch 2: 100%|██████████| 205/205 [02:31<00:00, 1.35it/s, train\_loss=0.420, val\_loss=0.458]Metric val\_loss improved by 0.077 >= min\_delta = 0.01. New best score: 0.458

Epoch 3: 100%|██████████| 205/205 [02:32<00:00, 1.35it/s, train\_loss=0.381, val\_loss=0.498]📉 Epoch 4: train\_loss=0.3808 | val\_loss=0.4976

Epoch 4: 100%|██████████| 205/205 [02:32<00:00, 1.35it/s, train\_loss=0.392, val\_loss=0.430]📉 Epoch 5: train\_loss=0.3919 | val\_loss=0.4305

Epoch 4: 100%|██████████| 205/205 [02:32<00:00, 1.35it/s, train\_loss=0.392, val\_loss=0.430]Metric val\_loss improved by 0.028 >= min\_delta = 0.01. New best score: 0.430

Epoch 5: 100%|██████████| 205/205 [02:32<00:00, 1.35it/s, train\_loss=0.329, val\_loss=0.446]📉 Epoch 6: train\_loss=0.3294 | val\_loss=0.4460

Epoch 6: 100%|██████████| 205/205 [02:32<00:00, 1.35it/s, train\_loss=0.334, val\_loss=0.455]📉 Epoch 7: train\_loss=0.3342 | val\_loss=0.4546

Epoch 7: 100%|██████████| 205/205 [02:32<00:00, 1.35it/s, train\_loss=0.334, val\_loss=0.437]📉 Epoch 8: train\_loss=0.3339 | val\_loss=0.4375

Epoch 7: 100%|██████████| 205/205 [02:32<00:00, 1.35it/s, train\_loss=0.334, val\_loss=0.437]Monitored metric val\_loss did not improve in the last 3 records. Best score: 0.430. Signaling Trainer to stop.

Epoch 7: 100%|██████████| 205/205 [02:32<00:00, 1.35it/s, train\_loss=0.334, val\_loss=0.437]

`predict()` was called with `n > output\_chunk\_length`: using auto-regression to forecast the values after `output\_chunk\_length` points. The model will access `(n - output\_chunk\_length)` future values of your `past\_covariates` (relative to the first predicted time step). To hide this warning, set `show\_warnings=False`.

GPU available: True (cuda), used: True

TPU available: False, using: 0 TPU cores

HPU available: False, using: 0 HPUs

LOCAL\_RANK: 0 - CUDA\_VISIBLE\_DEVICES: [0]

Predicting DataLoader 0: 100%|██████████| 1/1 [00:12<00:00, 0.08it/s][I 2025-05-10 17:46:36,990] Trial 9 finished with value: 297.21044921875 and parameters: {'input\_chunk\_length': 72, 'output\_chunk\_length': 48, 'hidden\_size': 80, 'lstm\_layers': 3, 'num\_attention\_heads': 1, 'dropout': 0.3, 'n\_epochs': 30, 'lr': 0.001, 'batch\_size': 128}. Best is trial 0 with value: 284.9240417480469.

🔍 DEBUG: Dataövergripande information:

▶️ train: 2021-01-02 00:00:00 → 2023-12-31 23:00:00 (26256 timsteg)

✅ val: 2024-01-01 00:00:00 → 2024-12-31 23:00:00 (8784 timsteg)

⏪ past\_cov\_train: 2021-01-02 00:00:00 → 2023-12-31 23:00:00

⏪ past\_cov\_val: 2023-12-28 00:00:00 → 2024-12-31 23:00:00

⏩ future\_cov\_train: 2021-01-02 00:00:00 → 2023-12-31 23:00:00

⏩ future\_cov\_val: 2023-12-28 00:00:00 → 2024-12-31 23:00:00

GPU available: True (cuda), used: True

TPU available: False, using: 0 TPU cores

HPU available: False, using: 0 HPUs

LOCAL\_RANK: 0 - CUDA\_VISIBLE\_DEVICES: [0]

| Name | Type | Params | Mode

------------------------------------------------------------------------------------------------

0 | train\_metrics | MetricCollection | 0 | train

1 | val\_metrics | MetricCollection | 0 | train

2 | input\_embeddings | \_MultiEmbedding | 0 | train

3 | static\_covariates\_vsn | \_VariableSelectionNetwork | 0 | train

4 | encoder\_vsn | \_VariableSelectionNetwork | 201 K | train

5 | decoder\_vsn | \_VariableSelectionNetwork | 101 K | train

6 | static\_context\_grn | \_GatedResidualNetwork | 6.6 K | train

7 | static\_context\_hidden\_encoder\_grn | \_GatedResidualNetwork | 6.6 K | train

8 | static\_context\_cell\_encoder\_grn | \_GatedResidualNetwork | 6.6 K | train

9 | static\_context\_enrichment | \_GatedResidualNetwork | 6.6 K | train

10 | lstm\_encoder | LSTM | 26.2 K | train

11 | lstm\_decoder | LSTM | 26.2 K | train

12 | post\_lstm\_gan | \_GateAddNorm | 3.4 K | train

13 | static\_enrichment\_grn | \_GatedResidualNetwork | 8.2 K | train

14 | multihead\_attn | \_InterpretableMultiHeadAttention | 4.1 K | train

15 | post\_attn\_gan | \_GateAddNorm | 3.4 K | train

16 | feed\_forward\_block | \_GatedResidualNetwork | 6.6 K | train

17 | pre\_output\_gan | \_GateAddNorm | 3.4 K | train

18 | output\_layer | Linear | 123 | train

------------------------------------------------------------------------------------------------

409 K Trainable params

0 Non-trainable params

409 K Total params

1.639 Total estimated model params size (MB)

3062 Modules in train mode

0 Modules in eval mode

🚀 Startar ny gridsearch-körning med hyperparametrar:

input\_chunk\_length: 48

output\_chunk\_length: 48

hidden\_size: 40

lstm\_layers: 2

num\_attention\_heads: 4

dropout: 0.5

optimizer\_kwargs: {'lr': 0.001}

Epoch 0: 100%|██████████| 205/205 [01:49<00:00, 1.88it/s, train\_loss=0.788, val\_loss=0.786]📉 Epoch 1: train\_loss=0.7879 | val\_loss=0.7861

Epoch 0: 100%|██████████| 205/205 [01:49<00:00, 1.88it/s, train\_loss=0.788, val\_loss=0.786]Metric val\_loss improved. New best score: 0.786

Epoch 1: 100%|██████████| 205/205 [01:48<00:00, 1.88it/s, train\_loss=0.469, val\_loss=0.624]📉 Epoch 2: train\_loss=0.4694 | val\_loss=0.6243

Epoch 1: 100%|██████████| 205/205 [01:48<00:00, 1.88it/s, train\_loss=0.469, val\_loss=0.624]Metric val\_loss improved by 0.162 >= min\_delta = 0.01. New best score: 0.624

Epoch 2: 100%|██████████| 205/205 [01:49<00:00, 1.88it/s, train\_loss=0.404, val\_loss=0.502]📉 Epoch 3: train\_loss=0.4042 | val\_loss=0.5022

Epoch 2: 100%|██████████| 205/205 [01:49<00:00, 1.88it/s, train\_loss=0.404, val\_loss=0.502]Metric val\_loss improved by 0.122 >= min\_delta = 0.01. New best score: 0.502

Epoch 3: 100%|██████████| 205/205 [01:49<00:00, 1.88it/s, train\_loss=0.429, val\_loss=0.462]📉 Epoch 4: train\_loss=0.4291 | val\_loss=0.4618

Epoch 3: 100%|██████████| 205/205 [01:49<00:00, 1.88it/s, train\_loss=0.429, val\_loss=0.462]Metric val\_loss improved by 0.040 >= min\_delta = 0.01. New best score: 0.462

Epoch 4: 100%|██████████| 205/205 [01:48<00:00, 1.88it/s, train\_loss=0.343, val\_loss=0.457]📉 Epoch 5: train\_loss=0.3430 | val\_loss=0.4566

Epoch 5: 100%|██████████| 205/205 [01:48<00:00, 1.88it/s, train\_loss=0.410, val\_loss=0.410]📉 Epoch 6: train\_loss=0.4100 | val\_loss=0.4100

Epoch 5: 100%|██████████| 205/205 [01:48<00:00, 1.88it/s, train\_loss=0.410, val\_loss=0.410]Metric val\_loss improved by 0.052 >= min\_delta = 0.01. New best score: 0.410

Epoch 6: 100%|██████████| 205/205 [01:48<00:00, 1.88it/s, train\_loss=0.369, val\_loss=0.429]📉 Epoch 7: train\_loss=0.3689 | val\_loss=0.4290

Epoch 7: 100%|██████████| 205/205 [01:48<00:00, 1.88it/s, train\_loss=0.384, val\_loss=0.427]📉 Epoch 8: train\_loss=0.3840 | val\_loss=0.4270

Epoch 8: 100%|██████████| 205/205 [01:48<00:00, 1.89it/s, train\_loss=0.371, val\_loss=0.396]📉 Epoch 9: train\_loss=0.3705 | val\_loss=0.3956

Epoch 8: 100%|██████████| 205/205 [01:48<00:00, 1.89it/s, train\_loss=0.371, val\_loss=0.396]Metric val\_loss improved by 0.014 >= min\_delta = 0.01. New best score: 0.396

Epoch 9: 100%|██████████| 205/205 [01:48<00:00, 1.89it/s, train\_loss=0.353, val\_loss=0.477]📉 Epoch 10: train\_loss=0.3532 | val\_loss=0.4772

Epoch 10: 100%|██████████| 205/205 [01:48<00:00, 1.88it/s, train\_loss=0.366, val\_loss=0.405]📉 Epoch 11: train\_loss=0.3659 | val\_loss=0.4051

Epoch 11: 100%|██████████| 205/205 [01:48<00:00, 1.89it/s, train\_loss=0.333, val\_loss=0.503]📉 Epoch 12: train\_loss=0.3325 | val\_loss=0.5028

Epoch 11: 100%|██████████| 205/205 [01:48<00:00, 1.89it/s, train\_loss=0.333, val\_loss=0.503]Monitored metric val\_loss did not improve in the last 3 records. Best score: 0.396. Signaling Trainer to stop.

Epoch 11: 100%|██████████| 205/205 [01:48<00:00, 1.89it/s, train\_loss=0.333, val\_loss=0.503]

`predict()` was called with `n > output\_chunk\_length`: using auto-regression to forecast the values after `output\_chunk\_length` points. The model will access `(n - output\_chunk\_length)` future values of your `past\_covariates` (relative to the first predicted time step). To hide this warning, set `show\_warnings=False`.

GPU available: True (cuda), used: True

TPU available: False, using: 0 TPU cores

HPU available: False, using: 0 HPUs

LOCAL\_RANK: 0 - CUDA\_VISIBLE\_DEVICES: [0]

Predicting DataLoader 0: 100%|██████████| 1/1 [00:12<00:00, 0.08it/s][I 2025-05-10 18:08:37,513] Trial 10 finished with value: 343.7170715332031 and parameters: {'input\_chunk\_length': 48, 'output\_chunk\_length': 48, 'hidden\_size': 40, 'lstm\_layers': 2, 'num\_attention\_heads': 4, 'dropout': 0.5, 'n\_epochs': 30, 'lr': 0.001, 'batch\_size': 128}. Best is trial 0 with value: 284.9240417480469.

🔍 DEBUG: Dataövergripande information:

▶️ train: 2021-01-02 00:00:00 → 2023-12-31 23:00:00 (26256 timsteg)

✅ val: 2024-01-01 00:00:00 → 2024-12-31 23:00:00 (8784 timsteg)

⏪ past\_cov\_train: 2021-01-02 00:00:00 → 2023-12-31 23:00:00

⏪ past\_cov\_val: 2023-12-28 00:00:00 → 2024-12-31 23:00:00

⏩ future\_cov\_train: 2021-01-02 00:00:00 → 2023-12-31 23:00:00

⏩ future\_cov\_val: 2023-12-28 00:00:00 → 2024-12-31 23:00:00

GPU available: True (cuda), used: True

TPU available: False, using: 0 TPU cores

HPU available: False, using: 0 HPUs

LOCAL\_RANK: 0 - CUDA\_VISIBLE\_DEVICES: [0]

| Name | Type | Params | Mode

------------------------------------------------------------------------------------------------

0 | train\_metrics | MetricCollection | 0 | train

1 | val\_metrics | MetricCollection | 0 | train

2 | input\_embeddings | \_MultiEmbedding | 0 | train

3 | static\_covariates\_vsn | \_VariableSelectionNetwork | 0 | train

4 | encoder\_vsn | \_VariableSelectionNetwork | 386 K | train

5 | decoder\_vsn | \_VariableSelectionNetwork | 185 K | train

6 | static\_context\_grn | \_GatedResidualNetwork | 26.1 K | train

7 | static\_context\_hidden\_encoder\_grn | \_GatedResidualNetwork | 26.1 K | train

8 | static\_context\_cell\_encoder\_grn | \_GatedResidualNetwork | 26.1 K | train

9 | static\_context\_enrichment | \_GatedResidualNetwork | 26.1 K | train

10 | lstm\_encoder | LSTM | 155 K | train

11 | lstm\_decoder | LSTM | 155 K | train

12 | post\_lstm\_gan | \_GateAddNorm | 13.1 K | train

13 | static\_enrichment\_grn | \_GatedResidualNetwork | 32.5 K | train

14 | multihead\_attn | \_InterpretableMultiHeadAttention | 16.2 K | train

15 | post\_attn\_gan | \_GateAddNorm | 13.1 K | train

16 | feed\_forward\_block | \_GatedResidualNetwork | 26.1 K | train

17 | pre\_output\_gan | \_GateAddNorm | 13.1 K | train

18 | output\_layer | Linear | 243 | train

------------------------------------------------------------------------------------------------

1.1 M Trainable params

0 Non-trainable params

1.1 M Total params

4.403 Total estimated model params size (MB)

3062 Modules in train mode

0 Modules in eval mode

🚀 Startar ny gridsearch-körning med hyperparametrar:

input\_chunk\_length: 72

output\_chunk\_length: 48

hidden\_size: 80

lstm\_layers: 3

num\_attention\_heads: 4

dropout: 0.3

optimizer\_kwargs: {'lr': 0.001}

Epoch 0: 100%|██████████| 205/205 [02:31<00:00, 1.35it/s, train\_loss=0.473, val\_loss=0.599]📉 Epoch 1: train\_loss=0.4726 | val\_loss=0.5988

Epoch 0: 100%|██████████| 205/205 [02:31<00:00, 1.35it/s, train\_loss=0.473, val\_loss=0.599]Metric val\_loss improved. New best score: 0.599

Epoch 1: 100%|██████████| 205/205 [02:31<00:00, 1.35it/s, train\_loss=0.490, val\_loss=0.480]📉 Epoch 2: train\_loss=0.4904 | val\_loss=0.4800

Epoch 1: 100%|██████████| 205/205 [02:31<00:00, 1.35it/s, train\_loss=0.490, val\_loss=0.480]Metric val\_loss improved by 0.119 >= min\_delta = 0.01. New best score: 0.480

Epoch 2: 100%|██████████| 205/205 [02:31<00:00, 1.35it/s, train\_loss=0.389, val\_loss=0.416]📉 Epoch 3: train\_loss=0.3894 | val\_loss=0.4163

Epoch 2: 100%|██████████| 205/205 [02:31<00:00, 1.35it/s, train\_loss=0.389, val\_loss=0.416]Metric val\_loss improved by 0.064 >= min\_delta = 0.01. New best score: 0.416

Epoch 3: 100%|██████████| 205/205 [02:32<00:00, 1.35it/s, train\_loss=0.359, val\_loss=0.477]📉 Epoch 4: train\_loss=0.3590 | val\_loss=0.4773

Epoch 4: 100%|██████████| 205/205 [02:31<00:00, 1.35it/s, train\_loss=0.391, val\_loss=0.409]📉 Epoch 5: train\_loss=0.3905 | val\_loss=0.4089

Epoch 5: 100%|██████████| 205/205 [02:31<00:00, 1.35it/s, train\_loss=0.334, val\_loss=0.394]📉 Epoch 6: train\_loss=0.3338 | val\_loss=0.3939

Epoch 5: 100%|██████████| 205/205 [02:31<00:00, 1.35it/s, train\_loss=0.334, val\_loss=0.394]Metric val\_loss improved by 0.022 >= min\_delta = 0.01. New best score: 0.394

Epoch 6: 100%|██████████| 205/205 [02:31<00:00, 1.35it/s, train\_loss=0.314, val\_loss=0.393]📉 Epoch 7: train\_loss=0.3137 | val\_loss=0.3929

Epoch 7: 100%|██████████| 205/205 [02:31<00:00, 1.35it/s, train\_loss=0.305, val\_loss=0.394]📉 Epoch 8: train\_loss=0.3050 | val\_loss=0.3945

Epoch 8: 100%|██████████| 205/205 [02:32<00:00, 1.35it/s, train\_loss=0.310, val\_loss=0.459]📉 Epoch 9: train\_loss=0.3103 | val\_loss=0.4591

Epoch 8: 100%|██████████| 205/205 [02:32<00:00, 1.35it/s, train\_loss=0.310, val\_loss=0.459]Monitored metric val\_loss did not improve in the last 3 records. Best score: 0.394. Signaling Trainer to stop.

Epoch 8: 100%|██████████| 205/205 [02:32<00:00, 1.35it/s, train\_loss=0.310, val\_loss=0.459]

`predict()` was called with `n > output\_chunk\_length`: using auto-regression to forecast the values after `output\_chunk\_length` points. The model will access `(n - output\_chunk\_length)` future values of your `past\_covariates` (relative to the first predicted time step). To hide this warning, set `show\_warnings=False`.

GPU available: True (cuda), used: True

TPU available: False, using: 0 TPU cores

HPU available: False, using: 0 HPUs

LOCAL\_RANK: 0 - CUDA\_VISIBLE\_DEVICES: [0]

Predicting DataLoader 0: 100%|██████████| 1/1 [00:12<00:00, 0.08it/s][I 2025-05-10 18:31:38,776] Trial 11 finished with value: 309.7366943359375 and parameters: {'input\_chunk\_length': 72, 'output\_chunk\_length': 48, 'hidden\_size': 80, 'lstm\_layers': 3, 'num\_attention\_heads': 4, 'dropout': 0.3, 'n\_epochs': 30, 'lr': 0.001, 'batch\_size': 128}. Best is trial 0 with value: 284.9240417480469.

🔍 DEBUG: Dataövergripande information:

▶️ train: 2021-01-02 00:00:00 → 2023-12-31 23:00:00 (26256 timsteg)

✅ val: 2024-01-01 00:00:00 → 2024-12-31 23:00:00 (8784 timsteg)

⏪ past\_cov\_train: 2021-01-02 00:00:00 → 2023-12-31 23:00:00

⏪ past\_cov\_val: 2023-12-28 00:00:00 → 2024-12-31 23:00:00

⏩ future\_cov\_train: 2021-01-02 00:00:00 → 2023-12-31 23:00:00

⏩ future\_cov\_val: 2023-12-28 00:00:00 → 2024-12-31 23:00:00

GPU available: True (cuda), used: True

TPU available: False, using: 0 TPU cores

HPU available: False, using: 0 HPUs

LOCAL\_RANK: 0 - CUDA\_VISIBLE\_DEVICES: [0]

| Name | Type | Params | Mode

------------------------------------------------------------------------------------------------

0 | train\_metrics | MetricCollection | 0 | train

1 | val\_metrics | MetricCollection | 0 | train

2 | input\_embeddings | \_MultiEmbedding | 0 | train

3 | static\_covariates\_vsn | \_VariableSelectionNetwork | 0 | train

4 | encoder\_vsn | \_VariableSelectionNetwork | 386 K | train

5 | decoder\_vsn | \_VariableSelectionNetwork | 185 K | train

6 | static\_context\_grn | \_GatedResidualNetwork | 26.1 K | train

7 | static\_context\_hidden\_encoder\_grn | \_GatedResidualNetwork | 26.1 K | train

8 | static\_context\_cell\_encoder\_grn | \_GatedResidualNetwork | 26.1 K | train

9 | static\_context\_enrichment | \_GatedResidualNetwork | 26.1 K | train

10 | lstm\_encoder | LSTM | 155 K | train

11 | lstm\_decoder | LSTM | 155 K | train

12 | post\_lstm\_gan | \_GateAddNorm | 13.1 K | train

13 | static\_enrichment\_grn | \_GatedResidualNetwork | 32.5 K | train

14 | multihead\_attn | \_InterpretableMultiHeadAttention | 16.2 K | train

15 | post\_attn\_gan | \_GateAddNorm | 13.1 K | train

16 | feed\_forward\_block | \_GatedResidualNetwork | 26.1 K | train

17 | pre\_output\_gan | \_GateAddNorm | 13.1 K | train

18 | output\_layer | Linear | 243 | train

------------------------------------------------------------------------------------------------

1.1 M Trainable params

0 Non-trainable params

1.1 M Total params

4.403 Total estimated model params size (MB)

3062 Modules in train mode

0 Modules in eval mode

🚀 Startar ny gridsearch-körning med hyperparametrar:

input\_chunk\_length: 72

output\_chunk\_length: 48

hidden\_size: 80

lstm\_layers: 3

num\_attention\_heads: 4

dropout: 0.3

optimizer\_kwargs: {'lr': 0.001}

Epoch 0: 100%|██████████| 205/205 [02:31<00:00, 1.35it/s, train\_loss=0.467, val\_loss=0.615]📉 Epoch 1: train\_loss=0.4670 | val\_loss=0.6145

Epoch 0: 100%|██████████| 205/205 [02:31<00:00, 1.35it/s, train\_loss=0.467, val\_loss=0.615]Metric val\_loss improved. New best score: 0.615

Epoch 1: 100%|██████████| 205/205 [02:31<00:00, 1.35it/s, train\_loss=0.492, val\_loss=0.490]📉 Epoch 2: train\_loss=0.4924 | val\_loss=0.4905

Epoch 1: 100%|██████████| 205/205 [02:31<00:00, 1.35it/s, train\_loss=0.492, val\_loss=0.490]Metric val\_loss improved by 0.124 >= min\_delta = 0.01. New best score: 0.490

Epoch 2: 100%|██████████| 205/205 [02:33<00:00, 1.34it/s, train\_loss=0.383, val\_loss=0.416]📉 Epoch 3: train\_loss=0.3828 | val\_loss=0.4158

Epoch 2: 100%|██████████| 205/205 [02:33<00:00, 1.34it/s, train\_loss=0.383, val\_loss=0.416]Metric val\_loss improved by 0.075 >= min\_delta = 0.01. New best score: 0.416

Epoch 3: 100%|██████████| 205/205 [02:33<00:00, 1.34it/s, train\_loss=0.353, val\_loss=0.460]📉 Epoch 4: train\_loss=0.3530 | val\_loss=0.4599

Epoch 4: 100%|██████████| 205/205 [02:32<00:00, 1.34it/s, train\_loss=0.395, val\_loss=0.418]📉 Epoch 5: train\_loss=0.3953 | val\_loss=0.4177

Epoch 5: 100%|██████████| 205/205 [02:32<00:00, 1.34it/s, train\_loss=0.347, val\_loss=0.414]📉 Epoch 6: train\_loss=0.3468 | val\_loss=0.4144

Epoch 5: 100%|██████████| 205/205 [02:32<00:00, 1.34it/s, train\_loss=0.347, val\_loss=0.414]Monitored metric val\_loss did not improve in the last 3 records. Best score: 0.416. Signaling Trainer to stop.

Epoch 5: 100%|██████████| 205/205 [02:32<00:00, 1.34it/s, train\_loss=0.347, val\_loss=0.414]

`predict()` was called with `n > output\_chunk\_length`: using auto-regression to forecast the values after `output\_chunk\_length` points. The model will access `(n - output\_chunk\_length)` future values of your `past\_covariates` (relative to the first predicted time step). To hide this warning, set `show\_warnings=False`.

GPU available: True (cuda), used: True

TPU available: False, using: 0 TPU cores

HPU available: False, using: 0 HPUs

LOCAL\_RANK: 0 - CUDA\_VISIBLE\_DEVICES: [0]

Predicting DataLoader 0: 100%|██████████| 1/1 [00:12<00:00, 0.08it/s][I 2025-05-10 18:47:08,168] Trial 12 finished with value: 314.4595642089844 and parameters: {'input\_chunk\_length': 72, 'output\_chunk\_length': 48, 'hidden\_size': 80, 'lstm\_layers': 3, 'num\_attention\_heads': 4, 'dropout': 0.3, 'n\_epochs': 30, 'lr': 0.001, 'batch\_size': 128}. Best is trial 0 with value: 284.9240417480469.

🔍 DEBUG: Dataövergripande information:

▶️ train: 2021-01-02 00:00:00 → 2023-12-31 23:00:00 (26256 timsteg)

✅ val: 2024-01-01 00:00:00 → 2024-12-31 23:00:00 (8784 timsteg)

⏪ past\_cov\_train: 2021-01-02 00:00:00 → 2023-12-31 23:00:00

⏪ past\_cov\_val: 2023-12-28 00:00:00 → 2024-12-31 23:00:00

⏩ future\_cov\_train: 2021-01-02 00:00:00 → 2023-12-31 23:00:00

⏩ future\_cov\_val: 2023-12-28 00:00:00 → 2024-12-31 23:00:00

GPU available: True (cuda), used: True

TPU available: False, using: 0 TPU cores

HPU available: False, using: 0 HPUs

LOCAL\_RANK: 0 - CUDA\_VISIBLE\_DEVICES: [0]

| Name | Type | Params | Mode

------------------------------------------------------------------------------------------------

0 | train\_metrics | MetricCollection | 0 | train

1 | val\_metrics | MetricCollection | 0 | train

2 | input\_embeddings | \_MultiEmbedding | 0 | train

3 | static\_covariates\_vsn | \_VariableSelectionNetwork | 0 | train

4 | encoder\_vsn | \_VariableSelectionNetwork | 201 K | train

5 | decoder\_vsn | \_VariableSelectionNetwork | 101 K | train

6 | static\_context\_grn | \_GatedResidualNetwork | 6.6 K | train

7 | static\_context\_hidden\_encoder\_grn | \_GatedResidualNetwork | 6.6 K | train

8 | static\_context\_cell\_encoder\_grn | \_GatedResidualNetwork | 6.6 K | train

9 | static\_context\_enrichment | \_GatedResidualNetwork | 6.6 K | train

10 | lstm\_encoder | LSTM | 39.4 K | train

11 | lstm\_decoder | LSTM | 39.4 K | train

12 | post\_lstm\_gan | \_GateAddNorm | 3.4 K | train

13 | static\_enrichment\_grn | \_GatedResidualNetwork | 8.2 K | train

14 | multihead\_attn | \_InterpretableMultiHeadAttention | 4.1 K | train

15 | post\_attn\_gan | \_GateAddNorm | 3.4 K | train

16 | feed\_forward\_block | \_GatedResidualNetwork | 6.6 K | train

17 | pre\_output\_gan | \_GateAddNorm | 3.4 K | train

18 | output\_layer | Linear | 123 | train

------------------------------------------------------------------------------------------------

435 K Trainable params

0 Non-trainable params

435 K Total params

1.743 Total estimated model params size (MB)

3062 Modules in train mode

0 Modules in eval mode

🚀 Startar ny gridsearch-körning med hyperparametrar:

input\_chunk\_length: 48

output\_chunk\_length: 48

hidden\_size: 40

lstm\_layers: 3

num\_attention\_heads: 4

dropout: 0.3

optimizer\_kwargs: {'lr': 0.001}

Epoch 0: 100%|██████████| 205/205 [01:50<00:00, 1.86it/s, train\_loss=0.668, val\_loss=0.741]📉 Epoch 1: train\_loss=0.6677 | val\_loss=0.7406

Epoch 0: 100%|██████████| 205/205 [01:50<00:00, 1.86it/s, train\_loss=0.668, val\_loss=0.741]Metric val\_loss improved. New best score: 0.741

Epoch 1: 100%|██████████| 205/205 [01:49<00:00, 1.87it/s, train\_loss=0.408, val\_loss=0.525]📉 Epoch 2: train\_loss=0.4083 | val\_loss=0.5250

Epoch 1: 100%|██████████| 205/205 [01:49<00:00, 1.87it/s, train\_loss=0.408, val\_loss=0.525]Metric val\_loss improved by 0.216 >= min\_delta = 0.01. New best score: 0.525

Epoch 2: 100%|██████████| 205/205 [01:49<00:00, 1.88it/s, train\_loss=0.381, val\_loss=0.521]📉 Epoch 3: train\_loss=0.3806 | val\_loss=0.5209

Epoch 3: 100%|██████████| 205/205 [01:49<00:00, 1.88it/s, train\_loss=0.395, val\_loss=0.469]📉 Epoch 4: train\_loss=0.3954 | val\_loss=0.4691

Epoch 3: 100%|██████████| 205/205 [01:49<00:00, 1.88it/s, train\_loss=0.395, val\_loss=0.469]Metric val\_loss improved by 0.056 >= min\_delta = 0.01. New best score: 0.469

Epoch 4: 100%|██████████| 205/205 [01:49<00:00, 1.87it/s, train\_loss=0.365, val\_loss=0.434]📉 Epoch 5: train\_loss=0.3649 | val\_loss=0.4339

Epoch 4: 100%|██████████| 205/205 [01:49<00:00, 1.87it/s, train\_loss=0.365, val\_loss=0.434]Metric val\_loss improved by 0.035 >= min\_delta = 0.01. New best score: 0.434

Epoch 5: 100%|██████████| 205/205 [01:49<00:00, 1.88it/s, train\_loss=0.335, val\_loss=0.492]📉 Epoch 6: train\_loss=0.3347 | val\_loss=0.4922

Epoch 6: 100%|██████████| 205/205 [01:49<00:00, 1.88it/s, train\_loss=0.392, val\_loss=0.468]📉 Epoch 7: train\_loss=0.3916 | val\_loss=0.4680

Epoch 7: 100%|██████████| 205/205 [01:49<00:00, 1.88it/s, train\_loss=0.309, val\_loss=0.393]📉 Epoch 8: train\_loss=0.3092 | val\_loss=0.3934

Epoch 7: 100%|██████████| 205/205 [01:49<00:00, 1.88it/s, train\_loss=0.309, val\_loss=0.393]Metric val\_loss improved by 0.041 >= min\_delta = 0.01. New best score: 0.393

Epoch 8: 100%|██████████| 205/205 [01:49<00:00, 1.88it/s, train\_loss=0.284, val\_loss=0.397]📉 Epoch 9: train\_loss=0.2844 | val\_loss=0.3970

Epoch 9: 100%|██████████| 205/205 [01:49<00:00, 1.87it/s, train\_loss=0.291, val\_loss=0.443]📉 Epoch 10: train\_loss=0.2911 | val\_loss=0.4425

Epoch 10: 100%|██████████| 205/205 [01:48<00:00, 1.88it/s, train\_loss=0.315, val\_loss=0.430]📉 Epoch 11: train\_loss=0.3149 | val\_loss=0.4299

Epoch 10: 100%|██████████| 205/205 [01:48<00:00, 1.88it/s, train\_loss=0.315, val\_loss=0.430]Monitored metric val\_loss did not improve in the last 3 records. Best score: 0.393. Signaling Trainer to stop.

Epoch 10: 100%|██████████| 205/205 [01:48<00:00, 1.88it/s, train\_loss=0.315, val\_loss=0.430]

`predict()` was called with `n > output\_chunk\_length`: using auto-regression to forecast the values after `output\_chunk\_length` points. The model will access `(n - output\_chunk\_length)` future values of your `past\_covariates` (relative to the first predicted time step). To hide this warning, set `show\_warnings=False`.

GPU available: True (cuda), used: True

TPU available: False, using: 0 TPU cores

HPU available: False, using: 0 HPUs

LOCAL\_RANK: 0 - CUDA\_VISIBLE\_DEVICES: [0]

Predicting DataLoader 0: 100%|██████████| 1/1 [00:12<00:00, 0.08it/s][I 2025-05-10 19:07:25,662] Trial 13 finished with value: 318.0419616699219 and parameters: {'input\_chunk\_length': 48, 'output\_chunk\_length': 48, 'hidden\_size': 40, 'lstm\_layers': 3, 'num\_attention\_heads': 4, 'dropout': 0.3, 'n\_epochs': 30, 'lr': 0.001, 'batch\_size': 128}. Best is trial 0 with value: 284.9240417480469.

🔍 DEBUG: Dataövergripande information:

▶️ train: 2021-01-02 00:00:00 → 2023-12-31 23:00:00 (26256 timsteg)

✅ val: 2024-01-01 00:00:00 → 2024-12-31 23:00:00 (8784 timsteg)

⏪ past\_cov\_train: 2021-01-02 00:00:00 → 2023-12-31 23:00:00

⏪ past\_cov\_val: 2023-12-28 00:00:00 → 2024-12-31 23:00:00

⏩ future\_cov\_train: 2021-01-02 00:00:00 → 2023-12-31 23:00:00

⏩ future\_cov\_val: 2023-12-28 00:00:00 → 2024-12-31 23:00:00

GPU available: True (cuda), used: True

TPU available: False, using: 0 TPU cores

HPU available: False, using: 0 HPUs

LOCAL\_RANK: 0 - CUDA\_VISIBLE\_DEVICES: [0]

| Name | Type | Params | Mode

------------------------------------------------------------------------------------------------

0 | train\_metrics | MetricCollection | 0 | train

1 | val\_metrics | MetricCollection | 0 | train

2 | input\_embeddings | \_MultiEmbedding | 0 | train

3 | static\_covariates\_vsn | \_VariableSelectionNetwork | 0 | train

4 | encoder\_vsn | \_VariableSelectionNetwork | 201 K | train

5 | decoder\_vsn | \_VariableSelectionNetwork | 101 K | train

6 | static\_context\_grn | \_GatedResidualNetwork | 6.6 K | train

7 | static\_context\_hidden\_encoder\_grn | \_GatedResidualNetwork | 6.6 K | train

8 | static\_context\_cell\_encoder\_grn | \_GatedResidualNetwork | 6.6 K | train

9 | static\_context\_enrichment | \_GatedResidualNetwork | 6.6 K | train

10 | lstm\_encoder | LSTM | 13.1 K | train

11 | lstm\_decoder | LSTM | 13.1 K | train

12 | post\_lstm\_gan | \_GateAddNorm | 3.4 K | train

13 | static\_enrichment\_grn | \_GatedResidualNetwork | 8.2 K | train

14 | multihead\_attn | \_InterpretableMultiHeadAttention | 6.5 K | train

15 | post\_attn\_gan | \_GateAddNorm | 3.4 K | train

16 | feed\_forward\_block | \_GatedResidualNetwork | 6.6 K | train

17 | pre\_output\_gan | \_GateAddNorm | 3.4 K | train

18 | output\_layer | Linear | 123 | train

------------------------------------------------------------------------------------------------

385 K Trainable params

0 Non-trainable params

385 K Total params

1.543 Total estimated model params size (MB)

3056 Modules in train mode

0 Modules in eval mode

🚀 Startar ny gridsearch-körning med hyperparametrar:

input\_chunk\_length: 72

output\_chunk\_length: 48

hidden\_size: 40

lstm\_layers: 1

num\_attention\_heads: 1

dropout: 0.3

optimizer\_kwargs: {'lr': 0.001}

Epoch 0: 100%|██████████| 205/205 [02:23<00:00, 1.43it/s, train\_loss=0.803, val\_loss=0.794]📉 Epoch 1: train\_loss=0.8027 | val\_loss=0.7944

Epoch 0: 100%|██████████| 205/205 [02:23<00:00, 1.43it/s, train\_loss=0.803, val\_loss=0.794]Metric val\_loss improved. New best score: 0.794

Epoch 1: 100%|██████████| 205/205 [02:23<00:00, 1.43it/s, train\_loss=0.437, val\_loss=0.523]📉 Epoch 2: train\_loss=0.4365 | val\_loss=0.5226

Epoch 1: 100%|██████████| 205/205 [02:23<00:00, 1.43it/s, train\_loss=0.437, val\_loss=0.523]Metric val\_loss improved by 0.272 >= min\_delta = 0.01. New best score: 0.523

Epoch 2: 100%|██████████| 205/205 [02:23<00:00, 1.43it/s, train\_loss=0.426, val\_loss=0.477]📉 Epoch 3: train\_loss=0.4259 | val\_loss=0.4767

Epoch 2: 100%|██████████| 205/205 [02:23<00:00, 1.43it/s, train\_loss=0.426, val\_loss=0.477]Metric val\_loss improved by 0.046 >= min\_delta = 0.01. New best score: 0.477

Epoch 3: 100%|██████████| 205/205 [02:23<00:00, 1.43it/s, train\_loss=0.350, val\_loss=0.521]📉 Epoch 4: train\_loss=0.3497 | val\_loss=0.5210

Epoch 4: 100%|██████████| 205/205 [02:23<00:00, 1.43it/s, train\_loss=0.374, val\_loss=0.518]📉 Epoch 5: train\_loss=0.3743 | val\_loss=0.5177

Epoch 5: 100%|██████████| 205/205 [02:23<00:00, 1.43it/s, train\_loss=0.339, val\_loss=0.434]📉 Epoch 6: train\_loss=0.3389 | val\_loss=0.4337

Epoch 5: 100%|██████████| 205/205 [02:23<00:00, 1.43it/s, train\_loss=0.339, val\_loss=0.434]Metric val\_loss improved by 0.043 >= min\_delta = 0.01. New best score: 0.434

Epoch 6: 100%|██████████| 205/205 [02:23<00:00, 1.43it/s, train\_loss=0.305, val\_loss=0.448]📉 Epoch 7: train\_loss=0.3047 | val\_loss=0.4482

Epoch 7: 100%|██████████| 205/205 [02:23<00:00, 1.43it/s, train\_loss=0.353, val\_loss=0.404]📉 Epoch 8: train\_loss=0.3531 | val\_loss=0.4040

Epoch 7: 100%|██████████| 205/205 [02:23<00:00, 1.43it/s, train\_loss=0.353, val\_loss=0.404]Metric val\_loss improved by 0.030 >= min\_delta = 0.01. New best score: 0.404

Epoch 8: 100%|██████████| 205/205 [02:23<00:00, 1.43it/s, train\_loss=0.304, val\_loss=0.391]📉 Epoch 9: train\_loss=0.3041 | val\_loss=0.3909

Epoch 8: 100%|██████████| 205/205 [02:23<00:00, 1.43it/s, train\_loss=0.304, val\_loss=0.391]Metric val\_loss improved by 0.013 >= min\_delta = 0.01. New best score: 0.391

Epoch 9: 100%|██████████| 205/205 [02:23<00:00, 1.43it/s, train\_loss=0.288, val\_loss=0.383]📉 Epoch 10: train\_loss=0.2880 | val\_loss=0.3827

Epoch 10: 100%|██████████| 205/205 [02:23<00:00, 1.43it/s, train\_loss=0.291, val\_loss=0.409]📉 Epoch 11: train\_loss=0.2910 | val\_loss=0.4090

Epoch 11: 100%|██████████| 205/205 [02:23<00:00, 1.43it/s, train\_loss=0.255, val\_loss=0.414]📉 Epoch 12: train\_loss=0.2554 | val\_loss=0.4139

Epoch 11: 100%|██████████| 205/205 [02:23<00:00, 1.43it/s, train\_loss=0.255, val\_loss=0.414]Monitored metric val\_loss did not improve in the last 3 records. Best score: 0.391. Signaling Trainer to stop.

Epoch 11: 100%|██████████| 205/205 [02:23<00:00, 1.43it/s, train\_loss=0.255, val\_loss=0.414]

`predict()` was called with `n > output\_chunk\_length`: using auto-regression to forecast the values after `output\_chunk\_length` points. The model will access `(n - output\_chunk\_length)` future values of your `past\_covariates` (relative to the first predicted time step). To hide this warning, set `show\_warnings=False`.

GPU available: True (cuda), used: True

TPU available: False, using: 0 TPU cores

HPU available: False, using: 0 HPUs

LOCAL\_RANK: 0 - CUDA\_VISIBLE\_DEVICES: [0]

Predicting DataLoader 0: 100%|██████████| 1/1 [00:11<00:00, 0.09it/s][I 2025-05-10 19:36:22,110] Trial 14 finished with value: 312.9854736328125 and parameters: {'input\_chunk\_length': 72, 'output\_chunk\_length': 48, 'hidden\_size': 40, 'lstm\_layers': 1, 'num\_attention\_heads': 1, 'dropout': 0.3, 'n\_epochs': 30, 'lr': 0.001, 'batch\_size': 128}. Best is trial 0 with value: 284.9240417480469.

🔍 DEBUG: Dataövergripande information:

▶️ train: 2021-01-02 00:00:00 → 2023-12-31 23:00:00 (26256 timsteg)

✅ val: 2024-01-01 00:00:00 → 2024-12-31 23:00:00 (8784 timsteg)

⏪ past\_cov\_train: 2021-01-02 00:00:00 → 2023-12-31 23:00:00

⏪ past\_cov\_val: 2023-12-28 00:00:00 → 2024-12-31 23:00:00

⏩ future\_cov\_train: 2021-01-02 00:00:00 → 2023-12-31 23:00:00

⏩ future\_cov\_val: 2023-12-28 00:00:00 → 2024-12-31 23:00:00

GPU available: True (cuda), used: True

TPU available: False, using: 0 TPU cores

HPU available: False, using: 0 HPUs

LOCAL\_RANK: 0 - CUDA\_VISIBLE\_DEVICES: [0]

| Name | Type | Params | Mode

------------------------------------------------------------------------------------------------

0 | train\_metrics | MetricCollection | 0 | train

1 | val\_metrics | MetricCollection | 0 | train

2 | input\_embeddings | \_MultiEmbedding | 0 | train

3 | static\_covariates\_vsn | \_VariableSelectionNetwork | 0 | train

4 | encoder\_vsn | \_VariableSelectionNetwork | 386 K | train

5 | decoder\_vsn | \_VariableSelectionNetwork | 185 K | train

6 | static\_context\_grn | \_GatedResidualNetwork | 26.1 K | train

7 | static\_context\_hidden\_encoder\_grn | \_GatedResidualNetwork | 26.1 K | train

8 | static\_context\_cell\_encoder\_grn | \_GatedResidualNetwork | 26.1 K | train

9 | static\_context\_enrichment | \_GatedResidualNetwork | 26.1 K | train

10 | lstm\_encoder | LSTM | 103 K | train

11 | lstm\_decoder | LSTM | 103 K | train

12 | post\_lstm\_gan | \_GateAddNorm | 13.1 K | train

13 | static\_enrichment\_grn | \_GatedResidualNetwork | 32.5 K | train

14 | multihead\_attn | \_InterpretableMultiHeadAttention | 16.2 K | train

15 | post\_attn\_gan | \_GateAddNorm | 13.1 K | train

16 | feed\_forward\_block | \_GatedResidualNetwork | 26.1 K | train

17 | pre\_output\_gan | \_GateAddNorm | 13.1 K | train

18 | output\_layer | Linear | 243 | train

------------------------------------------------------------------------------------------------

997 K Trainable params

0 Non-trainable params

997 K Total params

3.988 Total estimated model params size (MB)

3062 Modules in train mode

0 Modules in eval mode

🚀 Startar ny gridsearch-körning med hyperparametrar:

input\_chunk\_length: 72

output\_chunk\_length: 48

hidden\_size: 80

lstm\_layers: 2

num\_attention\_heads: 4

dropout: 0.4

optimizer\_kwargs: {'lr': 0.001}

Epoch 0: 100%|██████████| 205/205 [02:31<00:00, 1.35it/s, train\_loss=0.671, val\_loss=0.578]📉 Epoch 1: train\_loss=0.6712 | val\_loss=0.5777

Epoch 0: 100%|██████████| 205/205 [02:31<00:00, 1.35it/s, train\_loss=0.671, val\_loss=0.578]Metric val\_loss improved. New best score: 0.578

Epoch 1: 100%|██████████| 205/205 [02:31<00:00, 1.35it/s, train\_loss=0.414, val\_loss=0.592]📉 Epoch 2: train\_loss=0.4137 | val\_loss=0.5918

Epoch 2: 100%|██████████| 205/205 [02:31<00:00, 1.35it/s, train\_loss=0.463, val\_loss=0.566]📉 Epoch 3: train\_loss=0.4635 | val\_loss=0.5661

Epoch 2: 100%|██████████| 205/205 [02:31<00:00, 1.35it/s, train\_loss=0.463, val\_loss=0.566]Metric val\_loss improved by 0.012 >= min\_delta = 0.01. New best score: 0.566

Epoch 3: 100%|██████████| 205/205 [02:31<00:00, 1.35it/s, train\_loss=0.395, val\_loss=0.446]📉 Epoch 4: train\_loss=0.3946 | val\_loss=0.4457

Epoch 3: 100%|██████████| 205/205 [02:31<00:00, 1.35it/s, train\_loss=0.395, val\_loss=0.446]Metric val\_loss improved by 0.120 >= min\_delta = 0.01. New best score: 0.446

Epoch 4: 100%|██████████| 205/205 [02:31<00:00, 1.35it/s, train\_loss=0.367, val\_loss=0.397]📉 Epoch 5: train\_loss=0.3674 | val\_loss=0.3971

Epoch 4: 100%|██████████| 205/205 [02:31<00:00, 1.35it/s, train\_loss=0.367, val\_loss=0.397]Metric val\_loss improved by 0.049 >= min\_delta = 0.01. New best score: 0.397

Epoch 5: 100%|██████████| 205/205 [02:31<00:00, 1.35it/s, train\_loss=0.411, val\_loss=0.473]📉 Epoch 6: train\_loss=0.4105 | val\_loss=0.4731

Epoch 6: 100%|██████████| 205/205 [02:31<00:00, 1.35it/s, train\_loss=0.340, val\_loss=0.459]📉 Epoch 7: train\_loss=0.3402 | val\_loss=0.4590

Epoch 7: 100%|██████████| 205/205 [02:31<00:00, 1.35it/s, train\_loss=0.322, val\_loss=0.433]📉 Epoch 8: train\_loss=0.3217 | val\_loss=0.4331

Epoch 7: 100%|██████████| 205/205 [02:31<00:00, 1.35it/s, train\_loss=0.322, val\_loss=0.433]Monitored metric val\_loss did not improve in the last 3 records. Best score: 0.397. Signaling Trainer to stop.

Epoch 7: 100%|██████████| 205/205 [02:31<00:00, 1.35it/s, train\_loss=0.322, val\_loss=0.433]

`predict()` was called with `n > output\_chunk\_length`: using auto-regression to forecast the values after `output\_chunk\_length` points. The model will access `(n - output\_chunk\_length)` future values of your `past\_covariates` (relative to the first predicted time step). To hide this warning, set `show\_warnings=False`.

GPU available: True (cuda), used: True

TPU available: False, using: 0 TPU cores

HPU available: False, using: 0 HPUs

LOCAL\_RANK: 0 - CUDA\_VISIBLE\_DEVICES: [0]

Predicting DataLoader 0: 100%|██████████| 1/1 [00:12<00:00, 0.08it/s][I 2025-05-10 19:56:49,541] Trial 15 finished with value: 308.0473327636719 and parameters: {'input\_chunk\_length': 72, 'output\_chunk\_length': 48, 'hidden\_size': 80, 'lstm\_layers': 2, 'num\_attention\_heads': 4, 'dropout': 0.4, 'n\_epochs': 30, 'lr': 0.001, 'batch\_size': 128}. Best is trial 0 with value: 284.9240417480469.

🔍 DEBUG: Dataövergripande information:

▶️ train: 2021-01-02 00:00:00 → 2023-12-31 23:00:00 (26256 timsteg)

✅ val: 2024-01-01 00:00:00 → 2024-12-31 23:00:00 (8784 timsteg)

⏪ past\_cov\_train: 2021-01-02 00:00:00 → 2023-12-31 23:00:00

⏪ past\_cov\_val: 2023-12-28 00:00:00 → 2024-12-31 23:00:00

⏩ future\_cov\_train: 2021-01-02 00:00:00 → 2023-12-31 23:00:00

⏩ future\_cov\_val: 2023-12-28 00:00:00 → 2024-12-31 23:00:00

GPU available: True (cuda), used: True

TPU available: False, using: 0 TPU cores

HPU available: False, using: 0 HPUs

LOCAL\_RANK: 0 - CUDA\_VISIBLE\_DEVICES: [0]

| Name | Type | Params | Mode

------------------------------------------------------------------------------------------------

0 | train\_metrics | MetricCollection | 0 | train

1 | val\_metrics | MetricCollection | 0 | train

2 | input\_embeddings | \_MultiEmbedding | 0 | train

3 | static\_covariates\_vsn | \_VariableSelectionNetwork | 0 | train

4 | encoder\_vsn | \_VariableSelectionNetwork | 727 K | train

5 | decoder\_vsn | \_VariableSelectionNetwork | 312 K | train

6 | static\_context\_grn | \_GatedResidualNetwork | 103 K | train

7 | static\_context\_hidden\_encoder\_grn | \_GatedResidualNetwork | 103 K | train

8 | static\_context\_cell\_encoder\_grn | \_GatedResidualNetwork | 103 K | train

9 | static\_context\_enrichment | \_GatedResidualNetwork | 103 K | train

10 | lstm\_encoder | LSTM | 618 K | train

11 | lstm\_decoder | LSTM | 618 K | train

12 | post\_lstm\_gan | \_GateAddNorm | 51.8 K | train

13 | static\_enrichment\_grn | \_GatedResidualNetwork | 128 K | train

14 | multihead\_attn | \_InterpretableMultiHeadAttention | 102 K | train

15 | post\_attn\_gan | \_GateAddNorm | 51.8 K | train

16 | feed\_forward\_block | \_GatedResidualNetwork | 103 K | train

17 | pre\_output\_gan | \_GateAddNorm | 51.8 K | train

18 | output\_layer | Linear | 483 | train

------------------------------------------------------------------------------------------------

3.2 M Trainable params

0 Non-trainable params

3.2 M Total params

12.721 Total estimated model params size (MB)

3056 Modules in train mode

0 Modules in eval mode

🚀 Startar ny gridsearch-körning med hyperparametrar:

input\_chunk\_length: 72

output\_chunk\_length: 48

hidden\_size: 160

lstm\_layers: 3

num\_attention\_heads: 1

dropout: 0.3

optimizer\_kwargs: {'lr': 0.01}

Epoch 0: 100%|██████████| 205/205 [03:06<00:00, 1.10it/s, train\_loss=0.735, val\_loss=0.922]📉 Epoch 1: train\_loss=0.7347 | val\_loss=0.9223

Epoch 0: 100%|██████████| 205/205 [03:06<00:00, 1.10it/s, train\_loss=0.735, val\_loss=0.922]Metric val\_loss improved. New best score: 0.922

Epoch 1: 100%|██████████| 205/205 [03:06<00:00, 1.10it/s, train\_loss=0.858, val\_loss=0.764]📉 Epoch 2: train\_loss=0.8583 | val\_loss=0.7644

Epoch 1: 100%|██████████| 205/205 [03:06<00:00, 1.10it/s, train\_loss=0.858, val\_loss=0.764]Metric val\_loss improved by 0.158 >= min\_delta = 0.01. New best score: 0.764

Epoch 2: 100%|██████████| 205/205 [03:06<00:00, 1.10it/s, train\_loss=0.865, val\_loss=0.678]📉 Epoch 3: train\_loss=0.8653 | val\_loss=0.6782

Epoch 2: 100%|██████████| 205/205 [03:07<00:00, 1.10it/s, train\_loss=0.865, val\_loss=0.678]Metric val\_loss improved by 0.086 >= min\_delta = 0.01. New best score: 0.678

Epoch 3: 100%|██████████| 205/205 [03:06<00:00, 1.10it/s, train\_loss=0.551, val\_loss=0.504]📉 Epoch 4: train\_loss=0.5510 | val\_loss=0.5043

Epoch 3: 100%|██████████| 205/205 [03:06<00:00, 1.10it/s, train\_loss=0.551, val\_loss=0.504]Metric val\_loss improved by 0.174 >= min\_delta = 0.01. New best score: 0.504

Epoch 4: 100%|██████████| 205/205 [03:06<00:00, 1.10it/s, train\_loss=0.421, val\_loss=0.444]📉 Epoch 5: train\_loss=0.4215 | val\_loss=0.4444

Epoch 4: 100%|██████████| 205/205 [03:06<00:00, 1.10it/s, train\_loss=0.421, val\_loss=0.444]Metric val\_loss improved by 0.060 >= min\_delta = 0.01. New best score: 0.444

Epoch 5: 100%|██████████| 205/205 [03:06<00:00, 1.10it/s, train\_loss=0.367, val\_loss=0.491]📉 Epoch 6: train\_loss=0.3671 | val\_loss=0.4911

Epoch 6: 100%|██████████| 205/205 [03:07<00:00, 1.10it/s, train\_loss=0.408, val\_loss=0.558]📉 Epoch 7: train\_loss=0.4084 | val\_loss=0.5577

Epoch 7: 100%|██████████| 205/205 [03:06<00:00, 1.10it/s, train\_loss=0.400, val\_loss=0.463]📉 Epoch 8: train\_loss=0.4002 | val\_loss=0.4630

Epoch 7: 100%|██████████| 205/205 [03:06<00:00, 1.10it/s, train\_loss=0.400, val\_loss=0.463]Monitored metric val\_loss did not improve in the last 3 records. Best score: 0.444. Signaling Trainer to stop.

Epoch 7: 100%|██████████| 205/205 [03:06<00:00, 1.10it/s, train\_loss=0.400, val\_loss=0.463]

`predict()` was called with `n > output\_chunk\_length`: using auto-regression to forecast the values after `output\_chunk\_length` points. The model will access `(n - output\_chunk\_length)` future values of your `past\_covariates` (relative to the first predicted time step). To hide this warning, set `show\_warnings=False`.

GPU available: True (cuda), used: True

TPU available: False, using: 0 TPU cores

HPU available: False, using: 0 HPUs

LOCAL\_RANK: 0 - CUDA\_VISIBLE\_DEVICES: [0]

Predicting DataLoader 0: 100%|██████████| 1/1 [00:11<00:00, 0.08it/s][I 2025-05-10 20:21:58,031] Trial 16 finished with value: 342.36932373046875 and parameters: {'input\_chunk\_length': 72, 'output\_chunk\_length': 48, 'hidden\_size': 160, 'lstm\_layers': 3, 'num\_attention\_heads': 1, 'dropout': 0.3, 'n\_epochs': 30, 'lr': 0.01, 'batch\_size': 128}. Best is trial 0 with value: 284.9240417480469.

🔍 DEBUG: Dataövergripande information:

▶️ train: 2021-01-02 00:00:00 → 2023-12-31 23:00:00 (26256 timsteg)

✅ val: 2024-01-01 00:00:00 → 2024-12-31 23:00:00 (8784 timsteg)

⏪ past\_cov\_train: 2021-01-02 00:00:00 → 2023-12-31 23:00:00

⏪ past\_cov\_val: 2023-12-28 00:00:00 → 2024-12-31 23:00:00

⏩ future\_cov\_train: 2021-01-02 00:00:00 → 2023-12-31 23:00:00

⏩ future\_cov\_val: 2023-12-28 00:00:00 → 2024-12-31 23:00:00

GPU available: True (cuda), used: True

TPU available: False, using: 0 TPU cores

HPU available: False, using: 0 HPUs

LOCAL\_RANK: 0 - CUDA\_VISIBLE\_DEVICES: [0]

| Name | Type | Params | Mode

------------------------------------------------------------------------------------------------

0 | train\_metrics | MetricCollection | 0 | train

1 | val\_metrics | MetricCollection | 0 | train

2 | input\_embeddings | \_MultiEmbedding | 0 | train

3 | static\_covariates\_vsn | \_VariableSelectionNetwork | 0 | train

4 | encoder\_vsn | \_VariableSelectionNetwork | 1.2 M | train

5 | decoder\_vsn | \_VariableSelectionNetwork | 565 K | train

6 | static\_context\_grn | \_GatedResidualNetwork | 411 K | train

7 | static\_context\_hidden\_encoder\_grn | \_GatedResidualNetwork | 411 K | train

8 | static\_context\_cell\_encoder\_grn | \_GatedResidualNetwork | 411 K | train

9 | static\_context\_enrichment | \_GatedResidualNetwork | 411 K | train

10 | lstm\_encoder | LSTM | 2.5 M | train

11 | lstm\_decoder | LSTM | 2.5 M | train

12 | post\_lstm\_gan | \_GateAddNorm | 206 K | train

13 | static\_enrichment\_grn | \_GatedResidualNetwork | 513 K | train

14 | multihead\_attn | \_InterpretableMultiHeadAttention | 256 K | train

15 | post\_attn\_gan | \_GateAddNorm | 206 K | train

16 | feed\_forward\_block | \_GatedResidualNetwork | 411 K | train

17 | pre\_output\_gan | \_GateAddNorm | 206 K | train

18 | output\_layer | Linear | 963 | train

------------------------------------------------------------------------------------------------

10.2 M Trainable params

0 Non-trainable params

10.2 M Total params

40.726 Total estimated model params size (MB)

3062 Modules in train mode

0 Modules in eval mode

🚀 Startar ny gridsearch-körning med hyperparametrar:

input\_chunk\_length: 48

output\_chunk\_length: 48

hidden\_size: 320

lstm\_layers: 3

num\_attention\_heads: 4

dropout: 0.4

optimizer\_kwargs: {'lr': 0.001}

Epoch 0: 100%|██████████| 205/205 [03:11<00:00, 1.07it/s, train\_loss=0.616, val\_loss=0.686]📉 Epoch 1: train\_loss=0.6164 | val\_loss=0.6861

Epoch 0: 100%|██████████| 205/205 [03:11<00:00, 1.07it/s, train\_loss=0.616, val\_loss=0.686]Metric val\_loss improved. New best score: 0.686

Epoch 1: 100%|██████████| 205/205 [03:11<00:00, 1.07it/s, train\_loss=0.565, val\_loss=0.571]📉 Epoch 2: train\_loss=0.5651 | val\_loss=0.5712

Epoch 1: 100%|██████████| 205/205 [03:11<00:00, 1.07it/s, train\_loss=0.565, val\_loss=0.571]Metric val\_loss improved by 0.115 >= min\_delta = 0.01. New best score: 0.571

Epoch 2: 100%|██████████| 205/205 [03:11<00:00, 1.07it/s, train\_loss=0.443, val\_loss=0.474]📉 Epoch 3: train\_loss=0.4426 | val\_loss=0.4738

Epoch 2: 100%|██████████| 205/205 [03:11<00:00, 1.07it/s, train\_loss=0.443, val\_loss=0.474]Metric val\_loss improved by 0.097 >= min\_delta = 0.01. New best score: 0.474

Epoch 3: 100%|██████████| 205/205 [03:11<00:00, 1.07it/s, train\_loss=0.362, val\_loss=0.480]📉 Epoch 4: train\_loss=0.3618 | val\_loss=0.4796

Epoch 4: 100%|██████████| 205/205 [03:11<00:00, 1.07it/s, train\_loss=0.316, val\_loss=0.533]📉 Epoch 5: train\_loss=0.3165 | val\_loss=0.5327

Epoch 5: 100%|██████████| 205/205 [03:11<00:00, 1.07it/s, train\_loss=0.334, val\_loss=0.559]📉 Epoch 6: train\_loss=0.3337 | val\_loss=0.5590

Epoch 5: 100%|██████████| 205/205 [03:11<00:00, 1.07it/s, train\_loss=0.334, val\_loss=0.559]Monitored metric val\_loss did not improve in the last 3 records. Best score: 0.474. Signaling Trainer to stop.

Epoch 5: 100%|██████████| 205/205 [03:11<00:00, 1.07it/s, train\_loss=0.334, val\_loss=0.559]

`predict()` was called with `n > output\_chunk\_length`: using auto-regression to forecast the values after `output\_chunk\_length` points. The model will access `(n - output\_chunk\_length)` future values of your `past\_covariates` (relative to the first predicted time step). To hide this warning, set `show\_warnings=False`.

GPU available: True (cuda), used: True

TPU available: False, using: 0 TPU cores

HPU available: False, using: 0 HPUs

LOCAL\_RANK: 0 - CUDA\_VISIBLE\_DEVICES: [0]

Predicting DataLoader 0: 100%|██████████| 1/1 [00:12<00:00, 0.08it/s][I 2025-05-10 20:41:19,858] Trial 17 finished with value: 353.5141296386719 and parameters: {'input\_chunk\_length': 48, 'output\_chunk\_length': 48, 'hidden\_size': 320, 'lstm\_layers': 3, 'num\_attention\_heads': 4, 'dropout': 0.4, 'n\_epochs': 30, 'lr': 0.001, 'batch\_size': 128}. Best is trial 0 with value: 284.9240417480469.

🔍 DEBUG: Dataövergripande information:

▶️ train: 2021-01-02 00:00:00 → 2023-12-31 23:00:00 (26256 timsteg)

✅ val: 2024-01-01 00:00:00 → 2024-12-31 23:00:00 (8784 timsteg)

⏪ past\_cov\_train: 2021-01-02 00:00:00 → 2023-12-31 23:00:00

⏪ past\_cov\_val: 2023-12-28 00:00:00 → 2024-12-31 23:00:00

⏩ future\_cov\_train: 2021-01-02 00:00:00 → 2023-12-31 23:00:00

⏩ future\_cov\_val: 2023-12-28 00:00:00 → 2024-12-31 23:00:00

GPU available: True (cuda), used: True

TPU available: False, using: 0 TPU cores

HPU available: False, using: 0 HPUs

LOCAL\_RANK: 0 - CUDA\_VISIBLE\_DEVICES: [0]

| Name | Type | Params | Mode

------------------------------------------------------------------------------------------------

0 | train\_metrics | MetricCollection | 0 | train

1 | val\_metrics | MetricCollection | 0 | train

2 | input\_embeddings | \_MultiEmbedding | 0 | train

3 | static\_covariates\_vsn | \_VariableSelectionNetwork | 0 | train

4 | encoder\_vsn | \_VariableSelectionNetwork | 66.3 K | train

5 | decoder\_vsn | \_VariableSelectionNetwork | 33.0 K | train

6 | static\_context\_grn | \_GatedResidualNetwork | 460 | train

7 | static\_context\_hidden\_encoder\_grn | \_GatedResidualNetwork | 460 | train

8 | static\_context\_cell\_encoder\_grn | \_GatedResidualNetwork | 460 | train

9 | static\_context\_enrichment | \_GatedResidualNetwork | 460 | train

10 | lstm\_encoder | LSTM | 880 | train

11 | lstm\_decoder | LSTM | 880 | train

12 | post\_lstm\_gan | \_GateAddNorm | 240 | train

13 | static\_enrichment\_grn | \_GatedResidualNetwork | 560 | train

14 | multihead\_attn | \_InterpretableMultiHeadAttention | 218 | train

15 | post\_attn\_gan | \_GateAddNorm | 240 | train

16 | feed\_forward\_block | \_GatedResidualNetwork | 460 | train

17 | pre\_output\_gan | \_GateAddNorm | 240 | train

18 | output\_layer | Linear | 33 | train

------------------------------------------------------------------------------------------------

103 K Trainable params

0 Non-trainable params

103 K Total params

0.416 Total estimated model params size (MB)

3062 Modules in train mode

0 Modules in eval mode

🚀 Startar ny gridsearch-körning med hyperparametrar:

input\_chunk\_length: 72

output\_chunk\_length: 48

hidden\_size: 10

lstm\_layers: 1

num\_attention\_heads: 4

dropout: 0.1

optimizer\_kwargs: {'lr': 0.001}

Epoch 0: 100%|██████████| 205/205 [02:25<00:00, 1.41it/s, train\_loss=2.410, val\_loss=2.670]📉 Epoch 1: train\_loss=2.4073 | val\_loss=2.6667

Epoch 0: 100%|██████████| 205/205 [02:25<00:00, 1.40it/s, train\_loss=2.410, val\_loss=2.670]Metric val\_loss improved. New best score: 2.667

Epoch 1: 100%|██████████| 205/205 [02:25<00:00, 1.41it/s, train\_loss=0.747, val\_loss=0.833]📉 Epoch 2: train\_loss=0.7474 | val\_loss=0.8328

Epoch 1: 100%|██████████| 205/205 [02:25<00:00, 1.41it/s, train\_loss=0.747, val\_loss=0.833]Metric val\_loss improved by 1.834 >= min\_delta = 0.01. New best score: 0.833

Epoch 2: 100%|██████████| 205/205 [02:25<00:00, 1.41it/s, train\_loss=0.449, val\_loss=0.532]📉 Epoch 3: train\_loss=0.4486 | val\_loss=0.5317

Epoch 2: 100%|██████████| 205/205 [02:25<00:00, 1.40it/s, train\_loss=0.449, val\_loss=0.532]Metric val\_loss improved by 0.301 >= min\_delta = 0.01. New best score: 0.532

Epoch 3: 100%|██████████| 205/205 [02:25<00:00, 1.41it/s, train\_loss=0.341, val\_loss=0.502]📉 Epoch 4: train\_loss=0.3412 | val\_loss=0.5019

Epoch 3: 100%|██████████| 205/205 [02:25<00:00, 1.41it/s, train\_loss=0.341, val\_loss=0.502]Metric val\_loss improved by 0.030 >= min\_delta = 0.01. New best score: 0.502

Epoch 4: 100%|██████████| 205/205 [02:26<00:00, 1.40it/s, train\_loss=0.385, val\_loss=0.473]📉 Epoch 5: train\_loss=0.3847 | val\_loss=0.4734

Epoch 4: 100%|██████████| 205/205 [02:26<00:00, 1.40it/s, train\_loss=0.385, val\_loss=0.473]Metric val\_loss improved by 0.029 >= min\_delta = 0.01. New best score: 0.473

Epoch 5: 100%|██████████| 205/205 [02:25<00:00, 1.41it/s, train\_loss=0.342, val\_loss=0.417]📉 Epoch 6: train\_loss=0.3417 | val\_loss=0.4170

Epoch 5: 100%|██████████| 205/205 [02:25<00:00, 1.41it/s, train\_loss=0.342, val\_loss=0.417]Metric val\_loss improved by 0.056 >= min\_delta = 0.01. New best score: 0.417

Epoch 6: 100%|██████████| 205/205 [02:25<00:00, 1.41it/s, train\_loss=0.405, val\_loss=0.428]📉 Epoch 7: train\_loss=0.4051 | val\_loss=0.4280

Epoch 7: 100%|██████████| 205/205 [02:25<00:00, 1.41it/s, train\_loss=0.332, val\_loss=0.435]📉 Epoch 8: train\_loss=0.3319 | val\_loss=0.4350

Epoch 8: 100%|██████████| 205/205 [02:25<00:00, 1.41it/s, train\_loss=0.324, val\_loss=0.446]📉 Epoch 9: train\_loss=0.3235 | val\_loss=0.4462

Epoch 8: 100%|██████████| 205/205 [02:25<00:00, 1.41it/s, train\_loss=0.324, val\_loss=0.446]Monitored metric val\_loss did not improve in the last 3 records. Best score: 0.417. Signaling Trainer to stop.

Epoch 8: 100%|██████████| 205/205 [02:25<00:00, 1.41it/s, train\_loss=0.324, val\_loss=0.446]

`predict()` was called with `n > output\_chunk\_length`: using auto-regression to forecast the values after `output\_chunk\_length` points. The model will access `(n - output\_chunk\_length)` future values of your `past\_covariates` (relative to the first predicted time step). To hide this warning, set `show\_warnings=False`.

GPU available: True (cuda), used: True

TPU available: False, using: 0 TPU cores

HPU available: False, using: 0 HPUs

LOCAL\_RANK: 0 - CUDA\_VISIBLE\_DEVICES: [0]

Predicting DataLoader 0: 100%|██████████| 1/1 [00:12<00:00, 0.08it/s][I 2025-05-10 21:03:25,895] Trial 18 finished with value: 343.9536437988281 and parameters: {'input\_chunk\_length': 72, 'output\_chunk\_length': 48, 'hidden\_size': 10, 'lstm\_layers': 1, 'num\_attention\_heads': 4, 'dropout': 0.1, 'n\_epochs': 30, 'lr': 0.001, 'batch\_size': 128}. Best is trial 0 with value: 284.9240417480469.

🔍 DEBUG: Dataövergripande information:

▶️ train: 2021-01-02 00:00:00 → 2023-12-31 23:00:00 (26256 timsteg)

✅ val: 2024-01-01 00:00:00 → 2024-12-31 23:00:00 (8784 timsteg)

⏪ past\_cov\_train: 2021-01-02 00:00:00 → 2023-12-31 23:00:00

⏪ past\_cov\_val: 2023-12-28 00:00:00 → 2024-12-31 23:00:00

⏩ future\_cov\_train: 2021-01-02 00:00:00 → 2023-12-31 23:00:00

⏩ future\_cov\_val: 2023-12-28 00:00:00 → 2024-12-31 23:00:00

GPU available: True (cuda), used: True

TPU available: False, using: 0 TPU cores

HPU available: False, using: 0 HPUs

LOCAL\_RANK: 0 - CUDA\_VISIBLE\_DEVICES: [0]

| Name | Type | Params | Mode

------------------------------------------------------------------------------------------------

0 | train\_metrics | MetricCollection | 0 | train

1 | val\_metrics | MetricCollection | 0 | train

2 | input\_embeddings | \_MultiEmbedding | 0 | train

3 | static\_covariates\_vsn | \_VariableSelectionNetwork | 0 | train

4 | encoder\_vsn | \_VariableSelectionNetwork | 201 K | train

5 | decoder\_vsn | \_VariableSelectionNetwork | 101 K | train

6 | static\_context\_grn | \_GatedResidualNetwork | 6.6 K | train

7 | static\_context\_hidden\_encoder\_grn | \_GatedResidualNetwork | 6.6 K | train

8 | static\_context\_cell\_encoder\_grn | \_GatedResidualNetwork | 6.6 K | train

9 | static\_context\_enrichment | \_GatedResidualNetwork | 6.6 K | train

10 | lstm\_encoder | LSTM | 26.2 K | train

11 | lstm\_decoder | LSTM | 26.2 K | train

12 | post\_lstm\_gan | \_GateAddNorm | 3.4 K | train

13 | static\_enrichment\_grn | \_GatedResidualNetwork | 8.2 K | train

14 | multihead\_attn | \_InterpretableMultiHeadAttention | 6.5 K | train

15 | post\_attn\_gan | \_GateAddNorm | 3.4 K | train

16 | feed\_forward\_block | \_GatedResidualNetwork | 6.6 K | train

17 | pre\_output\_gan | \_GateAddNorm | 3.4 K | train

18 | output\_layer | Linear | 123 | train

------------------------------------------------------------------------------------------------

412 K Trainable params

0 Non-trainable params

412 K Total params

1.648 Total estimated model params size (MB)

3056 Modules in train mode

0 Modules in eval mode

🚀 Startar ny gridsearch-körning med hyperparametrar:

input\_chunk\_length: 72

output\_chunk\_length: 48

hidden\_size: 40

lstm\_layers: 2

num\_attention\_heads: 1

dropout: 0.2

optimizer\_kwargs: {'lr': 0.01}

Epoch 0: 100%|██████████| 205/205 [02:23<00:00, 1.43it/s, train\_loss=0.627, val\_loss=0.687]📉 Epoch 1: train\_loss=0.6270 | val\_loss=0.6866

Epoch 0: 100%|██████████| 205/205 [02:23<00:00, 1.43it/s, train\_loss=0.627, val\_loss=0.687]Metric val\_loss improved. New best score: 0.687

Epoch 1: 100%|██████████| 205/205 [02:23<00:00, 1.43it/s, train\_loss=0.426, val\_loss=0.462]📉 Epoch 2: train\_loss=0.4258 | val\_loss=0.4621

Epoch 1: 100%|██████████| 205/205 [02:23<00:00, 1.43it/s, train\_loss=0.426, val\_loss=0.462]Metric val\_loss improved by 0.224 >= min\_delta = 0.01. New best score: 0.462

Epoch 2: 100%|██████████| 205/205 [02:23<00:00, 1.43it/s, train\_loss=0.361, val\_loss=0.447]📉 Epoch 3: train\_loss=0.3609 | val\_loss=0.4465

Epoch 2: 100%|██████████| 205/205 [02:23<00:00, 1.43it/s, train\_loss=0.361, val\_loss=0.447]Metric val\_loss improved by 0.016 >= min\_delta = 0.01. New best score: 0.447

Epoch 3: 100%|██████████| 205/205 [02:23<00:00, 1.43it/s, train\_loss=0.328, val\_loss=0.551]📉 Epoch 4: train\_loss=0.3284 | val\_loss=0.5508

Epoch 4: 100%|██████████| 205/205 [02:23<00:00, 1.43it/s, train\_loss=0.281, val\_loss=0.578]📉 Epoch 5: train\_loss=0.2809 | val\_loss=0.5780

Epoch 5: 100%|██████████| 205/205 [02:23<00:00, 1.43it/s, train\_loss=0.225, val\_loss=0.447]📉 Epoch 6: train\_loss=0.2252 | val\_loss=0.4466

Epoch 5: 100%|██████████| 205/205 [02:23<00:00, 1.43it/s, train\_loss=0.225, val\_loss=0.447]Monitored metric val\_loss did not improve in the last 3 records. Best score: 0.447. Signaling Trainer to stop.

Epoch 5: 100%|██████████| 205/205 [02:23<00:00, 1.43it/s, train\_loss=0.225, val\_loss=0.447]

`predict()` was called with `n > output\_chunk\_length`: using auto-regression to forecast the values after `output\_chunk\_length` points. The model will access `(n - output\_chunk\_length)` future values of your `past\_covariates` (relative to the first predicted time step). To hide this warning, set `show\_warnings=False`.

GPU available: True (cuda), used: True

TPU available: False, using: 0 TPU cores

HPU available: False, using: 0 HPUs

LOCAL\_RANK: 0 - CUDA\_VISIBLE\_DEVICES: [0]

Predicting DataLoader 0: 100%|██████████| 1/1 [00:11<00:00, 0.08it/s][I 2025-05-10 21:18:01,076] Trial 19 finished with value: 270.6866760253906 and parameters: {'input\_chunk\_length': 72, 'output\_chunk\_length': 48, 'hidden\_size': 40, 'lstm\_layers': 2, 'num\_attention\_heads': 1, 'dropout': 0.2, 'n\_epochs': 30, 'lr': 0.01, 'batch\_size': 128}. Best is trial 19 with value: 270.6866760253906.

✅ Saved new best params with score 270.6866760253906 to best\_params.json

🔍 DEBUG: Dataövergripande information:

▶️ train: 2021-01-02 00:00:00 → 2023-12-31 23:00:00 (26256 timsteg)

✅ val: 2024-01-01 00:00:00 → 2024-12-31 23:00:00 (8784 timsteg)

⏪ past\_cov\_train: 2021-01-02 00:00:00 → 2023-12-31 23:00:00

⏪ past\_cov\_val: 2023-12-28 00:00:00 → 2024-12-31 23:00:00

⏩ future\_cov\_train: 2021-01-02 00:00:00 → 2023-12-31 23:00:00

⏩ future\_cov\_val: 2023-12-28 00:00:00 → 2024-12-31 23:00:00

GPU available: True (cuda), used: True

TPU available: False, using: 0 TPU cores

HPU available: False, using: 0 HPUs

LOCAL\_RANK: 0 - CUDA\_VISIBLE\_DEVICES: [0]

| Name | Type | Params | Mode

------------------------------------------------------------------------------------------------

0 | train\_metrics | MetricCollection | 0 | train

1 | val\_metrics | MetricCollection | 0 | train

2 | input\_embeddings | \_MultiEmbedding | 0 | train

3 | static\_covariates\_vsn | \_VariableSelectionNetwork | 0 | train

4 | encoder\_vsn | \_VariableSelectionNetwork | 201 K | train

5 | decoder\_vsn | \_VariableSelectionNetwork | 101 K | train

6 | static\_context\_grn | \_GatedResidualNetwork | 6.6 K | train

7 | static\_context\_hidden\_encoder\_grn | \_GatedResidualNetwork | 6.6 K | train

8 | static\_context\_cell\_encoder\_grn | \_GatedResidualNetwork | 6.6 K | train

9 | static\_context\_enrichment | \_GatedResidualNetwork | 6.6 K | train

10 | lstm\_encoder | LSTM | 26.2 K | train

11 | lstm\_decoder | LSTM | 26.2 K | train

12 | post\_lstm\_gan | \_GateAddNorm | 3.4 K | train

13 | static\_enrichment\_grn | \_GatedResidualNetwork | 8.2 K | train

14 | multihead\_attn | \_InterpretableMultiHeadAttention | 4.1 K | train

15 | post\_attn\_gan | \_GateAddNorm | 3.4 K | train

16 | feed\_forward\_block | \_GatedResidualNetwork | 6.6 K | train

17 | pre\_output\_gan | \_GateAddNorm | 3.4 K | train

18 | output\_layer | Linear | 123 | train

------------------------------------------------------------------------------------------------

409 K Trainable params

0 Non-trainable params

409 K Total params

1.639 Total estimated model params size (MB)

3062 Modules in train mode

0 Modules in eval mode

🚀 Startar ny gridsearch-körning med hyperparametrar:

input\_chunk\_length: 48

output\_chunk\_length: 48

hidden\_size: 40

lstm\_layers: 2

num\_attention\_heads: 4

dropout: 0.2

optimizer\_kwargs: {'lr': 0.01}

Epoch 0: 100%|██████████| 205/205 [01:49<00:00, 1.88it/s, train\_loss=0.455, val\_loss=0.594]📉 Epoch 1: train\_loss=0.4547 | val\_loss=0.5941

Epoch 0: 100%|██████████| 205/205 [01:49<00:00, 1.88it/s, train\_loss=0.455, val\_loss=0.594]Metric val\_loss improved. New best score: 0.594

Epoch 1: 100%|██████████| 205/205 [01:48<00:00, 1.88it/s, train\_loss=0.376, val\_loss=0.478]📉 Epoch 2: train\_loss=0.3755 | val\_loss=0.4775

Epoch 1: 100%|██████████| 205/205 [01:48<00:00, 1.88it/s, train\_loss=0.376, val\_loss=0.478]Metric val\_loss improved by 0.117 >= min\_delta = 0.01. New best score: 0.478

Epoch 2: 100%|██████████| 205/205 [01:48<00:00, 1.88it/s, train\_loss=0.347, val\_loss=0.439]📉 Epoch 3: train\_loss=0.3475 | val\_loss=0.4394

Epoch 2: 100%|██████████| 205/205 [01:48<00:00, 1.88it/s, train\_loss=0.347, val\_loss=0.439]Metric val\_loss improved by 0.038 >= min\_delta = 0.01. New best score: 0.439

Epoch 3: 100%|██████████| 205/205 [01:49<00:00, 1.88it/s, train\_loss=0.288, val\_loss=0.455]📉 Epoch 4: train\_loss=0.2878 | val\_loss=0.4545

Epoch 4: 100%|██████████| 205/205 [01:48<00:00, 1.88it/s, train\_loss=0.262, val\_loss=0.607]📉 Epoch 5: train\_loss=0.2625 | val\_loss=0.6066

Epoch 5: 100%|██████████| 205/205 [01:49<00:00, 1.88it/s, train\_loss=0.240, val\_loss=0.546]📉 Epoch 6: train\_loss=0.2400 | val\_loss=0.5462

Epoch 5: 100%|██████████| 205/205 [01:49<00:00, 1.88it/s, train\_loss=0.240, val\_loss=0.546]Monitored metric val\_loss did not improve in the last 3 records. Best score: 0.439. Signaling Trainer to stop.

Epoch 5: 100%|██████████| 205/205 [01:49<00:00, 1.88it/s, train\_loss=0.240, val\_loss=0.546]

`predict()` was called with `n > output\_chunk\_length`: using auto-regression to forecast the values after `output\_chunk\_length` points. The model will access `(n - output\_chunk\_length)` future values of your `past\_covariates` (relative to the first predicted time step). To hide this warning, set `show\_warnings=False`.

GPU available: True (cuda), used: True

TPU available: False, using: 0 TPU cores

HPU available: False, using: 0 HPUs

LOCAL\_RANK: 0 - CUDA\_VISIBLE\_DEVICES: [0]

Predicting DataLoader 0: 100%|██████████| 1/1 [00:11<00:00, 0.08it/s][I 2025-05-10 21:29:08,577] Trial 20 finished with value: 321.9932861328125 and parameters: {'input\_chunk\_length': 48, 'output\_chunk\_length': 48, 'hidden\_size': 40, 'lstm\_layers': 2, 'num\_attention\_heads': 4, 'dropout': 0.2, 'n\_epochs': 30, 'lr': 0.01, 'batch\_size': 128}. Best is trial 19 with value: 270.6866760253906.

🔍 DEBUG: Dataövergripande information:

▶️ train: 2021-01-02 00:00:00 → 2023-12-31 23:00:00 (26256 timsteg)

✅ val: 2024-01-01 00:00:00 → 2024-12-31 23:00:00 (8784 timsteg)

⏪ past\_cov\_train: 2021-01-02 00:00:00 → 2023-12-31 23:00:00

⏪ past\_cov\_val: 2023-12-28 00:00:00 → 2024-12-31 23:00:00

⏩ future\_cov\_train: 2021-01-02 00:00:00 → 2023-12-31 23:00:00

⏩ future\_cov\_val: 2023-12-28 00:00:00 → 2024-12-31 23:00:00

GPU available: True (cuda), used: True

TPU available: False, using: 0 TPU cores

HPU available: False, using: 0 HPUs

LOCAL\_RANK: 0 - CUDA\_VISIBLE\_DEVICES: [0]

| Name | Type | Params | Mode

------------------------------------------------------------------------------------------------

0 | train\_metrics | MetricCollection | 0 | train

1 | val\_metrics | MetricCollection | 0 | train

2 | input\_embeddings | \_MultiEmbedding | 0 | train

3 | static\_covariates\_vsn | \_VariableSelectionNetwork | 0 | train

4 | encoder\_vsn | \_VariableSelectionNetwork | 201 K | train

5 | decoder\_vsn | \_VariableSelectionNetwork | 101 K | train

6 | static\_context\_grn | \_GatedResidualNetwork | 6.6 K | train

7 | static\_context\_hidden\_encoder\_grn | \_GatedResidualNetwork | 6.6 K | train

8 | static\_context\_cell\_encoder\_grn | \_GatedResidualNetwork | 6.6 K | train

9 | static\_context\_enrichment | \_GatedResidualNetwork | 6.6 K | train

10 | lstm\_encoder | LSTM | 26.2 K | train

11 | lstm\_decoder | LSTM | 26.2 K | train

12 | post\_lstm\_gan | \_GateAddNorm | 3.4 K | train

13 | static\_enrichment\_grn | \_GatedResidualNetwork | 8.2 K | train

14 | multihead\_attn | \_InterpretableMultiHeadAttention | 6.5 K | train

15 | post\_attn\_gan | \_GateAddNorm | 3.4 K | train

16 | feed\_forward\_block | \_GatedResidualNetwork | 6.6 K | train

17 | pre\_output\_gan | \_GateAddNorm | 3.4 K | train

18 | output\_layer | Linear | 123 | train

------------------------------------------------------------------------------------------------

412 K Trainable params

0 Non-trainable params

412 K Total params

1.648 Total estimated model params size (MB)

3056 Modules in train mode

0 Modules in eval mode

🚀 Startar ny gridsearch-körning med hyperparametrar:

input\_chunk\_length: 72

output\_chunk\_length: 48

hidden\_size: 40

lstm\_layers: 2

num\_attention\_heads: 1

dropout: 0.2

optimizer\_kwargs: {'lr': 0.01}

Epoch 0: 100%|██████████| 205/205 [02:23<00:00, 1.43it/s, train\_loss=0.604, val\_loss=0.593]📉 Epoch 1: train\_loss=0.6035 | val\_loss=0.5927

Epoch 0: 100%|██████████| 205/205 [02:23<00:00, 1.43it/s, train\_loss=0.604, val\_loss=0.593]Metric val\_loss improved. New best score: 0.593

Epoch 1: 100%|██████████| 205/205 [02:23<00:00, 1.43it/s, train\_loss=0.416, val\_loss=0.531]📉 Epoch 2: train\_loss=0.4159 | val\_loss=0.5306

Epoch 1: 100%|██████████| 205/205 [02:23<00:00, 1.43it/s, train\_loss=0.416, val\_loss=0.531]Metric val\_loss improved by 0.062 >= min\_delta = 0.01. New best score: 0.531

Epoch 2: 100%|██████████| 205/205 [02:23<00:00, 1.43it/s, train\_loss=0.358, val\_loss=0.436]📉 Epoch 3: train\_loss=0.3578 | val\_loss=0.4356

Epoch 2: 100%|██████████| 205/205 [02:23<00:00, 1.43it/s, train\_loss=0.358, val\_loss=0.436]Metric val\_loss improved by 0.095 >= min\_delta = 0.01. New best score: 0.436

Epoch 3: 100%|██████████| 205/205 [02:23<00:00, 1.43it/s, train\_loss=0.303, val\_loss=0.452]📉 Epoch 4: train\_loss=0.3026 | val\_loss=0.4518

Epoch 4: 100%|██████████| 205/205 [02:23<00:00, 1.43it/s, train\_loss=0.302, val\_loss=0.498]📉 Epoch 5: train\_loss=0.3017 | val\_loss=0.4984

Epoch 5: 100%|██████████| 205/205 [02:23<00:00, 1.43it/s, train\_loss=0.244, val\_loss=0.489]📉 Epoch 6: train\_loss=0.2441 | val\_loss=0.4888

Epoch 5: 100%|██████████| 205/205 [02:23<00:00, 1.43it/s, train\_loss=0.244, val\_loss=0.489]Monitored metric val\_loss did not improve in the last 3 records. Best score: 0.436. Signaling Trainer to stop.

Epoch 5: 100%|██████████| 205/205 [02:23<00:00, 1.43it/s, train\_loss=0.244, val\_loss=0.489]

`predict()` was called with `n > output\_chunk\_length`: using auto-regression to forecast the values after `output\_chunk\_length` points. The model will access `(n - output\_chunk\_length)` future values of your `past\_covariates` (relative to the first predicted time step). To hide this warning, set `show\_warnings=False`.

GPU available: True (cuda), used: True

TPU available: False, using: 0 TPU cores

HPU available: False, using: 0 HPUs

LOCAL\_RANK: 0 - CUDA\_VISIBLE\_DEVICES: [0]

Predicting DataLoader 0: 100%|██████████| 1/1 [00:11<00:00, 0.08it/s][I 2025-05-10 21:43:44,167] Trial 21 finished with value: 314.173095703125 and parameters: {'input\_chunk\_length': 72, 'output\_chunk\_length': 48, 'hidden\_size': 40, 'lstm\_layers': 2, 'num\_attention\_heads': 1, 'dropout': 0.2, 'n\_epochs': 30, 'lr': 0.01, 'batch\_size': 128}. Best is trial 19 with value: 270.6866760253906.

🔍 DEBUG: Dataövergripande information:

▶️ train: 2021-01-02 00:00:00 → 2023-12-31 23:00:00 (26256 timsteg)

✅ val: 2024-01-01 00:00:00 → 2024-12-31 23:00:00 (8784 timsteg)

⏪ past\_cov\_train: 2021-01-02 00:00:00 → 2023-12-31 23:00:00

⏪ past\_cov\_val: 2023-12-28 00:00:00 → 2024-12-31 23:00:00

⏩ future\_cov\_train: 2021-01-02 00:00:00 → 2023-12-31 23:00:00

⏩ future\_cov\_val: 2023-12-28 00:00:00 → 2024-12-31 23:00:00

GPU available: True (cuda), used: True

TPU available: False, using: 0 TPU cores

HPU available: False, using: 0 HPUs

LOCAL\_RANK: 0 - CUDA\_VISIBLE\_DEVICES: [0]

| Name | Type | Params | Mode

------------------------------------------------------------------------------------------------

0 | train\_metrics | MetricCollection | 0 | train

1 | val\_metrics | MetricCollection | 0 | train

2 | input\_embeddings | \_MultiEmbedding | 0 | train

3 | static\_covariates\_vsn | \_VariableSelectionNetwork | 0 | train

4 | encoder\_vsn | \_VariableSelectionNetwork | 386 K | train

5 | decoder\_vsn | \_VariableSelectionNetwork | 185 K | train

6 | static\_context\_grn | \_GatedResidualNetwork | 26.1 K | train

7 | static\_context\_hidden\_encoder\_grn | \_GatedResidualNetwork | 26.1 K | train

8 | static\_context\_cell\_encoder\_grn | \_GatedResidualNetwork | 26.1 K | train

9 | static\_context\_enrichment | \_GatedResidualNetwork | 26.1 K | train

10 | lstm\_encoder | LSTM | 103 K | train

11 | lstm\_decoder | LSTM | 103 K | train

12 | post\_lstm\_gan | \_GateAddNorm | 13.1 K | train

13 | static\_enrichment\_grn | \_GatedResidualNetwork | 32.5 K | train

14 | multihead\_attn | \_InterpretableMultiHeadAttention | 25.8 K | train

15 | post\_attn\_gan | \_GateAddNorm | 13.1 K | train

16 | feed\_forward\_block | \_GatedResidualNetwork | 26.1 K | train

17 | pre\_output\_gan | \_GateAddNorm | 13.1 K | train

18 | output\_layer | Linear | 243 | train

------------------------------------------------------------------------------------------------

1.0 M Trainable params

0 Non-trainable params

1.0 M Total params

4.027 Total estimated model params size (MB)

3056 Modules in train mode

0 Modules in eval mode

🚀 Startar ny gridsearch-körning med hyperparametrar:

input\_chunk\_length: 72

output\_chunk\_length: 48

hidden\_size: 80

lstm\_layers: 2

num\_attention\_heads: 1

dropout: 0.2

optimizer\_kwargs: {'lr': 0.01}

Epoch 0: 100%|██████████| 205/205 [02:32<00:00, 1.35it/s, train\_loss=0.606, val\_loss=0.870]📉 Epoch 1: train\_loss=0.6064 | val\_loss=0.8703

Epoch 0: 100%|██████████| 205/205 [02:32<00:00, 1.35it/s, train\_loss=0.606, val\_loss=0.870]Metric val\_loss improved. New best score: 0.870

Epoch 1: 100%|██████████| 205/205 [02:31<00:00, 1.35it/s, train\_loss=0.606, val\_loss=0.602]📉 Epoch 2: train\_loss=0.6057 | val\_loss=0.6021

Epoch 1: 100%|██████████| 205/205 [02:31<00:00, 1.35it/s, train\_loss=0.606, val\_loss=0.602]Metric val\_loss improved by 0.268 >= min\_delta = 0.01. New best score: 0.602

Epoch 2: 100%|██████████| 205/205 [02:31<00:00, 1.35it/s, train\_loss=0.597, val\_loss=0.609]📉 Epoch 3: train\_loss=0.5974 | val\_loss=0.6089

Epoch 3: 100%|██████████| 205/205 [02:31<00:00, 1.35it/s, train\_loss=0.602, val\_loss=0.602]📉 Epoch 4: train\_loss=0.6015 | val\_loss=0.6024

Epoch 4: 100%|██████████| 205/205 [02:31<00:00, 1.35it/s, train\_loss=0.597, val\_loss=0.567]📉 Epoch 5: train\_loss=0.5969 | val\_loss=0.5670

Epoch 4: 100%|██████████| 205/205 [02:31<00:00, 1.35it/s, train\_loss=0.597, val\_loss=0.567]Metric val\_loss improved by 0.035 >= min\_delta = 0.01. New best score: 0.567

Epoch 5: 100%|██████████| 205/205 [02:31<00:00, 1.35it/s, train\_loss=0.521, val\_loss=0.541]📉 Epoch 6: train\_loss=0.5214 | val\_loss=0.5411

Epoch 5: 100%|██████████| 205/205 [02:31<00:00, 1.35it/s, train\_loss=0.521, val\_loss=0.541]Metric val\_loss improved by 0.026 >= min\_delta = 0.01. New best score: 0.541

Epoch 6: 100%|██████████| 205/205 [02:31<00:00, 1.35it/s, train\_loss=0.396, val\_loss=0.419]📉 Epoch 7: train\_loss=0.3963 | val\_loss=0.4194

Epoch 6: 100%|██████████| 205/205 [02:31<00:00, 1.35it/s, train\_loss=0.396, val\_loss=0.419]Metric val\_loss improved by 0.122 >= min\_delta = 0.01. New best score: 0.419

Epoch 7: 100%|██████████| 205/205 [02:31<00:00, 1.35it/s, train\_loss=0.292, val\_loss=0.498]📉 Epoch 8: train\_loss=0.2921 | val\_loss=0.4979

Epoch 8: 100%|██████████| 205/205 [02:31<00:00, 1.35it/s, train\_loss=0.272, val\_loss=0.468]📉 Epoch 9: train\_loss=0.2717 | val\_loss=0.4685

Epoch 9: 100%|██████████| 205/205 [02:31<00:00, 1.35it/s, train\_loss=0.211, val\_loss=0.534]📉 Epoch 10: train\_loss=0.2105 | val\_loss=0.5337

Epoch 9: 100%|██████████| 205/205 [02:31<00:00, 1.35it/s, train\_loss=0.211, val\_loss=0.534]Monitored metric val\_loss did not improve in the last 3 records. Best score: 0.419. Signaling Trainer to stop.

Epoch 9: 100%|██████████| 205/205 [02:31<00:00, 1.35it/s, train\_loss=0.211, val\_loss=0.534]

`predict()` was called with `n > output\_chunk\_length`: using auto-regression to forecast the values after `output\_chunk\_length` points. The model will access `(n - output\_chunk\_length)` future values of your `past\_covariates` (relative to the first predicted time step). To hide this warning, set `show\_warnings=False`.

GPU available: True (cuda), used: True

TPU available: False, using: 0 TPU cores

HPU available: False, using: 0 HPUs

LOCAL\_RANK: 0 - CUDA\_VISIBLE\_DEVICES: [0]

Predicting DataLoader 0: 100%|██████████| 1/1 [00:11<00:00, 0.08it/s][I 2025-05-10 22:09:14,650] Trial 22 finished with value: 349.30328369140625 and parameters: {'input\_chunk\_length': 72, 'output\_chunk\_length': 48, 'hidden\_size': 80, 'lstm\_layers': 2, 'num\_attention\_heads': 1, 'dropout': 0.2, 'n\_epochs': 30, 'lr': 0.01, 'batch\_size': 128}. Best is trial 19 with value: 270.6866760253906.

🔍 DEBUG: Dataövergripande information:

▶️ train: 2021-01-02 00:00:00 → 2023-12-31 23:00:00 (26256 timsteg)

✅ val: 2024-01-01 00:00:00 → 2024-12-31 23:00:00 (8784 timsteg)

⏪ past\_cov\_train: 2021-01-02 00:00:00 → 2023-12-31 23:00:00

⏪ past\_cov\_val: 2023-12-28 00:00:00 → 2024-12-31 23:00:00

⏩ future\_cov\_train: 2021-01-02 00:00:00 → 2023-12-31 23:00:00

⏩ future\_cov\_val: 2023-12-28 00:00:00 → 2024-12-31 23:00:00

GPU available: True (cuda), used: True

TPU available: False, using: 0 TPU cores

HPU available: False, using: 0 HPUs

LOCAL\_RANK: 0 - CUDA\_VISIBLE\_DEVICES: [0]

| Name | Type | Params | Mode

------------------------------------------------------------------------------------------------

0 | train\_metrics | MetricCollection | 0 | train

1 | val\_metrics | MetricCollection | 0 | train

2 | input\_embeddings | \_MultiEmbedding | 0 | train

3 | static\_covariates\_vsn | \_VariableSelectionNetwork | 0 | train

4 | encoder\_vsn | \_VariableSelectionNetwork | 201 K | train

5 | decoder\_vsn | \_VariableSelectionNetwork | 101 K | train

6 | static\_context\_grn | \_GatedResidualNetwork | 6.6 K | train

7 | static\_context\_hidden\_encoder\_grn | \_GatedResidualNetwork | 6.6 K | train

8 | static\_context\_cell\_encoder\_grn | \_GatedResidualNetwork | 6.6 K | train

9 | static\_context\_enrichment | \_GatedResidualNetwork | 6.6 K | train

10 | lstm\_encoder | LSTM | 26.2 K | train

11 | lstm\_decoder | LSTM | 26.2 K | train

12 | post\_lstm\_gan | \_GateAddNorm | 3.4 K | train

13 | static\_enrichment\_grn | \_GatedResidualNetwork | 8.2 K | train

14 | multihead\_attn | \_InterpretableMultiHeadAttention | 6.5 K | train

15 | post\_attn\_gan | \_GateAddNorm | 3.4 K | train

16 | feed\_forward\_block | \_GatedResidualNetwork | 6.6 K | train

17 | pre\_output\_gan | \_GateAddNorm | 3.4 K | train

18 | output\_layer | Linear | 123 | train

------------------------------------------------------------------------------------------------

412 K Trainable params

0 Non-trainable params

412 K Total params

1.648 Total estimated model params size (MB)

3056 Modules in train mode

0 Modules in eval mode

🚀 Startar ny gridsearch-körning med hyperparametrar:

input\_chunk\_length: 72

output\_chunk\_length: 48

hidden\_size: 40

lstm\_layers: 2

num\_attention\_heads: 1

dropout: 0.5

optimizer\_kwargs: {'lr': 0.01}

Epoch 0: 100%|██████████| 205/205 [02:23<00:00, 1.43it/s, train\_loss=0.578, val\_loss=0.544]📉 Epoch 1: train\_loss=0.5778 | val\_loss=0.5441

Epoch 0: 100%|██████████| 205/205 [02:23<00:00, 1.43it/s, train\_loss=0.578, val\_loss=0.544]Metric val\_loss improved. New best score: 0.544

Epoch 1: 100%|██████████| 205/205 [02:23<00:00, 1.43it/s, train\_loss=0.413, val\_loss=0.532]📉 Epoch 2: train\_loss=0.4132 | val\_loss=0.5318

Epoch 1: 100%|██████████| 205/205 [02:23<00:00, 1.43it/s, train\_loss=0.413, val\_loss=0.532]Metric val\_loss improved by 0.012 >= min\_delta = 0.01. New best score: 0.532

Epoch 2: 100%|██████████| 205/205 [02:23<00:00, 1.43it/s, train\_loss=0.382, val\_loss=0.539]📉 Epoch 3: train\_loss=0.3821 | val\_loss=0.5392

Epoch 3: 100%|██████████| 205/205 [02:23<00:00, 1.42it/s, train\_loss=0.343, val\_loss=0.611]📉 Epoch 4: train\_loss=0.3429 | val\_loss=0.6110

Epoch 4: 100%|██████████| 205/205 [02:23<00:00, 1.43it/s, train\_loss=0.341, val\_loss=0.484]📉 Epoch 5: train\_loss=0.3407 | val\_loss=0.4839

Epoch 4: 100%|██████████| 205/205 [02:23<00:00, 1.43it/s, train\_loss=0.341, val\_loss=0.484]Metric val\_loss improved by 0.048 >= min\_delta = 0.01. New best score: 0.484

Epoch 5: 100%|██████████| 205/205 [02:23<00:00, 1.43it/s, train\_loss=0.273, val\_loss=0.472]📉 Epoch 6: train\_loss=0.2731 | val\_loss=0.4718

Epoch 5: 100%|██████████| 205/205 [02:23<00:00, 1.43it/s, train\_loss=0.273, val\_loss=0.472]Metric val\_loss improved by 0.012 >= min\_delta = 0.01. New best score: 0.472

Epoch 6: 100%|██████████| 205/205 [02:23<00:00, 1.43it/s, train\_loss=0.230, val\_loss=0.528]📉 Epoch 7: train\_loss=0.2296 | val\_loss=0.5279

Epoch 7: 100%|██████████| 205/205 [02:23<00:00, 1.43it/s, train\_loss=0.229, val\_loss=0.464]📉 Epoch 8: train\_loss=0.2294 | val\_loss=0.4635

Epoch 8: 100%|██████████| 205/205 [02:23<00:00, 1.43it/s, train\_loss=0.224, val\_loss=0.468]📉 Epoch 9: train\_loss=0.2245 | val\_loss=0.4683

Epoch 8: 100%|██████████| 205/205 [02:23<00:00, 1.43it/s, train\_loss=0.224, val\_loss=0.468]Monitored metric val\_loss did not improve in the last 3 records. Best score: 0.472. Signaling Trainer to stop.

Epoch 8: 100%|██████████| 205/205 [02:23<00:00, 1.43it/s, train\_loss=0.224, val\_loss=0.468]

`predict()` was called with `n > output\_chunk\_length`: using auto-regression to forecast the values after `output\_chunk\_length` points. The model will access `(n - output\_chunk\_length)` future values of your `past\_covariates` (relative to the first predicted time step). To hide this warning, set `show\_warnings=False`.

GPU available: True (cuda), used: True

TPU available: False, using: 0 TPU cores

HPU available: False, using: 0 HPUs

LOCAL\_RANK: 0 - CUDA\_VISIBLE\_DEVICES: [0]

Predicting DataLoader 0: 100%|██████████| 1/1 [00:11<00:00, 0.08it/s][I 2025-05-10 22:31:00,995] Trial 23 finished with value: 288.19940185546875 and parameters: {'input\_chunk\_length': 72, 'output\_chunk\_length': 48, 'hidden\_size': 40, 'lstm\_layers': 2, 'num\_attention\_heads': 1, 'dropout': 0.5, 'n\_epochs': 30, 'lr': 0.01, 'batch\_size': 128}. Best is trial 19 with value: 270.6866760253906.

🔍 DEBUG: Dataövergripande information:

▶️ train: 2021-01-02 00:00:00 → 2023-12-31 23:00:00 (26256 timsteg)

✅ val: 2024-01-01 00:00:00 → 2024-12-31 23:00:00 (8784 timsteg)

⏪ past\_cov\_train: 2021-01-02 00:00:00 → 2023-12-31 23:00:00

⏪ past\_cov\_val: 2023-12-28 00:00:00 → 2024-12-31 23:00:00

⏩ future\_cov\_train: 2021-01-02 00:00:00 → 2023-12-31 23:00:00

⏩ future\_cov\_val: 2023-12-28 00:00:00 → 2024-12-31 23:00:00

GPU available: True (cuda), used: True

TPU available: False, using: 0 TPU cores

HPU available: False, using: 0 HPUs

LOCAL\_RANK: 0 - CUDA\_VISIBLE\_DEVICES: [0]

| Name | Type | Params | Mode

------------------------------------------------------------------------------------------------

0 | train\_metrics | MetricCollection | 0 | train

1 | val\_metrics | MetricCollection | 0 | train

2 | input\_embeddings | \_MultiEmbedding | 0 | train

3 | static\_covariates\_vsn | \_VariableSelectionNetwork | 0 | train

4 | encoder\_vsn | \_VariableSelectionNetwork | 201 K | train

5 | decoder\_vsn | \_VariableSelectionNetwork | 101 K | train

6 | static\_context\_grn | \_GatedResidualNetwork | 6.6 K | train

7 | static\_context\_hidden\_encoder\_grn | \_GatedResidualNetwork | 6.6 K | train

8 | static\_context\_cell\_encoder\_grn | \_GatedResidualNetwork | 6.6 K | train

9 | static\_context\_enrichment | \_GatedResidualNetwork | 6.6 K | train

10 | lstm\_encoder | LSTM | 26.2 K | train

11 | lstm\_decoder | LSTM | 26.2 K | train

12 | post\_lstm\_gan | \_GateAddNorm | 3.4 K | train

13 | static\_enrichment\_grn | \_GatedResidualNetwork | 8.2 K | train

14 | multihead\_attn | \_InterpretableMultiHeadAttention | 6.5 K | train

15 | post\_attn\_gan | \_GateAddNorm | 3.4 K | train

16 | feed\_forward\_block | \_GatedResidualNetwork | 6.6 K | train

17 | pre\_output\_gan | \_GateAddNorm | 3.4 K | train

18 | output\_layer | Linear | 123 | train

------------------------------------------------------------------------------------------------

412 K Trainable params

0 Non-trainable params

412 K Total params

1.648 Total estimated model params size (MB)

3056 Modules in train mode

0 Modules in eval mode

🚀 Startar ny gridsearch-körning med hyperparametrar:

input\_chunk\_length: 72

output\_chunk\_length: 48

hidden\_size: 40

lstm\_layers: 2

num\_attention\_heads: 1

dropout: 0.5

optimizer\_kwargs: {'lr': 0.01}

Epoch 0: 100%|██████████| 205/205 [02:23<00:00, 1.43it/s, train\_loss=0.550, val\_loss=0.528]📉 Epoch 1: train\_loss=0.5504 | val\_loss=0.5279

Epoch 0: 100%|██████████| 205/205 [02:23<00:00, 1.43it/s, train\_loss=0.550, val\_loss=0.528]Metric val\_loss improved. New best score: 0.528

Epoch 1: 100%|██████████| 205/205 [02:23<00:00, 1.43it/s, train\_loss=0.410, val\_loss=0.521]📉 Epoch 2: train\_loss=0.4105 | val\_loss=0.5211

Epoch 2: 100%|██████████| 205/205 [02:23<00:00, 1.42it/s, train\_loss=0.372, val\_loss=0.483]📉 Epoch 3: train\_loss=0.3723 | val\_loss=0.4830

Epoch 2: 100%|██████████| 205/205 [02:23<00:00, 1.42it/s, train\_loss=0.372, val\_loss=0.483]Metric val\_loss improved by 0.045 >= min\_delta = 0.01. New best score: 0.483

Epoch 3: 100%|██████████| 205/205 [02:23<00:00, 1.43it/s, train\_loss=0.324, val\_loss=0.568]📉 Epoch 4: train\_loss=0.3244 | val\_loss=0.5680

Epoch 4: 100%|██████████| 205/205 [02:23<00:00, 1.42it/s, train\_loss=0.331, val\_loss=0.443]📉 Epoch 5: train\_loss=0.3311 | val\_loss=0.4426

Epoch 4: 100%|██████████| 205/205 [02:23<00:00, 1.42it/s, train\_loss=0.331, val\_loss=0.443]Metric val\_loss improved by 0.040 >= min\_delta = 0.01. New best score: 0.443

Epoch 5: 100%|██████████| 205/205 [02:23<00:00, 1.43it/s, train\_loss=0.293, val\_loss=0.448]📉 Epoch 6: train\_loss=0.2929 | val\_loss=0.4478

Epoch 6: 100%|██████████| 205/205 [02:24<00:00, 1.42it/s, train\_loss=0.231, val\_loss=0.505]📉 Epoch 7: train\_loss=0.2309 | val\_loss=0.5050

Epoch 7: 100%|██████████| 205/205 [02:23<00:00, 1.43it/s, train\_loss=0.226, val\_loss=0.484]📉 Epoch 8: train\_loss=0.2260 | val\_loss=0.4843

Epoch 7: 100%|██████████| 205/205 [02:23<00:00, 1.43it/s, train\_loss=0.226, val\_loss=0.484]Monitored metric val\_loss did not improve in the last 3 records. Best score: 0.443. Signaling Trainer to stop.

Epoch 7: 100%|██████████| 205/205 [02:23<00:00, 1.43it/s, train\_loss=0.226, val\_loss=0.484]

`predict()` was called with `n > output\_chunk\_length`: using auto-regression to forecast the values after `output\_chunk\_length` points. The model will access `(n - output\_chunk\_length)` future values of your `past\_covariates` (relative to the first predicted time step). To hide this warning, set `show\_warnings=False`.

GPU available: True (cuda), used: True

TPU available: False, using: 0 TPU cores

HPU available: False, using: 0 HPUs

LOCAL\_RANK: 0 - CUDA\_VISIBLE\_DEVICES: [0]

Predicting DataLoader 0: 100%|██████████| 1/1 [00:11<00:00, 0.09it/s][I 2025-05-10 22:50:23,857] Trial 24 finished with value: 304.5734558105469 and parameters: {'input\_chunk\_length': 72, 'output\_chunk\_length': 48, 'hidden\_size': 40, 'lstm\_layers': 2, 'num\_attention\_heads': 1, 'dropout': 0.5, 'n\_epochs': 30, 'lr': 0.01, 'batch\_size': 128}. Best is trial 19 with value: 270.6866760253906.

🔍 DEBUG: Dataövergripande information:

▶️ train: 2021-01-02 00:00:00 → 2023-12-31 23:00:00 (26256 timsteg)

✅ val: 2024-01-01 00:00:00 → 2024-12-31 23:00:00 (8784 timsteg)

⏪ past\_cov\_train: 2021-01-02 00:00:00 → 2023-12-31 23:00:00

⏪ past\_cov\_val: 2023-12-28 00:00:00 → 2024-12-31 23:00:00

⏩ future\_cov\_train: 2021-01-02 00:00:00 → 2023-12-31 23:00:00

⏩ future\_cov\_val: 2023-12-28 00:00:00 → 2024-12-31 23:00:00

GPU available: True (cuda), used: True

TPU available: False, using: 0 TPU cores

HPU available: False, using: 0 HPUs

LOCAL\_RANK: 0 - CUDA\_VISIBLE\_DEVICES: [0]

| Name | Type | Params | Mode

------------------------------------------------------------------------------------------------

0 | train\_metrics | MetricCollection | 0 | train

1 | val\_metrics | MetricCollection | 0 | train

2 | input\_embeddings | \_MultiEmbedding | 0 | train

3 | static\_covariates\_vsn | \_VariableSelectionNetwork | 0 | train

4 | encoder\_vsn | \_VariableSelectionNetwork | 201 K | train

5 | decoder\_vsn | \_VariableSelectionNetwork | 101 K | train

6 | static\_context\_grn | \_GatedResidualNetwork | 6.6 K | train

7 | static\_context\_hidden\_encoder\_grn | \_GatedResidualNetwork | 6.6 K | train

8 | static\_context\_cell\_encoder\_grn | \_GatedResidualNetwork | 6.6 K | train

9 | static\_context\_enrichment | \_GatedResidualNetwork | 6.6 K | train

10 | lstm\_encoder | LSTM | 26.2 K | train

11 | lstm\_decoder | LSTM | 26.2 K | train

12 | post\_lstm\_gan | \_GateAddNorm | 3.4 K | train

13 | static\_enrichment\_grn | \_GatedResidualNetwork | 8.2 K | train

14 | multihead\_attn | \_InterpretableMultiHeadAttention | 6.5 K | train

15 | post\_attn\_gan | \_GateAddNorm | 3.4 K | train

16 | feed\_forward\_block | \_GatedResidualNetwork | 6.6 K | train

17 | pre\_output\_gan | \_GateAddNorm | 3.4 K | train

18 | output\_layer | Linear | 123 | train

------------------------------------------------------------------------------------------------

412 K Trainable params

0 Non-trainable params

412 K Total params

1.648 Total estimated model params size (MB)

3056 Modules in train mode

0 Modules in eval mode

🚀 Startar ny gridsearch-körning med hyperparametrar:

input\_chunk\_length: 72

output\_chunk\_length: 48

hidden\_size: 40

lstm\_layers: 2

num\_attention\_heads: 1

dropout: 0.5

optimizer\_kwargs: {'lr': 0.01}

Epoch 0: 100%|██████████| 205/205 [02:23<00:00, 1.43it/s, train\_loss=0.556, val\_loss=0.521]📉 Epoch 1: train\_loss=0.5562 | val\_loss=0.5208

Epoch 0: 100%|██████████| 205/205 [02:23<00:00, 1.43it/s, train\_loss=0.556, val\_loss=0.521]Metric val\_loss improved. New best score: 0.521

Epoch 1: 100%|██████████| 205/205 [02:23<00:00, 1.43it/s, train\_loss=0.423, val\_loss=0.518]📉 Epoch 2: train\_loss=0.4235 | val\_loss=0.5181

Epoch 2: 100%|██████████| 205/205 [02:23<00:00, 1.43it/s, train\_loss=0.367, val\_loss=0.516]📉 Epoch 3: train\_loss=0.3671 | val\_loss=0.5156

Epoch 3: 100%|██████████| 205/205 [02:23<00:00, 1.43it/s, train\_loss=0.346, val\_loss=0.586]📉 Epoch 4: train\_loss=0.3460 | val\_loss=0.5859

Epoch 3: 100%|██████████| 205/205 [02:23<00:00, 1.43it/s, train\_loss=0.346, val\_loss=0.586]Monitored metric val\_loss did not improve in the last 3 records. Best score: 0.521. Signaling Trainer to stop.

Epoch 3: 100%|██████████| 205/205 [02:23<00:00, 1.43it/s, train\_loss=0.346, val\_loss=0.586]

`predict()` was called with `n > output\_chunk\_length`: using auto-regression to forecast the values after `output\_chunk\_length` points. The model will access `(n - output\_chunk\_length)` future values of your `past\_covariates` (relative to the first predicted time step). To hide this warning, set `show\_warnings=False`.

GPU available: True (cuda), used: True

TPU available: False, using: 0 TPU cores

HPU available: False, using: 0 HPUs

LOCAL\_RANK: 0 - CUDA\_VISIBLE\_DEVICES: [0]

Predicting DataLoader 0: 100%|██████████| 1/1 [00:11<00:00, 0.08it/s][I 2025-05-10 23:00:12,066] Trial 25 finished with value: 401.185791015625 and parameters: {'input\_chunk\_length': 72, 'output\_chunk\_length': 48, 'hidden\_size': 40, 'lstm\_layers': 2, 'num\_attention\_heads': 1, 'dropout': 0.5, 'n\_epochs': 30, 'lr': 0.01, 'batch\_size': 128}. Best is trial 19 with value: 270.6866760253906.

🔍 DEBUG: Dataövergripande information:

▶️ train: 2021-01-02 00:00:00 → 2023-12-31 23:00:00 (26256 timsteg)

✅ val: 2024-01-01 00:00:00 → 2024-12-31 23:00:00 (8784 timsteg)

⏪ past\_cov\_train: 2021-01-02 00:00:00 → 2023-12-31 23:00:00

⏪ past\_cov\_val: 2023-12-28 00:00:00 → 2024-12-31 23:00:00

⏩ future\_cov\_train: 2021-01-02 00:00:00 → 2023-12-31 23:00:00

⏩ future\_cov\_val: 2023-12-28 00:00:00 → 2024-12-31 23:00:00

GPU available: True (cuda), used: True

TPU available: False, using: 0 TPU cores

HPU available: False, using: 0 HPUs

LOCAL\_RANK: 0 - CUDA\_VISIBLE\_DEVICES: [0]

| Name | Type | Params | Mode

------------------------------------------------------------------------------------------------

0 | train\_metrics | MetricCollection | 0 | train

1 | val\_metrics | MetricCollection | 0 | train

2 | input\_embeddings | \_MultiEmbedding | 0 | train

3 | static\_covariates\_vsn | \_VariableSelectionNetwork | 0 | train

4 | encoder\_vsn | \_VariableSelectionNetwork | 201 K | train

5 | decoder\_vsn | \_VariableSelectionNetwork | 101 K | train

6 | static\_context\_grn | \_GatedResidualNetwork | 6.6 K | train

7 | static\_context\_hidden\_encoder\_grn | \_GatedResidualNetwork | 6.6 K | train

8 | static\_context\_cell\_encoder\_grn | \_GatedResidualNetwork | 6.6 K | train

9 | static\_context\_enrichment | \_GatedResidualNetwork | 6.6 K | train

10 | lstm\_encoder | LSTM | 26.2 K | train

11 | lstm\_decoder | LSTM | 26.2 K | train

12 | post\_lstm\_gan | \_GateAddNorm | 3.4 K | train

13 | static\_enrichment\_grn | \_GatedResidualNetwork | 8.2 K | train

14 | multihead\_attn | \_InterpretableMultiHeadAttention | 6.5 K | train

15 | post\_attn\_gan | \_GateAddNorm | 3.4 K | train

16 | feed\_forward\_block | \_GatedResidualNetwork | 6.6 K | train

17 | pre\_output\_gan | \_GateAddNorm | 3.4 K | train

18 | output\_layer | Linear | 123 | train

------------------------------------------------------------------------------------------------

412 K Trainable params

0 Non-trainable params

412 K Total params

1.648 Total estimated model params size (MB)

3056 Modules in train mode

0 Modules in eval mode

🚀 Startar ny gridsearch-körning med hyperparametrar:

input\_chunk\_length: 72

output\_chunk\_length: 48

hidden\_size: 40

lstm\_layers: 2

num\_attention\_heads: 1

dropout: 0.5

optimizer\_kwargs: {'lr': 0.01}

Epoch 0: 100%|██████████| 205/205 [02:23<00:00, 1.43it/s, train\_loss=0.586, val\_loss=0.558]📉 Epoch 1: train\_loss=0.5859 | val\_loss=0.5576

Epoch 0: 100%|██████████| 205/205 [02:23<00:00, 1.42it/s, train\_loss=0.586, val\_loss=0.558]Metric val\_loss improved. New best score: 0.558

Epoch 1: 100%|██████████| 205/205 [02:23<00:00, 1.43it/s, train\_loss=0.416, val\_loss=0.556]📉 Epoch 2: train\_loss=0.4156 | val\_loss=0.5561

Epoch 2: 100%|██████████| 205/205 [02:23<00:00, 1.43it/s, train\_loss=0.366, val\_loss=0.496]📉 Epoch 3: train\_loss=0.3655 | val\_loss=0.4958

Epoch 2: 100%|██████████| 205/205 [02:23<00:00, 1.43it/s, train\_loss=0.366, val\_loss=0.496]Metric val\_loss improved by 0.062 >= min\_delta = 0.01. New best score: 0.496

Epoch 3: 100%|██████████| 205/205 [02:23<00:00, 1.43it/s, train\_loss=0.347, val\_loss=0.543]📉 Epoch 4: train\_loss=0.3468 | val\_loss=0.5427

Epoch 4: 100%|██████████| 205/205 [02:23<00:00, 1.43it/s, train\_loss=0.341, val\_loss=0.540]📉 Epoch 5: train\_loss=0.3412 | val\_loss=0.5399

Epoch 5: 100%|██████████| 205/205 [02:23<00:00, 1.43it/s, train\_loss=0.265, val\_loss=0.471]📉 Epoch 6: train\_loss=0.2650 | val\_loss=0.4708

Epoch 5: 100%|██████████| 205/205 [02:23<00:00, 1.43it/s, train\_loss=0.265, val\_loss=0.471]Metric val\_loss improved by 0.025 >= min\_delta = 0.01. New best score: 0.471

Epoch 6: 100%|██████████| 205/205 [02:23<00:00, 1.43it/s, train\_loss=0.236, val\_loss=0.542]📉 Epoch 7: train\_loss=0.2357 | val\_loss=0.5418

Epoch 7: 100%|██████████| 205/205 [02:23<00:00, 1.43it/s, train\_loss=0.234, val\_loss=0.518]📉 Epoch 8: train\_loss=0.2338 | val\_loss=0.5178

Epoch 8: 100%|██████████| 205/205 [02:23<00:00, 1.43it/s, train\_loss=0.225, val\_loss=0.538]📉 Epoch 9: train\_loss=0.2247 | val\_loss=0.5379

Epoch 8: 100%|██████████| 205/205 [02:23<00:00, 1.43it/s, train\_loss=0.225, val\_loss=0.538]Monitored metric val\_loss did not improve in the last 3 records. Best score: 0.471. Signaling Trainer to stop.

Epoch 8: 100%|██████████| 205/205 [02:23<00:00, 1.43it/s, train\_loss=0.225, val\_loss=0.538]

`predict()` was called with `n > output\_chunk\_length`: using auto-regression to forecast the values after `output\_chunk\_length` points. The model will access `(n - output\_chunk\_length)` future values of your `past\_covariates` (relative to the first predicted time step). To hide this warning, set `show\_warnings=False`.

GPU available: True (cuda), used: True

TPU available: False, using: 0 TPU cores

HPU available: False, using: 0 HPUs

LOCAL\_RANK: 0 - CUDA\_VISIBLE\_DEVICES: [0]

Predicting DataLoader 0: 100%|██████████| 1/1 [00:11<00:00, 0.09it/s][I 2025-05-10 23:21:57,883] Trial 26 finished with value: 308.3460998535156 and parameters: {'input\_chunk\_length': 72, 'output\_chunk\_length': 48, 'hidden\_size': 40, 'lstm\_layers': 2, 'num\_attention\_heads': 1, 'dropout': 0.5, 'n\_epochs': 30, 'lr': 0.01, 'batch\_size': 128}. Best is trial 19 with value: 270.6866760253906.

🔍 DEBUG: Dataövergripande information:

▶️ train: 2021-01-02 00:00:00 → 2023-12-31 23:00:00 (26256 timsteg)

✅ val: 2024-01-01 00:00:00 → 2024-12-31 23:00:00 (8784 timsteg)

⏪ past\_cov\_train: 2021-01-02 00:00:00 → 2023-12-31 23:00:00

⏪ past\_cov\_val: 2023-12-28 00:00:00 → 2024-12-31 23:00:00

⏩ future\_cov\_train: 2021-01-02 00:00:00 → 2023-12-31 23:00:00

⏩ future\_cov\_val: 2023-12-28 00:00:00 → 2024-12-31 23:00:00

GPU available: True (cuda), used: True

TPU available: False, using: 0 TPU cores

HPU available: False, using: 0 HPUs

LOCAL\_RANK: 0 - CUDA\_VISIBLE\_DEVICES: [0]

| Name | Type | Params | Mode

------------------------------------------------------------------------------------------------

0 | train\_metrics | MetricCollection | 0 | train

1 | val\_metrics | MetricCollection | 0 | train

2 | input\_embeddings | \_MultiEmbedding | 0 | train

3 | static\_covariates\_vsn | \_VariableSelectionNetwork | 0 | train

4 | encoder\_vsn | \_VariableSelectionNetwork | 201 K | train

5 | decoder\_vsn | \_VariableSelectionNetwork | 101 K | train

6 | static\_context\_grn | \_GatedResidualNetwork | 6.6 K | train

7 | static\_context\_hidden\_encoder\_grn | \_GatedResidualNetwork | 6.6 K | train

8 | static\_context\_cell\_encoder\_grn | \_GatedResidualNetwork | 6.6 K | train

9 | static\_context\_enrichment | \_GatedResidualNetwork | 6.6 K | train

10 | lstm\_encoder | LSTM | 26.2 K | train

11 | lstm\_decoder | LSTM | 26.2 K | train

12 | post\_lstm\_gan | \_GateAddNorm | 3.4 K | train

13 | static\_enrichment\_grn | \_GatedResidualNetwork | 8.2 K | train

14 | multihead\_attn | \_InterpretableMultiHeadAttention | 6.5 K | train

15 | post\_attn\_gan | \_GateAddNorm | 3.4 K | train

16 | feed\_forward\_block | \_GatedResidualNetwork | 6.6 K | train

17 | pre\_output\_gan | \_GateAddNorm | 3.4 K | train

18 | output\_layer | Linear | 123 | train

------------------------------------------------------------------------------------------------

412 K Trainable params

0 Non-trainable params

412 K Total params

1.648 Total estimated model params size (MB)

3056 Modules in train mode

0 Modules in eval mode

🚀 Startar ny gridsearch-körning med hyperparametrar:

input\_chunk\_length: 72

output\_chunk\_length: 48

hidden\_size: 40

lstm\_layers: 2

num\_attention\_heads: 1

dropout: 0.4

optimizer\_kwargs: {'lr': 0.01}

Epoch 0: 100%|██████████| 205/205 [02:23<00:00, 1.42it/s, train\_loss=0.525, val\_loss=0.543]📉 Epoch 1: train\_loss=0.5248 | val\_loss=0.5433

Epoch 0: 100%|██████████| 205/205 [02:23<00:00, 1.42it/s, train\_loss=0.525, val\_loss=0.543]Metric val\_loss improved. New best score: 0.543

Epoch 1: 100%|██████████| 205/205 [02:23<00:00, 1.43it/s, train\_loss=0.389, val\_loss=0.488]📉 Epoch 2: train\_loss=0.3893 | val\_loss=0.4882

Epoch 1: 100%|██████████| 205/205 [02:23<00:00, 1.43it/s, train\_loss=0.389, val\_loss=0.488]Metric val\_loss improved by 0.055 >= min\_delta = 0.01. New best score: 0.488

Epoch 2: 100%|██████████| 205/205 [02:23<00:00, 1.43it/s, train\_loss=0.337, val\_loss=0.439]📉 Epoch 3: train\_loss=0.3375 | val\_loss=0.4392

Epoch 2: 100%|██████████| 205/205 [02:23<00:00, 1.43it/s, train\_loss=0.337, val\_loss=0.439]Metric val\_loss improved by 0.049 >= min\_delta = 0.01. New best score: 0.439

Epoch 3: 100%|██████████| 205/205 [02:23<00:00, 1.43it/s, train\_loss=0.318, val\_loss=0.494]📉 Epoch 4: train\_loss=0.3178 | val\_loss=0.4938

Epoch 4: 100%|██████████| 205/205 [02:23<00:00, 1.43it/s, train\_loss=0.312, val\_loss=0.507]📉 Epoch 5: train\_loss=0.3117 | val\_loss=0.5075

Epoch 5: 100%|██████████| 205/205 [02:23<00:00, 1.43it/s, train\_loss=0.244, val\_loss=0.505]📉 Epoch 6: train\_loss=0.2435 | val\_loss=0.5052

Epoch 5: 100%|██████████| 205/205 [02:23<00:00, 1.43it/s, train\_loss=0.244, val\_loss=0.505]Monitored metric val\_loss did not improve in the last 3 records. Best score: 0.439. Signaling Trainer to stop.

Epoch 5: 100%|██████████| 205/205 [02:23<00:00, 1.43it/s, train\_loss=0.244, val\_loss=0.505]

`predict()` was called with `n > output\_chunk\_length`: using auto-regression to forecast the values after `output\_chunk\_length` points. The model will access `(n - output\_chunk\_length)` future values of your `past\_covariates` (relative to the first predicted time step). To hide this warning, set `show\_warnings=False`.

GPU available: True (cuda), used: True

TPU available: False, using: 0 TPU cores

HPU available: False, using: 0 HPUs

LOCAL\_RANK: 0 - CUDA\_VISIBLE\_DEVICES: [0]

Predicting DataLoader 0: 100%|██████████| 1/1 [00:11<00:00, 0.09it/s][I 2025-05-10 23:36:33,900] Trial 27 finished with value: 293.2291564941406 and parameters: {'input\_chunk\_length': 72, 'output\_chunk\_length': 48, 'hidden\_size': 40, 'lstm\_layers': 2, 'num\_attention\_heads': 1, 'dropout': 0.4, 'n\_epochs': 30, 'lr': 0.01, 'batch\_size': 128}. Best is trial 19 with value: 270.6866760253906.

🔍 DEBUG: Dataövergripande information:

▶️ train: 2021-01-02 00:00:00 → 2023-12-31 23:00:00 (26256 timsteg)

✅ val: 2024-01-01 00:00:00 → 2024-12-31 23:00:00 (8784 timsteg)

⏪ past\_cov\_train: 2021-01-02 00:00:00 → 2023-12-31 23:00:00

⏪ past\_cov\_val: 2023-12-28 00:00:00 → 2024-12-31 23:00:00

⏩ future\_cov\_train: 2021-01-02 00:00:00 → 2023-12-31 23:00:00

⏩ future\_cov\_val: 2023-12-28 00:00:00 → 2024-12-31 23:00:00

GPU available: True (cuda), used: True

TPU available: False, using: 0 TPU cores

HPU available: False, using: 0 HPUs

LOCAL\_RANK: 0 - CUDA\_VISIBLE\_DEVICES: [0]

| Name | Type | Params | Mode

------------------------------------------------------------------------------------------------

0 | train\_metrics | MetricCollection | 0 | train

1 | val\_metrics | MetricCollection | 0 | train

2 | input\_embeddings | \_MultiEmbedding | 0 | train

3 | static\_covariates\_vsn | \_VariableSelectionNetwork | 0 | train

4 | encoder\_vsn | \_VariableSelectionNetwork | 201 K | train

5 | decoder\_vsn | \_VariableSelectionNetwork | 101 K | train

6 | static\_context\_grn | \_GatedResidualNetwork | 6.6 K | train

7 | static\_context\_hidden\_encoder\_grn | \_GatedResidualNetwork | 6.6 K | train

8 | static\_context\_cell\_encoder\_grn | \_GatedResidualNetwork | 6.6 K | train

9 | static\_context\_enrichment | \_GatedResidualNetwork | 6.6 K | train

10 | lstm\_encoder | LSTM | 26.2 K | train

11 | lstm\_decoder | LSTM | 26.2 K | train

12 | post\_lstm\_gan | \_GateAddNorm | 3.4 K | train

13 | static\_enrichment\_grn | \_GatedResidualNetwork | 8.2 K | train

14 | multihead\_attn | \_InterpretableMultiHeadAttention | 4.1 K | train

15 | post\_attn\_gan | \_GateAddNorm | 3.4 K | train

16 | feed\_forward\_block | \_GatedResidualNetwork | 6.6 K | train

17 | pre\_output\_gan | \_GateAddNorm | 3.4 K | train

18 | output\_layer | Linear | 123 | train

------------------------------------------------------------------------------------------------

409 K Trainable params

0 Non-trainable params

409 K Total params

1.639 Total estimated model params size (MB)

3062 Modules in train mode

0 Modules in eval mode

🚀 Startar ny gridsearch-körning med hyperparametrar:

input\_chunk\_length: 48

output\_chunk\_length: 48

hidden\_size: 40

lstm\_layers: 2

num\_attention\_heads: 4

dropout: 0.2

optimizer\_kwargs: {'lr': 0.01}

Epoch 0: 100%|██████████| 205/205 [01:48<00:00, 1.88it/s, train\_loss=0.457, val\_loss=0.625]📉 Epoch 1: train\_loss=0.4570 | val\_loss=0.6250

Epoch 0: 100%|██████████| 205/205 [01:49<00:00, 1.88it/s, train\_loss=0.457, val\_loss=0.625]Metric val\_loss improved. New best score: 0.625

Epoch 1: 100%|██████████| 205/205 [01:48<00:00, 1.88it/s, train\_loss=0.374, val\_loss=0.465]📉 Epoch 2: train\_loss=0.3743 | val\_loss=0.4650

Epoch 1: 100%|██████████| 205/205 [01:48<00:00, 1.88it/s, train\_loss=0.374, val\_loss=0.465]Metric val\_loss improved by 0.160 >= min\_delta = 0.01. New best score: 0.465

Epoch 2: 100%|██████████| 205/205 [01:48<00:00, 1.88it/s, train\_loss=0.345, val\_loss=0.441]📉 Epoch 3: train\_loss=0.3449 | val\_loss=0.4413

Epoch 2: 100%|██████████| 205/205 [01:48<00:00, 1.88it/s, train\_loss=0.345, val\_loss=0.441]Metric val\_loss improved by 0.024 >= min\_delta = 0.01. New best score: 0.441

Epoch 3: 100%|██████████| 205/205 [01:49<00:00, 1.88it/s, train\_loss=0.290, val\_loss=0.489]📉 Epoch 4: train\_loss=0.2895 | val\_loss=0.4885

Epoch 4: 100%|██████████| 205/205 [01:48<00:00, 1.88it/s, train\_loss=0.255, val\_loss=0.590]📉 Epoch 5: train\_loss=0.2547 | val\_loss=0.5896

Epoch 5: 100%|██████████| 205/205 [01:48<00:00, 1.88it/s, train\_loss=0.239, val\_loss=0.507]📉 Epoch 6: train\_loss=0.2388 | val\_loss=0.5070

Epoch 5: 100%|██████████| 205/205 [01:48<00:00, 1.88it/s, train\_loss=0.239, val\_loss=0.507]Monitored metric val\_loss did not improve in the last 3 records. Best score: 0.441. Signaling Trainer to stop.

Epoch 5: 100%|██████████| 205/205 [01:48<00:00, 1.88it/s, train\_loss=0.239, val\_loss=0.507]

`predict()` was called with `n > output\_chunk\_length`: using auto-regression to forecast the values after `output\_chunk\_length` points. The model will access `(n - output\_chunk\_length)` future values of your `past\_covariates` (relative to the first predicted time step). To hide this warning, set `show\_warnings=False`.

GPU available: True (cuda), used: True

TPU available: False, using: 0 TPU cores

HPU available: False, using: 0 HPUs

LOCAL\_RANK: 0 - CUDA\_VISIBLE\_DEVICES: [0]

Predicting DataLoader 0: 100%|██████████| 1/1 [00:11<00:00, 0.09it/s][I 2025-05-10 23:47:40,779] Trial 28 finished with value: 290.2784423828125 and parameters: {'input\_chunk\_length': 48, 'output\_chunk\_length': 48, 'hidden\_size': 40, 'lstm\_layers': 2, 'num\_attention\_heads': 4, 'dropout': 0.2, 'n\_epochs': 30, 'lr': 0.01, 'batch\_size': 128}. Best is trial 19 with value: 270.6866760253906.

🔍 DEBUG: Dataövergripande information:

▶️ train: 2021-01-02 00:00:00 → 2023-12-31 23:00:00 (26256 timsteg)

✅ val: 2024-01-01 00:00:00 → 2024-12-31 23:00:00 (8784 timsteg)

⏪ past\_cov\_train: 2021-01-02 00:00:00 → 2023-12-31 23:00:00

⏪ past\_cov\_val: 2023-12-28 00:00:00 → 2024-12-31 23:00:00

⏩ future\_cov\_train: 2021-01-02 00:00:00 → 2023-12-31 23:00:00

⏩ future\_cov\_val: 2023-12-28 00:00:00 → 2024-12-31 23:00:00

GPU available: True (cuda), used: True

TPU available: False, using: 0 TPU cores

HPU available: False, using: 0 HPUs

LOCAL\_RANK: 0 - CUDA\_VISIBLE\_DEVICES: [0]

| Name | Type | Params | Mode

------------------------------------------------------------------------------------------------

0 | train\_metrics | MetricCollection | 0 | train

1 | val\_metrics | MetricCollection | 0 | train

2 | input\_embeddings | \_MultiEmbedding | 0 | train

3 | static\_covariates\_vsn | \_VariableSelectionNetwork | 0 | train

4 | encoder\_vsn | \_VariableSelectionNetwork | 110 K | train

5 | decoder\_vsn | \_VariableSelectionNetwork | 55.4 K | train

6 | static\_context\_grn | \_GatedResidualNetwork | 1.7 K | train

7 | static\_context\_hidden\_encoder\_grn | \_GatedResidualNetwork | 1.7 K | train

8 | static\_context\_cell\_encoder\_grn | \_GatedResidualNetwork | 1.7 K | train

9 | static\_context\_enrichment | \_GatedResidualNetwork | 1.7 K | train

10 | lstm\_encoder | LSTM | 3.4 K | train

11 | lstm\_decoder | LSTM | 3.4 K | train

12 | post\_lstm\_gan | \_GateAddNorm | 880 | train

13 | static\_enrichment\_grn | \_GatedResidualNetwork | 2.1 K | train

14 | multihead\_attn | \_InterpretableMultiHeadAttention | 1.7 K | train

15 | post\_attn\_gan | \_GateAddNorm | 880 | train

16 | feed\_forward\_block | \_GatedResidualNetwork | 1.7 K | train

17 | pre\_output\_gan | \_GateAddNorm | 880 | train

18 | output\_layer | Linear | 63 | train

------------------------------------------------------------------------------------------------

187 K Trainable params

0 Non-trainable params

187 K Total params

0.748 Total estimated model params size (MB)

3056 Modules in train mode

0 Modules in eval mode

🚀 Startar ny gridsearch-körning med hyperparametrar:

input\_chunk\_length: 72

output\_chunk\_length: 48

hidden\_size: 20

lstm\_layers: 1

num\_attention\_heads: 1

dropout: 0.5

optimizer\_kwargs: {'lr': 0.01}

Epoch 0: 100%|██████████| 205/205 [02:22<00:00, 1.44it/s, train\_loss=0.658, val\_loss=0.699]📉 Epoch 1: train\_loss=0.6581 | val\_loss=0.6991

Epoch 0: 100%|██████████| 205/205 [02:22<00:00, 1.44it/s, train\_loss=0.658, val\_loss=0.699]Metric val\_loss improved. New best score: 0.699

Epoch 1: 100%|██████████| 205/205 [02:21<00:00, 1.45it/s, train\_loss=0.417, val\_loss=0.594]📉 Epoch 2: train\_loss=0.4172 | val\_loss=0.5942

Epoch 1: 100%|██████████| 205/205 [02:21<00:00, 1.45it/s, train\_loss=0.417, val\_loss=0.594]Metric val\_loss improved by 0.105 >= min\_delta = 0.01. New best score: 0.594

Epoch 2: 100%|██████████| 205/205 [02:21<00:00, 1.45it/s, train\_loss=0.385, val\_loss=0.457]📉 Epoch 3: train\_loss=0.3850 | val\_loss=0.4574

Epoch 2: 100%|██████████| 205/205 [02:21<00:00, 1.45it/s, train\_loss=0.385, val\_loss=0.457]Metric val\_loss improved by 0.137 >= min\_delta = 0.01. New best score: 0.457

Epoch 3: 100%|██████████| 205/205 [02:21<00:00, 1.45it/s, train\_loss=0.310, val\_loss=0.476]📉 Epoch 4: train\_loss=0.3102 | val\_loss=0.4758

Epoch 4: 100%|██████████| 205/205 [02:21<00:00, 1.45it/s, train\_loss=0.340, val\_loss=0.491]📉 Epoch 5: train\_loss=0.3399 | val\_loss=0.4912

Epoch 5: 100%|██████████| 205/205 [02:21<00:00, 1.44it/s, train\_loss=0.271, val\_loss=0.445]📉 Epoch 6: train\_loss=0.2715 | val\_loss=0.4453

Epoch 5: 100%|██████████| 205/205 [02:21<00:00, 1.44it/s, train\_loss=0.271, val\_loss=0.445]Metric val\_loss improved by 0.012 >= min\_delta = 0.01. New best score: 0.445

Epoch 6: 100%|██████████| 205/205 [02:21<00:00, 1.45it/s, train\_loss=0.301, val\_loss=0.460]📉 Epoch 7: train\_loss=0.3006 | val\_loss=0.4599

Epoch 7: 100%|██████████| 205/205 [02:21<00:00, 1.45it/s, train\_loss=0.255, val\_loss=0.487]📉 Epoch 8: train\_loss=0.2546 | val\_loss=0.4870

Epoch 8: 100%|██████████| 205/205 [02:21<00:00, 1.45it/s, train\_loss=0.265, val\_loss=0.445]📉 Epoch 9: train\_loss=0.2647 | val\_loss=0.4452

Epoch 8: 100%|██████████| 205/205 [02:21<00:00, 1.45it/s, train\_loss=0.265, val\_loss=0.445]Monitored metric val\_loss did not improve in the last 3 records. Best score: 0.445. Signaling Trainer to stop.

Epoch 8: 100%|██████████| 205/205 [02:21<00:00, 1.45it/s, train\_loss=0.265, val\_loss=0.445]

`predict()` was called with `n > output\_chunk\_length`: using auto-regression to forecast the values after `output\_chunk\_length` points. The model will access `(n - output\_chunk\_length)` future values of your `past\_covariates` (relative to the first predicted time step). To hide this warning, set `show\_warnings=False`.

GPU available: True (cuda), used: True

TPU available: False, using: 0 TPU cores

HPU available: False, using: 0 HPUs

LOCAL\_RANK: 0 - CUDA\_VISIBLE\_DEVICES: [0]

Predicting DataLoader 0: 100%|██████████| 1/1 [00:11<00:00, 0.09it/s]

[I 2025-05-11 00:09:10,908] Trial 29 finished with value: 290.8540954589844 and parameters: {'input\_chunk\_length': 72, 'output\_chunk\_length': 48, 'hidden\_size': 20, 'lstm\_layers': 1, 'num\_attention\_heads': 1, 'dropout': 0.5, 'n\_epochs': 30, 'lr': 0.01, 'batch\_size': 128}. Best is trial 19 with value: 270.6866760253906.

🔍 DEBUG: Dataövergripande information:

▶️ train: 2021-01-02 00:00:00 → 2023-12-31 23:00:00 (26256 timsteg)

✅ val: 2024-01-01 00:00:00 → 2024-12-31 23:00:00 (8784 timsteg)

⏪ past\_cov\_train: 2021-01-02 00:00:00 → 2023-12-31 23:00:00

⏪ past\_cov\_val: 2023-12-28 00:00:00 → 2024-12-31 23:00:00

⏩ future\_cov\_train: 2021-01-02 00:00:00 → 2023-12-31 23:00:00

⏩ future\_cov\_val: 2023-12-28 00:00:00 → 2024-12-31 23:00:00

GPU available: True (cuda), used: True

TPU available: False, using: 0 TPU cores

HPU available: False, using: 0 HPUs

LOCAL\_RANK: 0 - CUDA\_VISIBLE\_DEVICES: [0]

| Name | Type | Params | Mode

------------------------------------------------------------------------------------------------

0 | train\_metrics | MetricCollection | 0 | train

1 | val\_metrics | MetricCollection | 0 | train

2 | input\_embeddings | \_MultiEmbedding | 0 | train

3 | static\_covariates\_vsn | \_VariableSelectionNetwork | 0 | train

4 | encoder\_vsn | \_VariableSelectionNetwork | 201 K | train

5 | decoder\_vsn | \_VariableSelectionNetwork | 101 K | train

6 | static\_context\_grn | \_GatedResidualNetwork | 6.6 K | train

7 | static\_context\_hidden\_encoder\_grn | \_GatedResidualNetwork | 6.6 K | train

8 | static\_context\_cell\_encoder\_grn | \_GatedResidualNetwork | 6.6 K | train

9 | static\_context\_enrichment | \_GatedResidualNetwork | 6.6 K | train

10 | lstm\_encoder | LSTM | 26.2 K | train

11 | lstm\_decoder | LSTM | 26.2 K | train

12 | post\_lstm\_gan | \_GateAddNorm | 3.4 K | train

13 | static\_enrichment\_grn | \_GatedResidualNetwork | 8.2 K | train

14 | multihead\_attn | \_InterpretableMultiHeadAttention | 6.5 K | train

15 | post\_attn\_gan | \_GateAddNorm | 3.4 K | train

16 | feed\_forward\_block | \_GatedResidualNetwork | 6.6 K | train

17 | pre\_output\_gan | \_GateAddNorm | 3.4 K | train

18 | output\_layer | Linear | 123 | train

------------------------------------------------------------------------------------------------

412 K Trainable params

0 Non-trainable params

412 K Total params

1.648 Total estimated model params size (MB)

3056 Modules in train mode

0 Modules in eval mode

🚀 Startar ny gridsearch-körning med hyperparametrar:

input\_chunk\_length: 72

output\_chunk\_length: 48

hidden\_size: 40

lstm\_layers: 2

num\_attention\_heads: 1

dropout: 0.4

optimizer\_kwargs: {'lr': 0.01}

Epoch 0: 100%|██████████| 205/205 [02:23<00:00, 1.42it/s, train\_loss=0.524, val\_loss=0.561]📉 Epoch 1: train\_loss=0.5240 | val\_loss=0.5611

Epoch 0: 100%|██████████| 205/205 [02:23<00:00, 1.42it/s, train\_loss=0.524, val\_loss=0.561]Metric val\_loss improved. New best score: 0.561

Epoch 1: 100%|██████████| 205/205 [02:23<00:00, 1.43it/s, train\_loss=0.395, val\_loss=0.450]📉 Epoch 2: train\_loss=0.3945 | val\_loss=0.4499

Epoch 1: 100%|██████████| 205/205 [02:23<00:00, 1.43it/s, train\_loss=0.395, val\_loss=0.450]Metric val\_loss improved by 0.111 >= min\_delta = 0.01. New best score: 0.450

Epoch 2: 100%|██████████| 205/205 [02:23<00:00, 1.43it/s, train\_loss=0.342, val\_loss=0.477]📉 Epoch 3: train\_loss=0.3421 | val\_loss=0.4765

Epoch 3: 100%|██████████| 205/205 [02:23<00:00, 1.43it/s, train\_loss=0.309, val\_loss=0.482]📉 Epoch 4: train\_loss=0.3094 | val\_loss=0.4818

Epoch 4: 100%|██████████| 205/205 [02:23<00:00, 1.43it/s, train\_loss=0.341, val\_loss=0.503]📉 Epoch 5: train\_loss=0.3413 | val\_loss=0.5028

Epoch 4: 100%|██████████| 205/205 [02:23<00:00, 1.43it/s, train\_loss=0.341, val\_loss=0.503]Monitored metric val\_loss did not improve in the last 3 records. Best score: 0.450. Signaling Trainer to stop.

Epoch 4: 100%|██████████| 205/205 [02:23<00:00, 1.43it/s, train\_loss=0.341, val\_loss=0.503]

`predict()` was called with `n > output\_chunk\_length`: using auto-regression to forecast the values after `output\_chunk\_length` points. The model will access `(n - output\_chunk\_length)` future values of your `past\_covariates` (relative to the first predicted time step). To hide this warning, set `show\_warnings=False`.

GPU available: True (cuda), used: True

TPU available: False, using: 0 TPU cores

HPU available: False, using: 0 HPUs

LOCAL\_RANK: 0 - CUDA\_VISIBLE\_DEVICES: [0]

Predicting DataLoader 0: 100%|██████████| 1/1 [00:11<00:00, 0.08it/s][I 2025-05-11 00:21:22,972] Trial 30 finished with value: 331.35687255859375 and parameters: {'input\_chunk\_length': 72, 'output\_chunk\_length': 48, 'hidden\_size': 40, 'lstm\_layers': 2, 'num\_attention\_heads': 1, 'dropout': 0.4, 'n\_epochs': 30, 'lr': 0.01, 'batch\_size': 128}. Best is trial 19 with value: 270.6866760253906.

🔍 DEBUG: Dataövergripande information:

▶️ train: 2021-01-02 00:00:00 → 2023-12-31 23:00:00 (26256 timsteg)

✅ val: 2024-01-01 00:00:00 → 2024-12-31 23:00:00 (8784 timsteg)

⏪ past\_cov\_train: 2021-01-02 00:00:00 → 2023-12-31 23:00:00

⏪ past\_cov\_val: 2023-12-28 00:00:00 → 2024-12-31 23:00:00

⏩ future\_cov\_train: 2021-01-02 00:00:00 → 2023-12-31 23:00:00

⏩ future\_cov\_val: 2023-12-28 00:00:00 → 2024-12-31 23:00:00

GPU available: True (cuda), used: True

TPU available: False, using: 0 TPU cores

HPU available: False, using: 0 HPUs

LOCAL\_RANK: 0 - CUDA\_VISIBLE\_DEVICES: [0]

| Name | Type | Params | Mode

------------------------------------------------------------------------------------------------

0 | train\_metrics | MetricCollection | 0 | train

1 | val\_metrics | MetricCollection | 0 | train

2 | input\_embeddings | \_MultiEmbedding | 0 | train

3 | static\_covariates\_vsn | \_VariableSelectionNetwork | 0 | train

4 | encoder\_vsn | \_VariableSelectionNetwork | 201 K | train

5 | decoder\_vsn | \_VariableSelectionNetwork | 101 K | train

6 | static\_context\_grn | \_GatedResidualNetwork | 6.6 K | train

7 | static\_context\_hidden\_encoder\_grn | \_GatedResidualNetwork | 6.6 K | train

8 | static\_context\_cell\_encoder\_grn | \_GatedResidualNetwork | 6.6 K | train

9 | static\_context\_enrichment | \_GatedResidualNetwork | 6.6 K | train

10 | lstm\_encoder | LSTM | 26.2 K | train

11 | lstm\_decoder | LSTM | 26.2 K | train

12 | post\_lstm\_gan | \_GateAddNorm | 3.4 K | train

13 | static\_enrichment\_grn | \_GatedResidualNetwork | 8.2 K | train

14 | multihead\_attn | \_InterpretableMultiHeadAttention | 4.1 K | train

15 | post\_attn\_gan | \_GateAddNorm | 3.4 K | train

16 | feed\_forward\_block | \_GatedResidualNetwork | 6.6 K | train

17 | pre\_output\_gan | \_GateAddNorm | 3.4 K | train

18 | output\_layer | Linear | 123 | train

------------------------------------------------------------------------------------------------

409 K Trainable params

0 Non-trainable params

409 K Total params

1.639 Total estimated model params size (MB)

3062 Modules in train mode

0 Modules in eval mode

🚀 Startar ny gridsearch-körning med hyperparametrar:

input\_chunk\_length: 48

output\_chunk\_length: 48

hidden\_size: 40

lstm\_layers: 2

num\_attention\_heads: 4

dropout: 0.2

optimizer\_kwargs: {'lr': 0.01}

Epoch 0: 100%|██████████| 205/205 [01:48<00:00, 1.88it/s, train\_loss=0.454, val\_loss=0.618]📉 Epoch 1: train\_loss=0.4543 | val\_loss=0.6177

Epoch 0: 100%|██████████| 205/205 [01:49<00:00, 1.88it/s, train\_loss=0.454, val\_loss=0.618]Metric val\_loss improved. New best score: 0.618

Epoch 1: 100%|██████████| 205/205 [01:48<00:00, 1.88it/s, train\_loss=0.373, val\_loss=0.452]📉 Epoch 2: train\_loss=0.3734 | val\_loss=0.4524

Epoch 1: 100%|██████████| 205/205 [01:48<00:00, 1.88it/s, train\_loss=0.373, val\_loss=0.452]Metric val\_loss improved by 0.165 >= min\_delta = 0.01. New best score: 0.452

Epoch 2: 100%|██████████| 205/205 [01:49<00:00, 1.88it/s, train\_loss=0.339, val\_loss=0.478]📉 Epoch 3: train\_loss=0.3393 | val\_loss=0.4776

Epoch 3: 100%|██████████| 205/205 [01:48<00:00, 1.88it/s, train\_loss=0.287, val\_loss=0.424]📉 Epoch 4: train\_loss=0.2874 | val\_loss=0.4240

Epoch 3: 100%|██████████| 205/205 [01:48<00:00, 1.88it/s, train\_loss=0.287, val\_loss=0.424]Metric val\_loss improved by 0.028 >= min\_delta = 0.01. New best score: 0.424

Epoch 4: 100%|██████████| 205/205 [01:48<00:00, 1.88it/s, train\_loss=0.252, val\_loss=0.540]📉 Epoch 5: train\_loss=0.2522 | val\_loss=0.5404

Epoch 5: 100%|██████████| 205/205 [01:48<00:00, 1.88it/s, train\_loss=0.237, val\_loss=0.513]📉 Epoch 6: train\_loss=0.2369 | val\_loss=0.5128

Epoch 6: 100%|██████████| 205/205 [01:48<00:00, 1.88it/s, train\_loss=0.224, val\_loss=0.520]📉 Epoch 7: train\_loss=0.2241 | val\_loss=0.5201

Epoch 6: 100%|██████████| 205/205 [01:49<00:00, 1.88it/s, train\_loss=0.224, val\_loss=0.520]Monitored metric val\_loss did not improve in the last 3 records. Best score: 0.424. Signaling Trainer to stop.

Epoch 6: 100%|██████████| 205/205 [01:49<00:00, 1.88it/s, train\_loss=0.224, val\_loss=0.520]

`predict()` was called with `n > output\_chunk\_length`: using auto-regression to forecast the values after `output\_chunk\_length` points. The model will access `(n - output\_chunk\_length)` future values of your `past\_covariates` (relative to the first predicted time step). To hide this warning, set `show\_warnings=False`.

GPU available: True (cuda), used: True

TPU available: False, using: 0 TPU cores

HPU available: False, using: 0 HPUs

LOCAL\_RANK: 0 - CUDA\_VISIBLE\_DEVICES: [0]

Predicting DataLoader 0: 100%|██████████| 1/1 [00:11<00:00, 0.08it/s][I 2025-05-11 00:34:19,673] Trial 31 finished with value: 312.764892578125 and parameters: {'input\_chunk\_length': 48, 'output\_chunk\_length': 48, 'hidden\_size': 40, 'lstm\_layers': 2, 'num\_attention\_heads': 4, 'dropout': 0.2, 'n\_epochs': 30, 'lr': 0.01, 'batch\_size': 128}. Best is trial 19 with value: 270.6866760253906.

🔍 DEBUG: Dataövergripande information:

▶️ train: 2021-01-02 00:00:00 → 2023-12-31 23:00:00 (26256 timsteg)

✅ val: 2024-01-01 00:00:00 → 2024-12-31 23:00:00 (8784 timsteg)

⏪ past\_cov\_train: 2021-01-02 00:00:00 → 2023-12-31 23:00:00

⏪ past\_cov\_val: 2023-12-28 00:00:00 → 2024-12-31 23:00:00

⏩ future\_cov\_train: 2021-01-02 00:00:00 → 2023-12-31 23:00:00

⏩ future\_cov\_val: 2023-12-28 00:00:00 → 2024-12-31 23:00:00

GPU available: True (cuda), used: True

TPU available: False, using: 0 TPU cores

HPU available: False, using: 0 HPUs

LOCAL\_RANK: 0 - CUDA\_VISIBLE\_DEVICES: [0]

| Name | Type | Params | Mode

------------------------------------------------------------------------------------------------

0 | train\_metrics | MetricCollection | 0 | train

1 | val\_metrics | MetricCollection | 0 | train

2 | input\_embeddings | \_MultiEmbedding | 0 | train

3 | static\_covariates\_vsn | \_VariableSelectionNetwork | 0 | train

4 | encoder\_vsn | \_VariableSelectionNetwork | 201 K | train

5 | decoder\_vsn | \_VariableSelectionNetwork | 101 K | train

6 | static\_context\_grn | \_GatedResidualNetwork | 6.6 K | train

7 | static\_context\_hidden\_encoder\_grn | \_GatedResidualNetwork | 6.6 K | train

8 | static\_context\_cell\_encoder\_grn | \_GatedResidualNetwork | 6.6 K | train

9 | static\_context\_enrichment | \_GatedResidualNetwork | 6.6 K | train

10 | lstm\_encoder | LSTM | 26.2 K | train

11 | lstm\_decoder | LSTM | 26.2 K | train

12 | post\_lstm\_gan | \_GateAddNorm | 3.4 K | train

13 | static\_enrichment\_grn | \_GatedResidualNetwork | 8.2 K | train

14 | multihead\_attn | \_InterpretableMultiHeadAttention | 4.1 K | train

15 | post\_attn\_gan | \_GateAddNorm | 3.4 K | train

16 | feed\_forward\_block | \_GatedResidualNetwork | 6.6 K | train

17 | pre\_output\_gan | \_GateAddNorm | 3.4 K | train

18 | output\_layer | Linear | 123 | train

------------------------------------------------------------------------------------------------

409 K Trainable params

0 Non-trainable params

409 K Total params

1.639 Total estimated model params size (MB)

3062 Modules in train mode

0 Modules in eval mode

🚀 Startar ny gridsearch-körning med hyperparametrar:

input\_chunk\_length: 48

output\_chunk\_length: 48

hidden\_size: 40

lstm\_layers: 2

num\_attention\_heads: 4

dropout: 0.2

optimizer\_kwargs: {'lr': 0.01}

Epoch 0: 100%|██████████| 205/205 [01:48<00:00, 1.88it/s, train\_loss=0.454, val\_loss=0.604]📉 Epoch 1: train\_loss=0.4538 | val\_loss=0.6036

Epoch 0: 100%|██████████| 205/205 [01:48<00:00, 1.88it/s, train\_loss=0.454, val\_loss=0.604]Metric val\_loss improved. New best score: 0.604

Epoch 1: 100%|██████████| 205/205 [01:48<00:00, 1.88it/s, train\_loss=0.372, val\_loss=0.473]📉 Epoch 2: train\_loss=0.3720 | val\_loss=0.4728

Epoch 1: 100%|██████████| 205/205 [01:48<00:00, 1.88it/s, train\_loss=0.372, val\_loss=0.473]Metric val\_loss improved by 0.131 >= min\_delta = 0.01. New best score: 0.473

Epoch 2: 100%|██████████| 205/205 [01:48<00:00, 1.89it/s, train\_loss=0.336, val\_loss=0.462]📉 Epoch 3: train\_loss=0.3358 | val\_loss=0.4621

Epoch 2: 100%|██████████| 205/205 [01:48<00:00, 1.89it/s, train\_loss=0.336, val\_loss=0.462]Metric val\_loss improved by 0.011 >= min\_delta = 0.01. New best score: 0.462

Epoch 3: 100%|██████████| 205/205 [01:49<00:00, 1.88it/s, train\_loss=0.299, val\_loss=0.445]📉 Epoch 4: train\_loss=0.2990 | val\_loss=0.4453

Epoch 3: 100%|██████████| 205/205 [01:49<00:00, 1.88it/s, train\_loss=0.299, val\_loss=0.445]Metric val\_loss improved by 0.017 >= min\_delta = 0.01. New best score: 0.445

Epoch 4: 100%|██████████| 205/205 [01:48<00:00, 1.88it/s, train\_loss=0.267, val\_loss=0.516]📉 Epoch 5: train\_loss=0.2666 | val\_loss=0.5156

Epoch 5: 100%|██████████| 205/205 [01:48<00:00, 1.88it/s, train\_loss=0.235, val\_loss=0.465]📉 Epoch 6: train\_loss=0.2354 | val\_loss=0.4648

Epoch 6: 100%|██████████| 205/205 [01:48<00:00, 1.88it/s, train\_loss=0.213, val\_loss=0.491]📉 Epoch 7: train\_loss=0.2129 | val\_loss=0.4908

Epoch 6: 100%|██████████| 205/205 [01:48<00:00, 1.88it/s, train\_loss=0.213, val\_loss=0.491]Monitored metric val\_loss did not improve in the last 3 records. Best score: 0.445. Signaling Trainer to stop.

Epoch 6: 100%|██████████| 205/205 [01:48<00:00, 1.88it/s, train\_loss=0.213, val\_loss=0.491]

`predict()` was called with `n > output\_chunk\_length`: using auto-regression to forecast the values after `output\_chunk\_length` points. The model will access `(n - output\_chunk\_length)` future values of your `past\_covariates` (relative to the first predicted time step). To hide this warning, set `show\_warnings=False`.

GPU available: True (cuda), used: True

TPU available: False, using: 0 TPU cores

HPU available: False, using: 0 HPUs

LOCAL\_RANK: 0 - CUDA\_VISIBLE\_DEVICES: [0]

Predicting DataLoader 0: 100%|██████████| 1/1 [00:12<00:00, 0.08it/s][I 2025-05-11 00:47:16,088] Trial 32 finished with value: 295.8298034667969 and parameters: {'input\_chunk\_length': 48, 'output\_chunk\_length': 48, 'hidden\_size': 40, 'lstm\_layers': 2, 'num\_attention\_heads': 4, 'dropout': 0.2, 'n\_epochs': 30, 'lr': 0.01, 'batch\_size': 128}. Best is trial 19 with value: 270.6866760253906.

🔍 DEBUG: Dataövergripande information:

▶️ train: 2021-01-02 00:00:00 → 2023-12-31 23:00:00 (26256 timsteg)

✅ val: 2024-01-01 00:00:00 → 2024-12-31 23:00:00 (8784 timsteg)

⏪ past\_cov\_train: 2021-01-02 00:00:00 → 2023-12-31 23:00:00

⏪ past\_cov\_val: 2023-12-28 00:00:00 → 2024-12-31 23:00:00

⏩ future\_cov\_train: 2021-01-02 00:00:00 → 2023-12-31 23:00:00

⏩ future\_cov\_val: 2023-12-28 00:00:00 → 2024-12-31 23:00:00

GPU available: True (cuda), used: True

TPU available: False, using: 0 TPU cores

HPU available: False, using: 0 HPUs

LOCAL\_RANK: 0 - CUDA\_VISIBLE\_DEVICES: [0]

| Name | Type | Params | Mode

------------------------------------------------------------------------------------------------

0 | train\_metrics | MetricCollection | 0 | train

1 | val\_metrics | MetricCollection | 0 | train

2 | input\_embeddings | \_MultiEmbedding | 0 | train

3 | static\_covariates\_vsn | \_VariableSelectionNetwork | 0 | train

4 | encoder\_vsn | \_VariableSelectionNetwork | 201 K | train

5 | decoder\_vsn | \_VariableSelectionNetwork | 101 K | train

6 | static\_context\_grn | \_GatedResidualNetwork | 6.6 K | train

7 | static\_context\_hidden\_encoder\_grn | \_GatedResidualNetwork | 6.6 K | train

8 | static\_context\_cell\_encoder\_grn | \_GatedResidualNetwork | 6.6 K | train

9 | static\_context\_enrichment | \_GatedResidualNetwork | 6.6 K | train

10 | lstm\_encoder | LSTM | 26.2 K | train

11 | lstm\_decoder | LSTM | 26.2 K | train

12 | post\_lstm\_gan | \_GateAddNorm | 3.4 K | train

13 | static\_enrichment\_grn | \_GatedResidualNetwork | 8.2 K | train

14 | multihead\_attn | \_InterpretableMultiHeadAttention | 4.1 K | train

15 | post\_attn\_gan | \_GateAddNorm | 3.4 K | train

16 | feed\_forward\_block | \_GatedResidualNetwork | 6.6 K | train

17 | pre\_output\_gan | \_GateAddNorm | 3.4 K | train

18 | output\_layer | Linear | 123 | train

------------------------------------------------------------------------------------------------

409 K Trainable params

0 Non-trainable params

409 K Total params

1.639 Total estimated model params size (MB)

3062 Modules in train mode

0 Modules in eval mode

🚀 Startar ny gridsearch-körning med hyperparametrar:

input\_chunk\_length: 48

output\_chunk\_length: 48

hidden\_size: 40

lstm\_layers: 2

num\_attention\_heads: 4

dropout: 0.2

optimizer\_kwargs: {'lr': 0.01}

Epoch 0: 100%|██████████| 205/205 [01:49<00:00, 1.88it/s, train\_loss=0.454, val\_loss=0.600]📉 Epoch 1: train\_loss=0.4540 | val\_loss=0.5996

Epoch 0: 100%|██████████| 205/205 [01:49<00:00, 1.88it/s, train\_loss=0.454, val\_loss=0.600]Metric val\_loss improved. New best score: 0.600

Epoch 1: 100%|██████████| 205/205 [01:48<00:00, 1.88it/s, train\_loss=0.375, val\_loss=0.461]📉 Epoch 2: train\_loss=0.3752 | val\_loss=0.4608

Epoch 1: 100%|██████████| 205/205 [01:48<00:00, 1.88it/s, train\_loss=0.375, val\_loss=0.461]Metric val\_loss improved by 0.139 >= min\_delta = 0.01. New best score: 0.461

Epoch 2: 100%|██████████| 205/205 [01:48<00:00, 1.88it/s, train\_loss=0.330, val\_loss=0.475]📉 Epoch 3: train\_loss=0.3301 | val\_loss=0.4752

Epoch 3: 100%|██████████| 205/205 [01:48<00:00, 1.88it/s, train\_loss=0.308, val\_loss=0.456]📉 Epoch 4: train\_loss=0.3083 | val\_loss=0.4557

Epoch 4: 100%|██████████| 205/205 [01:49<00:00, 1.88it/s, train\_loss=0.263, val\_loss=0.538]📉 Epoch 5: train\_loss=0.2634 | val\_loss=0.5383

Epoch 4: 100%|██████████| 205/205 [01:49<00:00, 1.88it/s, train\_loss=0.263, val\_loss=0.538]Monitored metric val\_loss did not improve in the last 3 records. Best score: 0.461. Signaling Trainer to stop.

Epoch 4: 100%|██████████| 205/205 [01:49<00:00, 1.88it/s, train\_loss=0.263, val\_loss=0.538]

`predict()` was called with `n > output\_chunk\_length`: using auto-regression to forecast the values after `output\_chunk\_length` points. The model will access `(n - output\_chunk\_length)` future values of your `past\_covariates` (relative to the first predicted time step). To hide this warning, set `show\_warnings=False`.

GPU available: True (cuda), used: True

TPU available: False, using: 0 TPU cores

HPU available: False, using: 0 HPUs

LOCAL\_RANK: 0 - CUDA\_VISIBLE\_DEVICES: [0]

Predicting DataLoader 0: 100%|██████████| 1/1 [00:11<00:00, 0.08it/s]

[I 2025-05-11 00:56:35,056] Trial 33 finished with value: 345.09576416015625 and parameters: {'input\_chunk\_length': 48, 'output\_chunk\_length': 48, 'hidden\_size': 40, 'lstm\_layers': 2, 'num\_attention\_heads': 4, 'dropout': 0.2, 'n\_epochs': 30, 'lr': 0.01, 'batch\_size': 128}. Best is trial 19 with value: 270.6866760253906.

🔍 DEBUG: Dataövergripande information:

▶️ train: 2021-01-02 00:00:00 → 2023-12-31 23:00:00 (26256 timsteg)

✅ val: 2024-01-01 00:00:00 → 2024-12-31 23:00:00 (8784 timsteg)

⏪ past\_cov\_train: 2021-01-02 00:00:00 → 2023-12-31 23:00:00

⏪ past\_cov\_val: 2023-12-28 00:00:00 → 2024-12-31 23:00:00

⏩ future\_cov\_train: 2021-01-02 00:00:00 → 2023-12-31 23:00:00

⏩ future\_cov\_val: 2023-12-28 00:00:00 → 2024-12-31 23:00:00

GPU available: True (cuda), used: True

TPU available: False, using: 0 TPU cores

HPU available: False, using: 0 HPUs

LOCAL\_RANK: 0 - CUDA\_VISIBLE\_DEVICES: [0]

| Name | Type | Params | Mode

------------------------------------------------------------------------------------------------

0 | train\_metrics | MetricCollection | 0 | train

1 | val\_metrics | MetricCollection | 0 | train

2 | input\_embeddings | \_MultiEmbedding | 0 | train

3 | static\_covariates\_vsn | \_VariableSelectionNetwork | 0 | train

4 | encoder\_vsn | \_VariableSelectionNetwork | 66.3 K | train

5 | decoder\_vsn | \_VariableSelectionNetwork | 33.0 K | train

6 | static\_context\_grn | \_GatedResidualNetwork | 460 | train

7 | static\_context\_hidden\_encoder\_grn | \_GatedResidualNetwork | 460 | train

8 | static\_context\_cell\_encoder\_grn | \_GatedResidualNetwork | 460 | train

9 | static\_context\_enrichment | \_GatedResidualNetwork | 460 | train

10 | lstm\_encoder | LSTM | 1.8 K | train

11 | lstm\_decoder | LSTM | 1.8 K | train

12 | post\_lstm\_gan | \_GateAddNorm | 240 | train

13 | static\_enrichment\_grn | \_GatedResidualNetwork | 560 | train

14 | multihead\_attn | \_InterpretableMultiHeadAttention | 218 | train

15 | post\_attn\_gan | \_GateAddNorm | 240 | train

16 | feed\_forward\_block | \_GatedResidualNetwork | 460 | train

17 | pre\_output\_gan | \_GateAddNorm | 240 | train

18 | output\_layer | Linear | 33 | train

------------------------------------------------------------------------------------------------

105 K Trainable params

0 Non-trainable params

105 K Total params

0.423 Total estimated model params size (MB)

3062 Modules in train mode

0 Modules in eval mode

🚀 Startar ny gridsearch-körning med hyperparametrar:

input\_chunk\_length: 48

output\_chunk\_length: 48

hidden\_size: 10

lstm\_layers: 2

num\_attention\_heads: 4

dropout: 0.2

optimizer\_kwargs: {'lr': 0.0001}

Epoch 0: 100%|██████████| 205/205 [01:51<00:00, 1.84it/s, train\_loss=6.560, val\_loss=7.050]📉 Epoch 1: train\_loss=6.5640 | val\_loss=7.0525

Epoch 0: 100%|██████████| 205/205 [01:51<00:00, 1.84it/s, train\_loss=6.560, val\_loss=7.050]Metric val\_loss improved. New best score: 7.052

Epoch 1: 100%|██████████| 205/205 [01:51<00:00, 1.85it/s, train\_loss=6.240, val\_loss=6.630]📉 Epoch 2: train\_loss=6.2433 | val\_loss=6.6260

Epoch 1: 100%|██████████| 205/205 [01:51<00:00, 1.85it/s, train\_loss=6.240, val\_loss=6.630]Metric val\_loss improved by 0.426 >= min\_delta = 0.01. New best score: 6.626

Epoch 2: 100%|██████████| 205/205 [01:51<00:00, 1.85it/s, train\_loss=5.560, val\_loss=6.160]📉 Epoch 3: train\_loss=5.5572 | val\_loss=6.1558

Epoch 2: 100%|██████████| 205/205 [01:51<00:00, 1.85it/s, train\_loss=5.560, val\_loss=6.160]Metric val\_loss improved by 0.470 >= min\_delta = 0.01. New best score: 6.156

Epoch 3: 100%|██████████| 205/205 [01:51<00:00, 1.84it/s, train\_loss=5.140, val\_loss=5.630]📉 Epoch 4: train\_loss=5.1374 | val\_loss=5.6327

Epoch 3: 100%|██████████| 205/205 [01:51<00:00, 1.84it/s, train\_loss=5.140, val\_loss=5.630]Metric val\_loss improved by 0.523 >= min\_delta = 0.01. New best score: 5.633

Epoch 4: 100%|██████████| 205/205 [01:51<00:00, 1.85it/s, train\_loss=4.500, val\_loss=5.070]📉 Epoch 5: train\_loss=4.5039 | val\_loss=5.0732

Epoch 4: 100%|██████████| 205/205 [01:51<00:00, 1.85it/s, train\_loss=4.500, val\_loss=5.070]Metric val\_loss improved by 0.560 >= min\_delta = 0.01. New best score: 5.073

Epoch 5: 100%|██████████| 205/205 [01:51<00:00, 1.85it/s, train\_loss=4.250, val\_loss=4.500]📉 Epoch 6: train\_loss=4.2454 | val\_loss=4.4951

Epoch 5: 100%|██████████| 205/205 [01:51<00:00, 1.85it/s, train\_loss=4.250, val\_loss=4.500]Metric val\_loss improved by 0.578 >= min\_delta = 0.01. New best score: 4.495

Epoch 6: 100%|██████████| 205/205 [01:51<00:00, 1.85it/s, train\_loss=3.230, val\_loss=3.920]📉 Epoch 7: train\_loss=3.2283 | val\_loss=3.9165

Epoch 6: 100%|██████████| 205/205 [01:51<00:00, 1.85it/s, train\_loss=3.230, val\_loss=3.920]Metric val\_loss improved by 0.579 >= min\_delta = 0.01. New best score: 3.916

Epoch 7: 100%|██████████| 205/205 [01:51<00:00, 1.84it/s, train\_loss=3.110, val\_loss=3.360]📉 Epoch 8: train\_loss=3.1087 | val\_loss=3.3579

Epoch 7: 100%|██████████| 205/205 [01:51<00:00, 1.84it/s, train\_loss=3.110, val\_loss=3.360]Metric val\_loss improved by 0.559 >= min\_delta = 0.01. New best score: 3.358

Epoch 8: 100%|██████████| 205/205 [01:51<00:00, 1.84it/s, train\_loss=2.270, val\_loss=2.840]📉 Epoch 9: train\_loss=2.2712 | val\_loss=2.8386

Epoch 8: 100%|██████████| 205/205 [01:51<00:00, 1.84it/s, train\_loss=2.270, val\_loss=2.840]Metric val\_loss improved by 0.519 >= min\_delta = 0.01. New best score: 2.839

Epoch 9: 100%|██████████| 205/205 [01:51<00:00, 1.85it/s, train\_loss=1.880, val\_loss=2.370]📉 Epoch 10: train\_loss=1.8792 | val\_loss=2.3723

Epoch 9: 100%|██████████| 205/205 [01:51<00:00, 1.85it/s, train\_loss=1.880, val\_loss=2.370]Metric val\_loss improved by 0.466 >= min\_delta = 0.01. New best score: 2.372

Epoch 10: 100%|██████████| 205/205 [01:51<00:00, 1.84it/s, train\_loss=1.660, val\_loss=1.970]📉 Epoch 11: train\_loss=1.6565 | val\_loss=1.9695

Epoch 10: 100%|██████████| 205/205 [01:51<00:00, 1.84it/s, train\_loss=1.660, val\_loss=1.970]Metric val\_loss improved by 0.403 >= min\_delta = 0.01. New best score: 1.970

Epoch 11: 100%|██████████| 205/205 [01:51<00:00, 1.85it/s, train\_loss=1.350, val\_loss=1.630]📉 Epoch 12: train\_loss=1.3482 | val\_loss=1.6267

Epoch 11: 100%|██████████| 205/205 [01:51<00:00, 1.85it/s, train\_loss=1.350, val\_loss=1.630]Metric val\_loss improved by 0.343 >= min\_delta = 0.01. New best score: 1.627

Epoch 12: 100%|██████████| 205/205 [01:51<00:00, 1.84it/s, train\_loss=1.240, val\_loss=1.330]📉 Epoch 13: train\_loss=1.2381 | val\_loss=1.3311

Epoch 12: 100%|██████████| 205/205 [01:51<00:00, 1.84it/s, train\_loss=1.240, val\_loss=1.330]Metric val\_loss improved by 0.296 >= min\_delta = 0.01. New best score: 1.331

Epoch 13: 100%|██████████| 205/205 [01:51<00:00, 1.85it/s, train\_loss=0.909, val\_loss=1.110]📉 Epoch 14: train\_loss=0.9087 | val\_loss=1.1075

Epoch 13: 100%|██████████| 205/205 [01:51<00:00, 1.85it/s, train\_loss=0.909, val\_loss=1.110]Metric val\_loss improved by 0.224 >= min\_delta = 0.01. New best score: 1.108

Epoch 14: 100%|██████████| 205/205 [01:51<00:00, 1.85it/s, train\_loss=0.681, val\_loss=0.957]📉 Epoch 15: train\_loss=0.6810 | val\_loss=0.9569

Epoch 14: 100%|██████████| 205/205 [01:51<00:00, 1.85it/s, train\_loss=0.681, val\_loss=0.957]Metric val\_loss improved by 0.151 >= min\_delta = 0.01. New best score: 0.957

Epoch 15: 100%|██████████| 205/205 [01:51<00:00, 1.85it/s, train\_loss=0.617, val\_loss=0.838]📉 Epoch 16: train\_loss=0.6174 | val\_loss=0.8383

Epoch 15: 100%|██████████| 205/205 [01:51<00:00, 1.84it/s, train\_loss=0.617, val\_loss=0.838]Metric val\_loss improved by 0.119 >= min\_delta = 0.01. New best score: 0.838

Epoch 16: 100%|██████████| 205/205 [01:51<00:00, 1.84it/s, train\_loss=0.521, val\_loss=0.734]📉 Epoch 17: train\_loss=0.5214 | val\_loss=0.7337

Epoch 16: 100%|██████████| 205/205 [01:51<00:00, 1.84it/s, train\_loss=0.521, val\_loss=0.734]Metric val\_loss improved by 0.105 >= min\_delta = 0.01. New best score: 0.734

Epoch 17: 100%|██████████| 205/205 [01:51<00:00, 1.84it/s, train\_loss=0.462, val\_loss=0.624]📉 Epoch 18: train\_loss=0.4620 | val\_loss=0.6237

Epoch 17: 100%|██████████| 205/205 [01:51<00:00, 1.84it/s, train\_loss=0.462, val\_loss=0.624]Metric val\_loss improved by 0.110 >= min\_delta = 0.01. New best score: 0.624

Epoch 18: 100%|██████████| 205/205 [01:51<00:00, 1.84it/s, train\_loss=0.472, val\_loss=0.612]📉 Epoch 19: train\_loss=0.4722 | val\_loss=0.6116

Epoch 18: 100%|██████████| 205/205 [01:51<00:00, 1.84it/s, train\_loss=0.472, val\_loss=0.612]Metric val\_loss improved by 0.012 >= min\_delta = 0.01. New best score: 0.612

Epoch 19: 100%|██████████| 205/205 [01:51<00:00, 1.84it/s, train\_loss=0.450, val\_loss=0.594]📉 Epoch 20: train\_loss=0.4503 | val\_loss=0.5939

Epoch 19: 100%|██████████| 205/205 [01:51<00:00, 1.84it/s, train\_loss=0.450, val\_loss=0.594]Metric val\_loss improved by 0.018 >= min\_delta = 0.01. New best score: 0.594

Epoch 20: 100%|██████████| 205/205 [01:51<00:00, 1.85it/s, train\_loss=0.451, val\_loss=0.608]📉 Epoch 21: train\_loss=0.4511 | val\_loss=0.6083

Epoch 21: 100%|██████████| 205/205 [01:51<00:00, 1.84it/s, train\_loss=0.462, val\_loss=0.570]📉 Epoch 22: train\_loss=0.4618 | val\_loss=0.5704

Epoch 21: 100%|██████████| 205/205 [01:51<00:00, 1.84it/s, train\_loss=0.462, val\_loss=0.570]Metric val\_loss improved by 0.023 >= min\_delta = 0.01. New best score: 0.570

Epoch 22: 100%|██████████| 205/205 [01:51<00:00, 1.84it/s, train\_loss=0.450, val\_loss=0.599]📉 Epoch 23: train\_loss=0.4503 | val\_loss=0.5988

Epoch 23: 100%|██████████| 205/205 [01:51<00:00, 1.85it/s, train\_loss=0.468, val\_loss=0.609]📉 Epoch 24: train\_loss=0.4681 | val\_loss=0.6094

Epoch 24: 100%|██████████| 205/205 [01:50<00:00, 1.85it/s, train\_loss=0.434, val\_loss=0.556]📉 Epoch 25: train\_loss=0.4335 | val\_loss=0.5555

Epoch 24: 100%|██████████| 205/205 [01:50<00:00, 1.85it/s, train\_loss=0.434, val\_loss=0.556]Metric val\_loss improved by 0.015 >= min\_delta = 0.01. New best score: 0.556

Epoch 25: 100%|██████████| 205/205 [01:50<00:00, 1.85it/s, train\_loss=0.391, val\_loss=0.552]📉 Epoch 26: train\_loss=0.3913 | val\_loss=0.5518

Epoch 26: 100%|██████████| 205/205 [01:51<00:00, 1.84it/s, train\_loss=0.428, val\_loss=0.536]📉 Epoch 27: train\_loss=0.4276 | val\_loss=0.5364

Epoch 26: 100%|██████████| 205/205 [01:51<00:00, 1.84it/s, train\_loss=0.428, val\_loss=0.536]Metric val\_loss improved by 0.019 >= min\_delta = 0.01. New best score: 0.536

Epoch 27: 100%|██████████| 205/205 [01:51<00:00, 1.84it/s, train\_loss=0.454, val\_loss=0.538]📉 Epoch 28: train\_loss=0.4544 | val\_loss=0.5384

Epoch 28: 100%|██████████| 205/205 [01:51<00:00, 1.84it/s, train\_loss=0.377, val\_loss=0.536]📉 Epoch 29: train\_loss=0.3771 | val\_loss=0.5355

Epoch 29: 100%|██████████| 205/205 [01:51<00:00, 1.85it/s, train\_loss=0.369, val\_loss=0.536]📉 Epoch 30: train\_loss=0.3694 | val\_loss=0.5364

Epoch 29: 100%|██████████| 205/205 [01:51<00:00, 1.85it/s, train\_loss=0.369, val\_loss=0.536]Monitored metric val\_loss did not improve in the last 3 records. Best score: 0.536. Signaling Trainer to stop.

`Trainer.fit` stopped: `max\_epochs=30` reached.

Epoch 29: 100%|██████████| 205/205 [01:51<00:00, 1.85it/s, train\_loss=0.369, val\_loss=0.536]

`predict()` was called with `n > output\_chunk\_length`: using auto-regression to forecast the values after `output\_chunk\_length` points. The model will access `(n - output\_chunk\_length)` future values of your `past\_covariates` (relative to the first predicted time step). To hide this warning, set `show\_warnings=False`.

GPU available: True (cuda), used: True

TPU available: False, using: 0 TPU cores

HPU available: False, using: 0 HPUs

LOCAL\_RANK: 0 - CUDA\_VISIBLE\_DEVICES: [0]

Predicting DataLoader 0: 100%|██████████| 1/1 [00:12<00:00, 0.08it/s][I 2025-05-11 01:52:24,321] Trial 34 finished with value: 376.7163391113281 and parameters: {'input\_chunk\_length': 48, 'output\_chunk\_length': 48, 'hidden\_size': 10, 'lstm\_layers': 2, 'num\_attention\_heads': 4, 'dropout': 0.2, 'n\_epochs': 30, 'lr': 0.0001, 'batch\_size': 128}. Best is trial 19 with value: 270.6866760253906.

🔍 DEBUG: Dataövergripande information:

▶️ train: 2021-01-02 00:00:00 → 2023-12-31 23:00:00 (26256 timsteg)

✅ val: 2024-01-01 00:00:00 → 2024-12-31 23:00:00 (8784 timsteg)

⏪ past\_cov\_train: 2021-01-02 00:00:00 → 2023-12-31 23:00:00

⏪ past\_cov\_val: 2023-12-28 00:00:00 → 2024-12-31 23:00:00

⏩ future\_cov\_train: 2021-01-02 00:00:00 → 2023-12-31 23:00:00

⏩ future\_cov\_val: 2023-12-28 00:00:00 → 2024-12-31 23:00:00

GPU available: True (cuda), used: True

TPU available: False, using: 0 TPU cores

HPU available: False, using: 0 HPUs

LOCAL\_RANK: 0 - CUDA\_VISIBLE\_DEVICES: [0]

| Name | Type | Params | Mode

------------------------------------------------------------------------------------------------

0 | train\_metrics | MetricCollection | 0 | train

1 | val\_metrics | MetricCollection | 0 | train

2 | input\_embeddings | \_MultiEmbedding | 0 | train

3 | static\_covariates\_vsn | \_VariableSelectionNetwork | 0 | train

4 | encoder\_vsn | \_VariableSelectionNetwork | 983 K | train

5 | decoder\_vsn | \_VariableSelectionNetwork | 439 K | train

6 | static\_context\_grn | \_GatedResidualNetwork | 231 K | train

7 | static\_context\_hidden\_encoder\_grn | \_GatedResidualNetwork | 231 K | train

8 | static\_context\_cell\_encoder\_grn | \_GatedResidualNetwork | 231 K | train

9 | static\_context\_enrichment | \_GatedResidualNetwork | 231 K | train

10 | lstm\_encoder | LSTM | 462 K | train

11 | lstm\_decoder | LSTM | 462 K | train

12 | post\_lstm\_gan | \_GateAddNorm | 116 K | train

13 | static\_enrichment\_grn | \_GatedResidualNetwork | 289 K | train

14 | multihead\_attn | \_InterpretableMultiHeadAttention | 144 K | train

15 | post\_attn\_gan | \_GateAddNorm | 116 K | train

16 | feed\_forward\_block | \_GatedResidualNetwork | 231 K | train

17 | pre\_output\_gan | \_GateAddNorm | 116 K | train

18 | output\_layer | Linear | 723 | train

------------------------------------------------------------------------------------------------

4.3 M Trainable params

0 Non-trainable params

4.3 M Total params

17.157 Total estimated model params size (MB)

3062 Modules in train mode

0 Modules in eval mode

🚀 Startar ny gridsearch-körning med hyperparametrar:

input\_chunk\_length: 96

output\_chunk\_length: 48

hidden\_size: 240

lstm\_layers: 1

num\_attention\_heads: 4

dropout: 0.4

optimizer\_kwargs: {'lr': 0.01}

Epoch 0: 100%|██████████| 205/205 [04:35<00:00, 0.74it/s, train\_loss=0.430, val\_loss=0.851]📉 Epoch 1: train\_loss=0.4298 | val\_loss=0.8513

Epoch 0: 100%|██████████| 205/205 [04:36<00:00, 0.74it/s, train\_loss=0.430, val\_loss=0.851]Metric val\_loss improved. New best score: 0.851

Epoch 1: 100%|██████████| 205/205 [04:35<00:00, 0.74it/s, train\_loss=1.420, val\_loss=1.060]📉 Epoch 2: train\_loss=1.4193 | val\_loss=1.0609

Epoch 2: 100%|██████████| 205/205 [04:36<00:00, 0.74it/s, train\_loss=0.119, val\_loss=0.677]📉 Epoch 3: train\_loss=0.1185 | val\_loss=0.6773

Epoch 2: 100%|██████████| 205/205 [04:36<00:00, 0.74it/s, train\_loss=0.119, val\_loss=0.677]Metric val\_loss improved by 0.174 >= min\_delta = 0.01. New best score: 0.677

Epoch 3: 100%|██████████| 205/205 [04:35<00:00, 0.74it/s, train\_loss=0.606, val\_loss=0.728]📉 Epoch 4: train\_loss=0.6063 | val\_loss=0.7281

Epoch 4: 100%|██████████| 205/205 [04:36<00:00, 0.74it/s, train\_loss=0.694, val\_loss=0.704]📉 Epoch 5: train\_loss=0.6936 | val\_loss=0.7044

Epoch 5: 100%|██████████| 205/205 [04:36<00:00, 0.74it/s, train\_loss=0.636, val\_loss=0.606]📉 Epoch 6: train\_loss=0.6362 | val\_loss=0.6063

Epoch 5: 100%|██████████| 205/205 [04:36<00:00, 0.74it/s, train\_loss=0.636, val\_loss=0.606]Metric val\_loss improved by 0.071 >= min\_delta = 0.01. New best score: 0.606

Epoch 6: 100%|██████████| 205/205 [04:36<00:00, 0.74it/s, train\_loss=0.222, val\_loss=0.499]📉 Epoch 7: train\_loss=0.2218 | val\_loss=0.4993

Epoch 6: 100%|██████████| 205/205 [04:36<00:00, 0.74it/s, train\_loss=0.222, val\_loss=0.499]Metric val\_loss improved by 0.107 >= min\_delta = 0.01. New best score: 0.499

Epoch 7: 100%|██████████| 205/205 [04:36<00:00, 0.74it/s, train\_loss=0.173, val\_loss=0.621]📉 Epoch 8: train\_loss=0.1726 | val\_loss=0.6210

Epoch 8: 100%|██████████| 205/205 [04:36<00:00, 0.74it/s, train\_loss=0.424, val\_loss=0.716]📉 Epoch 9: train\_loss=0.4238 | val\_loss=0.7160

Epoch 9: 100%|██████████| 205/205 [04:36<00:00, 0.74it/s, train\_loss=0.141, val\_loss=0.421]📉 Epoch 10: train\_loss=0.1411 | val\_loss=0.4210

Epoch 9: 100%|██████████| 205/205 [04:36<00:00, 0.74it/s, train\_loss=0.141, val\_loss=0.421]Metric val\_loss improved by 0.078 >= min\_delta = 0.01. New best score: 0.421

Epoch 10: 100%|██████████| 205/205 [04:36<00:00, 0.74it/s, train\_loss=0.0644, val\_loss=0.582]📉 Epoch 11: train\_loss=0.0644 | val\_loss=0.5823

Epoch 11: 100%|██████████| 205/205 [04:35<00:00, 0.74it/s, train\_loss=0.227, val\_loss=0.610] 📉 Epoch 12: train\_loss=0.2274 | val\_loss=0.6101

Epoch 12: 100%|██████████| 205/205 [04:35<00:00, 0.74it/s, train\_loss=0.059, val\_loss=0.658]📉 Epoch 13: train\_loss=0.0590 | val\_loss=0.6578

Epoch 12: 100%|██████████| 205/205 [04:35<00:00, 0.74it/s, train\_loss=0.059, val\_loss=0.658]Monitored metric val\_loss did not improve in the last 3 records. Best score: 0.421. Signaling Trainer to stop.

Epoch 12: 100%|██████████| 205/205 [04:35<00:00, 0.74it/s, train\_loss=0.059, val\_loss=0.658]

`predict()` was called with `n > output\_chunk\_length`: using auto-regression to forecast the values after `output\_chunk\_length` points. The model will access `(n - output\_chunk\_length)` future values of your `past\_covariates` (relative to the first predicted time step). To hide this warning, set `show\_warnings=False`.

GPU available: True (cuda), used: True

TPU available: False, using: 0 TPU cores

HPU available: False, using: 0 HPUs

LOCAL\_RANK: 0 - CUDA\_VISIBLE\_DEVICES: [0]

Predicting DataLoader 0: 100%|██████████| 1/1 [00:12<00:00, 0.08it/s][I 2025-05-11 02:52:27,955] Trial 35 finished with value: 367.1859436035156 and parameters: {'input\_chunk\_length': 96, 'output\_chunk\_length': 48, 'hidden\_size': 240, 'lstm\_layers': 1, 'num\_attention\_heads': 4, 'dropout': 0.4, 'n\_epochs': 30, 'lr': 0.01, 'batch\_size': 128}. Best is trial 19 with value: 270.6866760253906.

🔍 DEBUG: Dataövergripande information:

▶️ train: 2021-01-02 00:00:00 → 2023-12-31 23:00:00 (26256 timsteg)

✅ val: 2024-01-01 00:00:00 → 2024-12-31 23:00:00 (8784 timsteg)

⏪ past\_cov\_train: 2021-01-02 00:00:00 → 2023-12-31 23:00:00

⏪ past\_cov\_val: 2023-12-28 00:00:00 → 2024-12-31 23:00:00

⏩ future\_cov\_train: 2021-01-02 00:00:00 → 2023-12-31 23:00:00

⏩ future\_cov\_val: 2023-12-28 00:00:00 → 2024-12-31 23:00:00

GPU available: True (cuda), used: True

TPU available: False, using: 0 TPU cores

HPU available: False, using: 0 HPUs

LOCAL\_RANK: 0 - CUDA\_VISIBLE\_DEVICES: [0]

| Name | Type | Params | Mode

------------------------------------------------------------------------------------------------

0 | train\_metrics | MetricCollection | 0 | train

1 | val\_metrics | MetricCollection | 0 | train

2 | input\_embeddings | \_MultiEmbedding | 0 | train

3 | static\_covariates\_vsn | \_VariableSelectionNetwork | 0 | train

4 | encoder\_vsn | \_VariableSelectionNetwork | 110 K | train

5 | decoder\_vsn | \_VariableSelectionNetwork | 55.4 K | train

6 | static\_context\_grn | \_GatedResidualNetwork | 1.7 K | train

7 | static\_context\_hidden\_encoder\_grn | \_GatedResidualNetwork | 1.7 K | train

8 | static\_context\_cell\_encoder\_grn | \_GatedResidualNetwork | 1.7 K | train

9 | static\_context\_enrichment | \_GatedResidualNetwork | 1.7 K | train

10 | lstm\_encoder | LSTM | 6.7 K | train

11 | lstm\_decoder | LSTM | 6.7 K | train

12 | post\_lstm\_gan | \_GateAddNorm | 880 | train

13 | static\_enrichment\_grn | \_GatedResidualNetwork | 2.1 K | train

14 | multihead\_attn | \_InterpretableMultiHeadAttention | 1.7 K | train

15 | post\_attn\_gan | \_GateAddNorm | 880 | train

16 | feed\_forward\_block | \_GatedResidualNetwork | 1.7 K | train

17 | pre\_output\_gan | \_GateAddNorm | 880 | train

18 | output\_layer | Linear | 63 | train

------------------------------------------------------------------------------------------------

193 K Trainable params

0 Non-trainable params

193 K Total params

0.775 Total estimated model params size (MB)

3056 Modules in train mode

0 Modules in eval mode

🚀 Startar ny gridsearch-körning med hyperparametrar:

input\_chunk\_length: 48

output\_chunk\_length: 48

hidden\_size: 20

lstm\_layers: 2

num\_attention\_heads: 1

dropout: 0.5

optimizer\_kwargs: {'lr': 0.0001}

Epoch 0: 100%|██████████| 205/205 [01:48<00:00, 1.89it/s, train\_loss=5.600, val\_loss=5.740]📉 Epoch 1: train\_loss=5.5996 | val\_loss=5.7369

Epoch 0: 100%|██████████| 205/205 [01:48<00:00, 1.89it/s, train\_loss=5.600, val\_loss=5.740]Metric val\_loss improved. New best score: 5.737

Epoch 1: 100%|██████████| 205/205 [01:48<00:00, 1.89it/s, train\_loss=5.080, val\_loss=4.910]📉 Epoch 2: train\_loss=5.0814 | val\_loss=4.9072

Epoch 1: 100%|██████████| 205/205 [01:48<00:00, 1.89it/s, train\_loss=5.080, val\_loss=4.910]Metric val\_loss improved by 0.830 >= min\_delta = 0.01. New best score: 4.907

Epoch 2: 100%|██████████| 205/205 [01:48<00:00, 1.89it/s, train\_loss=3.730, val\_loss=4.080]📉 Epoch 3: train\_loss=3.7276 | val\_loss=4.0754

Epoch 2: 100%|██████████| 205/205 [01:48<00:00, 1.89it/s, train\_loss=3.730, val\_loss=4.080]Metric val\_loss improved by 0.832 >= min\_delta = 0.01. New best score: 4.075

Epoch 3: 100%|██████████| 205/205 [01:48<00:00, 1.89it/s, train\_loss=2.970, val\_loss=3.230]📉 Epoch 4: train\_loss=2.9690 | val\_loss=3.2327

Epoch 3: 100%|██████████| 205/205 [01:48<00:00, 1.89it/s, train\_loss=2.970, val\_loss=3.230]Metric val\_loss improved by 0.843 >= min\_delta = 0.01. New best score: 3.233

Epoch 4: 100%|██████████| 205/205 [01:48<00:00, 1.89it/s, train\_loss=2.110, val\_loss=2.440]📉 Epoch 5: train\_loss=2.1061 | val\_loss=2.4423

Epoch 4: 100%|██████████| 205/205 [01:48<00:00, 1.89it/s, train\_loss=2.110, val\_loss=2.440]Metric val\_loss improved by 0.790 >= min\_delta = 0.01. New best score: 2.442

Epoch 5: 100%|██████████| 205/205 [01:48<00:00, 1.89it/s, train\_loss=1.630, val\_loss=1.790]📉 Epoch 6: train\_loss=1.6286 | val\_loss=1.7892

Epoch 5: 100%|██████████| 205/205 [01:48<00:00, 1.89it/s, train\_loss=1.630, val\_loss=1.790]Metric val\_loss improved by 0.653 >= min\_delta = 0.01. New best score: 1.789

Epoch 6: 100%|██████████| 205/205 [01:48<00:00, 1.89it/s, train\_loss=1.220, val\_loss=1.330]📉 Epoch 7: train\_loss=1.2221 | val\_loss=1.3321

Epoch 6: 100%|██████████| 205/205 [01:48<00:00, 1.89it/s, train\_loss=1.220, val\_loss=1.330]Metric val\_loss improved by 0.457 >= min\_delta = 0.01. New best score: 1.332

Epoch 7: 100%|██████████| 205/205 [01:48<00:00, 1.89it/s, train\_loss=0.941, val\_loss=1.060]📉 Epoch 8: train\_loss=0.9411 | val\_loss=1.0609

Epoch 7: 100%|██████████| 205/205 [01:48<00:00, 1.89it/s, train\_loss=0.941, val\_loss=1.060]Metric val\_loss improved by 0.271 >= min\_delta = 0.01. New best score: 1.061

Epoch 8: 100%|██████████| 205/205 [01:48<00:00, 1.88it/s, train\_loss=0.785, val\_loss=0.915]📉 Epoch 9: train\_loss=0.7846 | val\_loss=0.9151

Epoch 8: 100%|██████████| 205/205 [01:48<00:00, 1.88it/s, train\_loss=0.785, val\_loss=0.915]Metric val\_loss improved by 0.146 >= min\_delta = 0.01. New best score: 0.915

Epoch 9: 100%|██████████| 205/205 [01:48<00:00, 1.89it/s, train\_loss=0.741, val\_loss=0.812]📉 Epoch 10: train\_loss=0.7406 | val\_loss=0.8124

Epoch 9: 100%|██████████| 205/205 [01:48<00:00, 1.89it/s, train\_loss=0.741, val\_loss=0.812]Metric val\_loss improved by 0.103 >= min\_delta = 0.01. New best score: 0.812

Epoch 10: 100%|██████████| 205/205 [01:48<00:00, 1.89it/s, train\_loss=0.714, val\_loss=0.733]📉 Epoch 11: train\_loss=0.7137 | val\_loss=0.7329

Epoch 10: 100%|██████████| 205/205 [01:48<00:00, 1.89it/s, train\_loss=0.714, val\_loss=0.733]Metric val\_loss improved by 0.080 >= min\_delta = 0.01. New best score: 0.733

Epoch 11: 100%|██████████| 205/205 [01:48<00:00, 1.89it/s, train\_loss=0.548, val\_loss=0.724]📉 Epoch 12: train\_loss=0.5477 | val\_loss=0.7237

Epoch 12: 100%|██████████| 205/205 [01:48<00:00, 1.89it/s, train\_loss=0.529, val\_loss=0.606]📉 Epoch 13: train\_loss=0.5285 | val\_loss=0.6057

Epoch 12: 100%|██████████| 205/205 [01:48<00:00, 1.89it/s, train\_loss=0.529, val\_loss=0.606]Metric val\_loss improved by 0.127 >= min\_delta = 0.01. New best score: 0.606

Epoch 13: 100%|██████████| 205/205 [01:48<00:00, 1.89it/s, train\_loss=0.502, val\_loss=0.631]📉 Epoch 14: train\_loss=0.5016 | val\_loss=0.6308

Epoch 14: 100%|██████████| 205/205 [01:48<00:00, 1.89it/s, train\_loss=0.489, val\_loss=0.575]📉 Epoch 15: train\_loss=0.4889 | val\_loss=0.5747

Epoch 14: 100%|██████████| 205/205 [01:48<00:00, 1.89it/s, train\_loss=0.489, val\_loss=0.575]Metric val\_loss improved by 0.031 >= min\_delta = 0.01. New best score: 0.575

Epoch 15: 100%|██████████| 205/205 [01:48<00:00, 1.89it/s, train\_loss=0.464, val\_loss=0.551]📉 Epoch 16: train\_loss=0.4638 | val\_loss=0.5508

Epoch 15: 100%|██████████| 205/205 [01:48<00:00, 1.89it/s, train\_loss=0.464, val\_loss=0.551]Metric val\_loss improved by 0.024 >= min\_delta = 0.01. New best score: 0.551

Epoch 16: 100%|██████████| 205/205 [01:48<00:00, 1.88it/s, train\_loss=0.436, val\_loss=0.540]📉 Epoch 17: train\_loss=0.4357 | val\_loss=0.5401

Epoch 16: 100%|██████████| 205/205 [01:48<00:00, 1.88it/s, train\_loss=0.436, val\_loss=0.540]Metric val\_loss improved by 0.011 >= min\_delta = 0.01. New best score: 0.540

Epoch 17: 100%|██████████| 205/205 [01:48<00:00, 1.89it/s, train\_loss=0.426, val\_loss=0.533]📉 Epoch 18: train\_loss=0.4260 | val\_loss=0.5325

Epoch 18: 100%|██████████| 205/205 [01:48<00:00, 1.88it/s, train\_loss=0.384, val\_loss=0.496]📉 Epoch 19: train\_loss=0.3843 | val\_loss=0.4960

Epoch 18: 100%|██████████| 205/205 [01:48<00:00, 1.88it/s, train\_loss=0.384, val\_loss=0.496]Metric val\_loss improved by 0.044 >= min\_delta = 0.01. New best score: 0.496

Epoch 19: 100%|██████████| 205/205 [01:48<00:00, 1.89it/s, train\_loss=0.430, val\_loss=0.478]📉 Epoch 20: train\_loss=0.4298 | val\_loss=0.4783

Epoch 19: 100%|██████████| 205/205 [01:48<00:00, 1.89it/s, train\_loss=0.430, val\_loss=0.478]Metric val\_loss improved by 0.018 >= min\_delta = 0.01. New best score: 0.478

Epoch 20: 100%|██████████| 205/205 [01:48<00:00, 1.89it/s, train\_loss=0.395, val\_loss=0.491]📉 Epoch 21: train\_loss=0.3949 | val\_loss=0.4912

Epoch 21: 100%|██████████| 205/205 [01:48<00:00, 1.88it/s, train\_loss=0.455, val\_loss=0.472]📉 Epoch 22: train\_loss=0.4547 | val\_loss=0.4720

Epoch 22: 100%|██████████| 205/205 [01:48<00:00, 1.88it/s, train\_loss=0.441, val\_loss=0.451]📉 Epoch 23: train\_loss=0.4410 | val\_loss=0.4511

Epoch 22: 100%|██████████| 205/205 [01:48<00:00, 1.88it/s, train\_loss=0.441, val\_loss=0.451]Metric val\_loss improved by 0.027 >= min\_delta = 0.01. New best score: 0.451

Epoch 23: 100%|██████████| 205/205 [01:48<00:00, 1.89it/s, train\_loss=0.437, val\_loss=0.447]📉 Epoch 24: train\_loss=0.4365 | val\_loss=0.4469

Epoch 24: 100%|██████████| 205/205 [01:48<00:00, 1.89it/s, train\_loss=0.410, val\_loss=0.461]📉 Epoch 25: train\_loss=0.4101 | val\_loss=0.4605

Epoch 25: 100%|██████████| 205/205 [01:48<00:00, 1.89it/s, train\_loss=0.424, val\_loss=0.460]📉 Epoch 26: train\_loss=0.4241 | val\_loss=0.4599

Epoch 25: 100%|██████████| 205/205 [01:48<00:00, 1.89it/s, train\_loss=0.424, val\_loss=0.460]Monitored metric val\_loss did not improve in the last 3 records. Best score: 0.451. Signaling Trainer to stop.

Epoch 25: 100%|██████████| 205/205 [01:48<00:00, 1.89it/s, train\_loss=0.424, val\_loss=0.460]

`predict()` was called with `n > output\_chunk\_length`: using auto-regression to forecast the values after `output\_chunk\_length` points. The model will access `(n - output\_chunk\_length)` future values of your `past\_covariates` (relative to the first predicted time step). To hide this warning, set `show\_warnings=False`.

GPU available: True (cuda), used: True

TPU available: False, using: 0 TPU cores

HPU available: False, using: 0 HPUs

LOCAL\_RANK: 0 - CUDA\_VISIBLE\_DEVICES: [0]

Predicting DataLoader 0: 100%|██████████| 1/1 [00:11<00:00, 0.08it/s][I 2025-05-11 03:39:44,577] Trial 36 finished with value: 358.80511474609375 and parameters: {'input\_chunk\_length': 48, 'output\_chunk\_length': 48, 'hidden\_size': 20, 'lstm\_layers': 2, 'num\_attention\_heads': 1, 'dropout': 0.5, 'n\_epochs': 30, 'lr': 0.0001, 'batch\_size': 128}. Best is trial 19 with value: 270.6866760253906.

🔍 DEBUG: Dataövergripande information:

▶️ train: 2021-01-02 00:00:00 → 2023-12-31 23:00:00 (26256 timsteg)

✅ val: 2024-01-01 00:00:00 → 2024-12-31 23:00:00 (8784 timsteg)

⏪ past\_cov\_train: 2021-01-02 00:00:00 → 2023-12-31 23:00:00

⏪ past\_cov\_val: 2023-12-28 00:00:00 → 2024-12-31 23:00:00

⏩ future\_cov\_train: 2021-01-02 00:00:00 → 2023-12-31 23:00:00

⏩ future\_cov\_val: 2023-12-28 00:00:00 → 2024-12-31 23:00:00

GPU available: True (cuda), used: True

TPU available: False, using: 0 TPU cores

HPU available: False, using: 0 HPUs

LOCAL\_RANK: 0 - CUDA\_VISIBLE\_DEVICES: [0]

| Name | Type | Params | Mode

------------------------------------------------------------------------------------------------

0 | train\_metrics | MetricCollection | 0 | train

1 | val\_metrics | MetricCollection | 0 | train

2 | input\_embeddings | \_MultiEmbedding | 0 | train

3 | static\_covariates\_vsn | \_VariableSelectionNetwork | 0 | train

4 | encoder\_vsn | \_VariableSelectionNetwork | 727 K | train

5 | decoder\_vsn | \_VariableSelectionNetwork | 312 K | train

6 | static\_context\_grn | \_GatedResidualNetwork | 103 K | train

7 | static\_context\_hidden\_encoder\_grn | \_GatedResidualNetwork | 103 K | train

8 | static\_context\_cell\_encoder\_grn | \_GatedResidualNetwork | 103 K | train

9 | static\_context\_enrichment | \_GatedResidualNetwork | 103 K | train

10 | lstm\_encoder | LSTM | 412 K | train

11 | lstm\_decoder | LSTM | 412 K | train

12 | post\_lstm\_gan | \_GateAddNorm | 51.8 K | train

13 | static\_enrichment\_grn | \_GatedResidualNetwork | 128 K | train

14 | multihead\_attn | \_InterpretableMultiHeadAttention | 102 K | train

15 | post\_attn\_gan | \_GateAddNorm | 51.8 K | train

16 | feed\_forward\_block | \_GatedResidualNetwork | 103 K | train

17 | pre\_output\_gan | \_GateAddNorm | 51.8 K | train

18 | output\_layer | Linear | 483 | train

------------------------------------------------------------------------------------------------

2.8 M Trainable params

0 Non-trainable params

2.8 M Total params

11.073 Total estimated model params size (MB)

3056 Modules in train mode

0 Modules in eval mode

🚀 Startar ny gridsearch-körning med hyperparametrar:

input\_chunk\_length: 96

output\_chunk\_length: 48

hidden\_size: 160

lstm\_layers: 2

num\_attention\_heads: 1

dropout: 0.2

optimizer\_kwargs: {'lr': 0.01}

Epoch 0: 100%|██████████| 205/205 [03:50<00:00, 0.89it/s, train\_loss=0.554, val\_loss=0.702]📉 Epoch 1: train\_loss=0.5545 | val\_loss=0.7017

Epoch 0: 100%|██████████| 205/205 [03:50<00:00, 0.89it/s, train\_loss=0.554, val\_loss=0.702]Metric val\_loss improved. New best score: 0.702

Epoch 1: 100%|██████████| 205/205 [03:49<00:00, 0.89it/s, train\_loss=0.699, val\_loss=0.924]📉 Epoch 2: train\_loss=0.6992 | val\_loss=0.9238

Epoch 2: 100%|██████████| 205/205 [03:49<00:00, 0.89it/s, train\_loss=0.408, val\_loss=0.668]📉 Epoch 3: train\_loss=0.4083 | val\_loss=0.6678

Epoch 2: 100%|██████████| 205/205 [03:49<00:00, 0.89it/s, train\_loss=0.408, val\_loss=0.668]Metric val\_loss improved by 0.034 >= min\_delta = 0.01. New best score: 0.668

Epoch 3: 100%|██████████| 205/205 [03:49<00:00, 0.89it/s, train\_loss=0.245, val\_loss=0.673]📉 Epoch 4: train\_loss=0.2451 | val\_loss=0.6730

Epoch 4: 100%|██████████| 205/205 [03:49<00:00, 0.89it/s, train\_loss=0.361, val\_loss=0.701]📉 Epoch 5: train\_loss=0.3608 | val\_loss=0.7006

Epoch 5: 100%|██████████| 205/205 [03:49<00:00, 0.89it/s, train\_loss=0.592, val\_loss=0.704]📉 Epoch 6: train\_loss=0.5923 | val\_loss=0.7038

Epoch 5: 100%|██████████| 205/205 [03:49<00:00, 0.89it/s, train\_loss=0.592, val\_loss=0.704]Monitored metric val\_loss did not improve in the last 3 records. Best score: 0.668. Signaling Trainer to stop.

Epoch 5: 100%|██████████| 205/205 [03:49<00:00, 0.89it/s, train\_loss=0.592, val\_loss=0.704]

`predict()` was called with `n > output\_chunk\_length`: using auto-regression to forecast the values after `output\_chunk\_length` points. The model will access `(n - output\_chunk\_length)` future values of your `past\_covariates` (relative to the first predicted time step). To hide this warning, set `show\_warnings=False`.

GPU available: True (cuda), used: True

TPU available: False, using: 0 TPU cores

HPU available: False, using: 0 HPUs

LOCAL\_RANK: 0 - CUDA\_VISIBLE\_DEVICES: [0]

Predicting DataLoader 0: 100%|██████████| 1/1 [00:11<00:00, 0.08it/s][I 2025-05-11 04:02:56,156] Trial 37 finished with value: 566.0189819335938 and parameters: {'input\_chunk\_length': 96, 'output\_chunk\_length': 48, 'hidden\_size': 160, 'lstm\_layers': 2, 'num\_attention\_heads': 1, 'dropout': 0.2, 'n\_epochs': 30, 'lr': 0.01, 'batch\_size': 128}. Best is trial 19 with value: 270.6866760253906.

🔍 DEBUG: Dataövergripande information:

▶️ train: 2021-01-02 00:00:00 → 2023-12-31 23:00:00 (26256 timsteg)

✅ val: 2024-01-01 00:00:00 → 2024-12-31 23:00:00 (8784 timsteg)

⏪ past\_cov\_train: 2021-01-02 00:00:00 → 2023-12-31 23:00:00

⏪ past\_cov\_val: 2023-12-28 00:00:00 → 2024-12-31 23:00:00

⏩ future\_cov\_train: 2021-01-02 00:00:00 → 2023-12-31 23:00:00

⏩ future\_cov\_val: 2023-12-28 00:00:00 → 2024-12-31 23:00:00

GPU available: True (cuda), used: True

TPU available: False, using: 0 TPU cores

HPU available: False, using: 0 HPUs

LOCAL\_RANK: 0 - CUDA\_VISIBLE\_DEVICES: [0]

| Name | Type | Params | Mode

------------------------------------------------------------------------------------------------

0 | train\_metrics | MetricCollection | 0 | train

1 | val\_metrics | MetricCollection | 0 | train

2 | input\_embeddings | \_MultiEmbedding | 0 | train

3 | static\_covariates\_vsn | \_VariableSelectionNetwork | 0 | train

4 | encoder\_vsn | \_VariableSelectionNetwork | 201 K | train

5 | decoder\_vsn | \_VariableSelectionNetwork | 101 K | train

6 | static\_context\_grn | \_GatedResidualNetwork | 6.6 K | train

7 | static\_context\_hidden\_encoder\_grn | \_GatedResidualNetwork | 6.6 K | train

8 | static\_context\_cell\_encoder\_grn | \_GatedResidualNetwork | 6.6 K | train

9 | static\_context\_enrichment | \_GatedResidualNetwork | 6.6 K | train

10 | lstm\_encoder | LSTM | 39.4 K | train

11 | lstm\_decoder | LSTM | 39.4 K | train

12 | post\_lstm\_gan | \_GateAddNorm | 3.4 K | train

13 | static\_enrichment\_grn | \_GatedResidualNetwork | 8.2 K | train

14 | multihead\_attn | \_InterpretableMultiHeadAttention | 4.1 K | train

15 | post\_attn\_gan | \_GateAddNorm | 3.4 K | train

16 | feed\_forward\_block | \_GatedResidualNetwork | 6.6 K | train

17 | pre\_output\_gan | \_GateAddNorm | 3.4 K | train

18 | output\_layer | Linear | 123 | train

------------------------------------------------------------------------------------------------

435 K Trainable params

0 Non-trainable params

435 K Total params

1.743 Total estimated model params size (MB)

3062 Modules in train mode

0 Modules in eval mode

🚀 Startar ny gridsearch-körning med hyperparametrar:

input\_chunk\_length: 72

output\_chunk\_length: 48

hidden\_size: 40

lstm\_layers: 3

num\_attention\_heads: 4

dropout: 0.1

optimizer\_kwargs: {'lr': 0.01}

Epoch 0: 100%|██████████| 205/205 [02:23<00:00, 1.42it/s, train\_loss=0.491, val\_loss=0.552]📉 Epoch 1: train\_loss=0.4907 | val\_loss=0.5517

Epoch 0: 100%|██████████| 205/205 [02:23<00:00, 1.42it/s, train\_loss=0.491, val\_loss=0.552]Metric val\_loss improved. New best score: 0.552

Epoch 1: 100%|██████████| 205/205 [02:24<00:00, 1.42it/s, train\_loss=0.368, val\_loss=0.519]📉 Epoch 2: train\_loss=0.3681 | val\_loss=0.5185

Epoch 1: 100%|██████████| 205/205 [02:24<00:00, 1.42it/s, train\_loss=0.368, val\_loss=0.519]Metric val\_loss improved by 0.033 >= min\_delta = 0.01. New best score: 0.519

Epoch 2: 100%|██████████| 205/205 [02:23<00:00, 1.42it/s, train\_loss=0.469, val\_loss=0.538]📉 Epoch 3: train\_loss=0.4694 | val\_loss=0.5380

Epoch 3: 100%|██████████| 205/205 [02:23<00:00, 1.42it/s, train\_loss=0.286, val\_loss=0.490]📉 Epoch 4: train\_loss=0.2858 | val\_loss=0.4901

Epoch 3: 100%|██████████| 205/205 [02:23<00:00, 1.42it/s, train\_loss=0.286, val\_loss=0.490]Metric val\_loss improved by 0.028 >= min\_delta = 0.01. New best score: 0.490

Epoch 4: 100%|██████████| 205/205 [02:23<00:00, 1.43it/s, train\_loss=0.227, val\_loss=0.441]📉 Epoch 5: train\_loss=0.2266 | val\_loss=0.4409

Epoch 4: 100%|██████████| 205/205 [02:23<00:00, 1.43it/s, train\_loss=0.227, val\_loss=0.441]Metric val\_loss improved by 0.049 >= min\_delta = 0.01. New best score: 0.441

Epoch 5: 100%|██████████| 205/205 [02:23<00:00, 1.42it/s, train\_loss=0.276, val\_loss=0.479]📉 Epoch 6: train\_loss=0.2758 | val\_loss=0.4790

Epoch 6: 100%|██████████| 205/205 [02:24<00:00, 1.42it/s, train\_loss=0.199, val\_loss=0.574]📉 Epoch 7: train\_loss=0.1988 | val\_loss=0.5735

Epoch 7: 100%|██████████| 205/205 [02:23<00:00, 1.43it/s, train\_loss=0.195, val\_loss=0.546]📉 Epoch 8: train\_loss=0.1954 | val\_loss=0.5461

Epoch 7: 100%|██████████| 205/205 [02:23<00:00, 1.43it/s, train\_loss=0.195, val\_loss=0.546]Monitored metric val\_loss did not improve in the last 3 records. Best score: 0.441. Signaling Trainer to stop.

Epoch 7: 100%|██████████| 205/205 [02:23<00:00, 1.43it/s, train\_loss=0.195, val\_loss=0.546]

`predict()` was called with `n > output\_chunk\_length`: using auto-regression to forecast the values after `output\_chunk\_length` points. The model will access `(n - output\_chunk\_length)` future values of your `past\_covariates` (relative to the first predicted time step). To hide this warning, set `show\_warnings=False`.

GPU available: True (cuda), used: True

TPU available: False, using: 0 TPU cores

HPU available: False, using: 0 HPUs

LOCAL\_RANK: 0 - CUDA\_VISIBLE\_DEVICES: [0]

Predicting DataLoader 0: 100%|██████████| 1/1 [00:11<00:00, 0.08it/s][I 2025-05-11 04:22:21,333] Trial 38 finished with value: 310.32708740234375 and parameters: {'input\_chunk\_length': 72, 'output\_chunk\_length': 48, 'hidden\_size': 40, 'lstm\_layers': 3, 'num\_attention\_heads': 4, 'dropout': 0.1, 'n\_epochs': 30, 'lr': 0.01, 'batch\_size': 128}. Best is trial 19 with value: 270.6866760253906.

🔍 DEBUG: Dataövergripande information:

▶️ train: 2021-01-02 00:00:00 → 2023-12-31 23:00:00 (26256 timsteg)

✅ val: 2024-01-01 00:00:00 → 2024-12-31 23:00:00 (8784 timsteg)

⏪ past\_cov\_train: 2021-01-02 00:00:00 → 2023-12-31 23:00:00

⏪ past\_cov\_val: 2023-12-28 00:00:00 → 2024-12-31 23:00:00

⏩ future\_cov\_train: 2021-01-02 00:00:00 → 2023-12-31 23:00:00

⏩ future\_cov\_val: 2023-12-28 00:00:00 → 2024-12-31 23:00:00

GPU available: True (cuda), used: True

TPU available: False, using: 0 TPU cores

HPU available: False, using: 0 HPUs

LOCAL\_RANK: 0 - CUDA\_VISIBLE\_DEVICES: [0]

| Name | Type | Params | Mode

------------------------------------------------------------------------------------------------

0 | train\_metrics | MetricCollection | 0 | train

1 | val\_metrics | MetricCollection | 0 | train

2 | input\_embeddings | \_MultiEmbedding | 0 | train

3 | static\_covariates\_vsn | \_VariableSelectionNetwork | 0 | train

4 | encoder\_vsn | \_VariableSelectionNetwork | 1.2 M | train

5 | decoder\_vsn | \_VariableSelectionNetwork | 565 K | train

6 | static\_context\_grn | \_GatedResidualNetwork | 411 K | train

7 | static\_context\_hidden\_encoder\_grn | \_GatedResidualNetwork | 411 K | train

8 | static\_context\_cell\_encoder\_grn | \_GatedResidualNetwork | 411 K | train

9 | static\_context\_enrichment | \_GatedResidualNetwork | 411 K | train

10 | lstm\_encoder | LSTM | 1.6 M | train

11 | lstm\_decoder | LSTM | 1.6 M | train

12 | post\_lstm\_gan | \_GateAddNorm | 206 K | train

13 | static\_enrichment\_grn | \_GatedResidualNetwork | 513 K | train

14 | multihead\_attn | \_InterpretableMultiHeadAttention | 410 K | train

15 | post\_attn\_gan | \_GateAddNorm | 206 K | train

16 | feed\_forward\_block | \_GatedResidualNetwork | 411 K | train

17 | pre\_output\_gan | \_GateAddNorm | 206 K | train

18 | output\_layer | Linear | 963 | train

------------------------------------------------------------------------------------------------

8.7 M Trainable params

0 Non-trainable params

8.7 M Total params

34.767 Total estimated model params size (MB)

3056 Modules in train mode

0 Modules in eval mode

🚀 Startar ny gridsearch-körning med hyperparametrar:

input\_chunk\_length: 96

output\_chunk\_length: 48

hidden\_size: 320

lstm\_layers: 2

num\_attention\_heads: 1

dropout: 0.4

optimizer\_kwargs: {'lr': 0.0001}

Epoch 0: 100%|██████████| 205/205 [05:08<00:00, 0.66it/s, train\_loss=0.566, val\_loss=0.664]📉 Epoch 1: train\_loss=0.5662 | val\_loss=0.6642

Epoch 0: 100%|██████████| 205/205 [05:08<00:00, 0.66it/s, train\_loss=0.566, val\_loss=0.664]Metric val\_loss improved. New best score: 0.664

Epoch 1: 100%|██████████| 205/205 [05:08<00:00, 0.66it/s, train\_loss=0.490, val\_loss=0.529]📉 Epoch 2: train\_loss=0.4899 | val\_loss=0.5289

Epoch 1: 100%|██████████| 205/205 [05:08<00:00, 0.66it/s, train\_loss=0.490, val\_loss=0.529]Metric val\_loss improved by 0.135 >= min\_delta = 0.01. New best score: 0.529

Epoch 2: 100%|██████████| 205/205 [05:08<00:00, 0.66it/s, train\_loss=0.272, val\_loss=0.457]📉 Epoch 3: train\_loss=0.2718 | val\_loss=0.4566

Epoch 2: 100%|██████████| 205/205 [05:08<00:00, 0.66it/s, train\_loss=0.272, val\_loss=0.457]Metric val\_loss improved by 0.072 >= min\_delta = 0.01. New best score: 0.457

Epoch 3: 100%|██████████| 205/205 [05:08<00:00, 0.66it/s, train\_loss=0.278, val\_loss=0.460]📉 Epoch 4: train\_loss=0.2778 | val\_loss=0.4598

Epoch 4: 100%|██████████| 205/205 [05:08<00:00, 0.66it/s, train\_loss=0.130, val\_loss=0.408]📉 Epoch 5: train\_loss=0.1301 | val\_loss=0.4083

Epoch 4: 100%|██████████| 205/205 [05:08<00:00, 0.66it/s, train\_loss=0.130, val\_loss=0.408]Metric val\_loss improved by 0.048 >= min\_delta = 0.01. New best score: 0.408

Epoch 5: 100%|██████████| 205/205 [05:08<00:00, 0.66it/s, train\_loss=0.130, val\_loss=0.390]📉 Epoch 6: train\_loss=0.1297 | val\_loss=0.3900

Epoch 5: 100%|██████████| 205/205 [05:08<00:00, 0.66it/s, train\_loss=0.130, val\_loss=0.390]Metric val\_loss improved by 0.018 >= min\_delta = 0.01. New best score: 0.390

Epoch 6: 100%|██████████| 205/205 [05:09<00:00, 0.66it/s, train\_loss=0.330, val\_loss=0.443]📉 Epoch 7: train\_loss=0.3304 | val\_loss=0.4429

Epoch 7: 100%|██████████| 205/205 [05:08<00:00, 0.66it/s, train\_loss=0.425, val\_loss=0.385]📉 Epoch 8: train\_loss=0.4250 | val\_loss=0.3849

Epoch 8: 100%|██████████| 205/205 [05:08<00:00, 0.66it/s, train\_loss=0.238, val\_loss=0.392]📉 Epoch 9: train\_loss=0.2380 | val\_loss=0.3920

Epoch 8: 100%|██████████| 205/205 [05:08<00:00, 0.66it/s, train\_loss=0.238, val\_loss=0.392]Monitored metric val\_loss did not improve in the last 3 records. Best score: 0.390. Signaling Trainer to stop.

Epoch 8: 100%|██████████| 205/205 [05:08<00:00, 0.66it/s, train\_loss=0.238, val\_loss=0.392]

`predict()` was called with `n > output\_chunk\_length`: using auto-regression to forecast the values after `output\_chunk\_length` points. The model will access `(n - output\_chunk\_length)` future values of your `past\_covariates` (relative to the first predicted time step). To hide this warning, set `show\_warnings=False`.

GPU available: True (cuda), used: True

TPU available: False, using: 0 TPU cores

HPU available: False, using: 0 HPUs

LOCAL\_RANK: 0 - CUDA\_VISIBLE\_DEVICES: [0]

Predicting DataLoader 0: 100%|██████████| 1/1 [00:12<00:00, 0.08it/s]

[I 2025-05-11 05:08:54,155] Trial 39 finished with value: 284.7579345703125 and parameters: {'input\_chunk\_length': 96, 'output\_chunk\_length': 48, 'hidden\_size': 320, 'lstm\_layers': 2, 'num\_attention\_heads': 1, 'dropout': 0.4, 'n\_epochs': 30, 'lr': 0.0001, 'batch\_size': 128}. Best is trial 19 with value: 270.6866760253906.

Best hyperparameters: {'input\_chunk\_length': 72, 'output\_chunk\_length': 48, 'hidden\_size': 40, 'lstm\_layers': 2, 'num\_attention\_heads': 1, 'dropout': 0.2, 'n\_epochs': 30, 'lr': 0.01, 'batch\_size': 128}

Best MAE score: 270.6866760253906

number value params\_batch\_size params\_dropout params\_hidden\_size \

19 19 270.686676 128 0.2 40

39 39 284.757935 128 0.4 320

0 0 284.924042 128 0.4 40

23 23 288.199402 128 0.5 40

28 28 290.278442 128 0.2 40

29 29 290.854095 128 0.5 20

27 27 293.229156 128 0.4 40

32 32 295.829803 128 0.2 40

9 9 297.210449 128 0.3 80

24 24 304.573456 128 0.5 40

15 15 308.047333 128 0.4 80

26 26 308.346100 128 0.5 40

11 11 309.736694 128 0.3 80

1 1 309.894257 128 0.4 20

38 38 310.327087 128 0.1 40

31 31 312.764893 128 0.2 40

14 14 312.985474 128 0.3 40

21 21 314.173096 128 0.2 40

8 8 314.412079 128 0.1 240

12 12 314.459564 128 0.3 80

13 13 318.041962 128 0.3 40

20 20 321.993286 128 0.2 40

30 30 331.356873 128 0.4 40

7 7 333.611694 128 0.2 160

16 16 342.369324 128 0.3 160

10 10 343.717072 128 0.5 40

18 18 343.953644 128 0.1 10

33 33 345.095764 128 0.2 40

22 22 349.303284 128 0.2 80

17 17 353.514130 128 0.4 320

36 36 358.805115 128 0.5 20

35 35 367.185944 128 0.4 240

34 34 376.716339 128 0.2 10

3 3 384.718140 128 0.5 240

4 4 399.214264 128 0.2 10

25 25 401.185791 128 0.5 40

2 2 402.206421 128 0.4 240

5 5 407.454193 128 0.1 20

6 6 562.400696 128 0.4 320

37 37 566.018982 128 0.2 160

params\_input\_chunk\_length params\_lr params\_lstm\_layers params\_n\_epochs \

19 72 0.0100 2 30

39 96 0.0001 2 30

0 72 0.0010 3 30

23 72 0.0100 2 30

28 48 0.0100 2 30

29 72 0.0100 1 30

27 72 0.0100 2 30

32 48 0.0100 2 30

9 72 0.0010 3 30

24 72 0.0100 2 30

15 72 0.0010 2 30

26 72 0.0100 2 30

11 72 0.0010 3 30

1 72 0.0100 1 30

38 72 0.0100 3 30

31 48 0.0100 2 30

14 72 0.0010 1 30

21 72 0.0100 2 30

8 72 0.0001 3 30

12 72 0.0010 3 30

13 48 0.0010 3 30

20 48 0.0100 2 30

30 72 0.0100 2 30

7 96 0.0010 2 30

16 72 0.0100 3 30

10 48 0.0010 2 30

18 72 0.0010 1 30

33 48 0.0100 2 30

22 72 0.0100 2 30

17 48 0.0010 3 30

36 48 0.0001 2 30

35 96 0.0100 1 30

34 48 0.0001 2 30

3 96 0.0001 3 30

4 72 0.0001 1 30

25 72 0.0100 2 30

2 96 0.0010 3 30

5 96 0.0001 3 30

6 96 0.0100 3 30

37 96 0.0100 2 30

params\_num\_attention\_heads params\_output\_chunk\_length

19 1 48

39 1 48

0 4 48

23 1 48

28 4 48

29 1 48

27 1 48

32 4 48

9 1 48

24 1 48

15 4 48

26 1 48

11 4 48

1 1 48

38 4 48

31 4 48

14 1 48

21 1 48

8 4 48

12 4 48

13 4 48

20 4 48

30 1 48

7 4 48

16 1 48

10 4 48

18 4 48

33 4 48

22 1 48

17 4 48

36 1 48

35 4 48

34 4 48

3 1 48

4 1 48

25 1 48

2 1 48

5 1 48

6 1 48

37 1 48

🔍 DEBUG: Dataövergripande information:

▶️ train: 2021-01-02 00:00:00 → 2023-12-31 23:00:00 (26256 timsteg)

❌ val: None

⏪ past\_cov\_train: 2021-01-02 00:00:00 → 2023-12-31 23:00:00

⏩ future\_cov\_train: 2021-01-02 00:00:00 → 2023-12-31 23:00:00

GPU available: True (cuda), used: True

TPU available: False, using: 0 TPU cores

HPU available: False, using: 0 HPUs

LOCAL\_RANK: 0 - CUDA\_VISIBLE\_DEVICES: [0]

| Name | Type | Params | Mode

------------------------------------------------------------------------------------------------

0 | train\_metrics | MetricCollection | 0 | train

1 | val\_metrics | MetricCollection | 0 | train

2 | input\_embeddings | \_MultiEmbedding | 0 | train

3 | static\_covariates\_vsn | \_VariableSelectionNetwork | 0 | train

4 | encoder\_vsn | \_VariableSelectionNetwork | 201 K | train

5 | decoder\_vsn | \_VariableSelectionNetwork | 101 K | train

6 | static\_context\_grn | \_GatedResidualNetwork | 6.6 K | train

7 | static\_context\_hidden\_encoder\_grn | \_GatedResidualNetwork | 6.6 K | train

8 | static\_context\_cell\_encoder\_grn | \_GatedResidualNetwork | 6.6 K | train

9 | static\_context\_enrichment | \_GatedResidualNetwork | 6.6 K | train

10 | lstm\_encoder | LSTM | 26.2 K | train

11 | lstm\_decoder | LSTM | 26.2 K | train

12 | post\_lstm\_gan | \_GateAddNorm | 3.4 K | train

13 | static\_enrichment\_grn | \_GatedResidualNetwork | 8.2 K | train

14 | multihead\_attn | \_InterpretableMultiHeadAttention | 6.5 K | train

15 | post\_attn\_gan | \_GateAddNorm | 3.4 K | train

16 | feed\_forward\_block | \_GatedResidualNetwork | 6.6 K | train

17 | pre\_output\_gan | \_GateAddNorm | 3.4 K | train

18 | output\_layer | Linear | 123 | train

------------------------------------------------------------------------------------------------

412 K Trainable params

0 Non-trainable params

412 K Total params

1.648 Total estimated model params size (MB)

3056 Modules in train mode

0 Modules in eval mode

🚀 Startar ny gridsearch-körning med hyperparametrar:

input\_chunk\_length: 72

output\_chunk\_length: 48

hidden\_size: 40

lstm\_layers: 2

num\_attention\_heads: 1

dropout: 0.2

optimizer\_kwargs: {'lr': 0.01}

Epoch 0: 100%|██████████| 205/205 [02:04<00:00, 1.64it/s, train\_loss=0.622]📉 Epoch 1: train\_loss=0.6221

Epoch 1: 100%|██████████| 205/205 [02:04<00:00, 1.65it/s, train\_loss=0.381]📉 Epoch 2: train\_loss=0.3806

Epoch 2: 100%|██████████| 205/205 [02:04<00:00, 1.65it/s, train\_loss=0.320]📉 Epoch 3: train\_loss=0.3203

Epoch 3: 100%|██████████| 205/205 [02:04<00:00, 1.65it/s, train\_loss=0.306]📉 Epoch 4: train\_loss=0.3056

Epoch 4: 100%|██████████| 205/205 [02:04<00:00, 1.65it/s, train\_loss=0.322]📉 Epoch 5: train\_loss=0.3221

Epoch 5: 100%|██████████| 205/205 [02:04<00:00, 1.64it/s, train\_loss=0.204]📉 Epoch 6: train\_loss=0.2038

Epoch 6: 100%|██████████| 205/205 [02:04<00:00, 1.65it/s, train\_loss=0.219]📉 Epoch 7: train\_loss=0.2194

Epoch 7: 100%|██████████| 205/205 [02:04<00:00, 1.65it/s, train\_loss=0.211]📉 Epoch 8: train\_loss=0.2107

Epoch 8: 100%|██████████| 205/205 [02:04<00:00, 1.65it/s, train\_loss=0.170]📉 Epoch 9: train\_loss=0.1697

Epoch 9: 100%|██████████| 205/205 [02:04<00:00, 1.65it/s, train\_loss=0.166]📉 Epoch 10: train\_loss=0.1659

Epoch 10: 97%|█████████▋| 198/205 [02:00<00:04, 1.64it/s, train\_loss=0.172]