In [1]:
```
#| default_exp models.nhits
```

In [10]:
```
#| hide
%load_ext autoreload
%autoreload 2
```

NHITS

Long-horizon forecasting is challenging because of the *volatility* of the predictions and the *computational complexity*. To solve this problem we created the Neural Hierarchical Interpolation for Time Series (NHITS). `NHITS` builds upon `NBEATS` and specializes its partial outputs in the different frequencies of the time series through hierarchical interpolation and multi-rate input processing. On the long-horizon forecasting task `NHITS` improved accuracy by 25% on AAAI's best paper award the `Informer`, while being 50x faster.

The model is composed of several MLPs with ReLU non-linearities. Blocks are connected via doubly residual stacking principle with the backcast $\mathbf{\tilde{y}}_{t-L:t,l}$ and forecast $\mathbf{\hat{y}}_{t+1:t+H,l}$ outputs of the l-th block. Multi-rate input pooling, hierarchical interpolation and backcast residual connections together induce the specialization of the additive predictions in different signal bands, reducing memory footprint and computational time, thus improving the architecture parsimony and accuracy.

References

-Boris N. Oreshkin, Dmitri Carpov, Nicolas Chapados, Yoshua Bengio (2019). "N-BEATS: Neural basis expansion analysis for interpretable time series forecasting".
-Cristian Challu, Kin G. Olivares, Boris N. Oreshkin, Federico Garza, Max Mergenthaler-Canseco, Artur Dubrawski (2023). "NHITS: Neural Hierarchical Interpolation for Time Series Forecasting". Accepted at the Thirty-Seventh AAAI Conference on Artificial Intelligence.
-Zhou, H.; Zhang, S.; Peng, J.; Zhang, S.; Li, J.; Xiong, H.; and Zhang, W. (2020). "Informer: Beyond Efficient Transformer for Long Sequence Time-Series Forecasting". Association for the Advancement of Artificial Intelligence Conference 2021 (AAAI 2021).

 Figure 1. Neural Hierarchical Interpolation for Time Series (NHITS).

In [11]:
```
#| hide
import os
os.environ["PYTORCH_ENABLE_MPS_FALLBACK"] = "1"
os.environ["CUDA_VISIBLE_DEVICES"] = "0"
```

Loading [MathJax]/extensions/Safe.js

In [12]:
```
#| export
from typing import Tuple, Optional

import numpy as np
import torch
import torch.nn as nn
import torch.nn.functional as F

from neuralforecast.losses.pytorch import MAE
from neuralforecast.common._base_windows import BaseWindows
```

In [13]:
```
#| hide
from fastcore.test import test_eq
from nbdev.showdoc import show_doc
from neuralforecast.utils import generate_series
```

In [14]:
```
#| hide
import logging
import warnings

logging.getLogger("pytorch_lightning").setLevel(logging.ERROR)
warnings.filterwarnings("ignore")

import matplotlib.pyplot as plt

#plt.rcParams["axes.grid"]=True
plt.rcParams['font.family'] = 'serif'
#plt.rcParams["figure.figsize"] = (4,2)
```

In [15]:
```
#| export
class _IdentityBasis(nn.Module):
    def __init__(self, backcast_size: int, forecast_size: int,
                 interpolation_mode: str, out_features: int=1):
        super().__init__()
        assert (interpolation_mode in ['linear','nearest']) or ('cubic' in i
        self.forecast_size = forecast_size
        self.backcast_size = backcast_size
        self.interpolation_mode = interpolation_mode
        self.out_features = out_features

    def forward(self, theta: torch.Tensor) -> Tuple[torch.Tensor, torch.Tens

        backcast = theta[:, :self.backcast_size]
        knots = theta[:, self.backcast_size:]

        # Interpolation is performed on default dim=-1 := H
        knots = knots.reshape(len(knots), self.out_features, -1)
        if self.interpolation_mode in ['nearest', 'linear']:
            #knots = knots[:,None,:]
            forecast = F.interpolate(knots, size=self.forecast_size, mode=se
            #forecast = forecast[:,0,:]
        elif 'cubic' in self.interpolation_mode:
            if self.out_features>1:
                raise Exception('Cubic interpolation not available with mult
```

Loading [MathJax]/extensions/Safe.js

```
            batch_size = len(backcast)
            knots = knots[:,None,:,:]
            forecast = torch.zeros((len(knots), self.forecast_size)).to(knot
            n_batches = int(np.ceil(len(knots)/batch_size))
            for i in range(n_batches):
                forecast_i = F.interpolate(knots[i*batch_size:(i+1)*batch_si
                                           size=self.forecast_size, mode='bi
                forecast[i*batch_size:(i+1)*batch_size] += forecast_i[:,0,0,
            forecast = forecast[:,None,:] # [B,H] -> [B,None,H]

        # [B,Q,H] -> [B,H,Q]
        forecast = forecast.permute(0, 2, 1)
        return backcast, forecast
```

In [16]:

```
#| exporti
ACTIVATIONS = ['ReLU',
               'Softplus',
               'Tanh',
               'SELU',
               'LeakyReLU',
               'PReLU',
               'Sigmoid']

POOLING = ['MaxPool1d',
           'AvgPool1d']

class NHITSBlock(nn.Module):
    """
    NHITS block which takes a basis function as an argument.
    """
    def __init__(self,
                 input_size: int,
                 h: int,
                 n_theta: int,
                 mlp_units: list,
                 basis: nn.Module,
                 futr_input_size: int,
                 hist_input_size: int,
                 stat_input_size: int,
                 n_pool_kernel_size: int,
                 pooling_mode: str,
                 dropout_prob: float,
                 activation: str):
        super().__init__()

        pooled_hist_size = int(np.ceil(input_size/n_pool_kernel_size))
        pooled_futr_size = int(np.ceil((input_size+h)/n_pool_kernel_size))

        input_size = pooled_hist_size + \
                     hist_input_size * pooled_hist_size + \
                     futr_input_size * pooled_futr_size + stat_input_size

        self.dropout_prob = dropout_prob
        self.futr_input_size = futr_input_size
        self.hist_input_size = hist_input_size
        self.stat_input_size = stat_input_size
```

Loading [MathJax]/extensions/Safe.js

```
        assert activation in ACTIVATIONS, f'{activation} is not in {ACTIVATI
        assert pooling_mode in POOLING, f'{pooling_mode} is not in {POOLING}

        activ = getattr(nn, activation)()

        self.pooling_layer = getattr(nn, pooling_mode)(kernel_size=n_pool_ke
                                                       stride=n_pool_kernel_

        # Block MLPs
        hidden_layers = [nn.Linear(in_features=input_size,
                                   out_features=mlp_units[0][0])]
        for layer in mlp_units:
            hidden_layers.append(nn.Linear(in_features=layer[0],
                                           out_features=layer[1]))
            hidden_layers.append(activ)

            if self.dropout_prob>0:
                #raise NotImplementedError('dropout')
                hidden_layers.append(nn.Dropout(p=self.dropout_prob))

        output_layer = [nn.Linear(in_features=mlp_units[-1][1], out_features
        layers = hidden_layers + output_layer
        self.layers = nn.Sequential(*layers)
        self.basis = basis

    def forward(self, insample_y: torch.Tensor, futr_exog: torch.Tensor,
                hist_exog: torch.Tensor, stat_exog: torch.Tensor) -> Tuple[t

        # Pooling
        # Pool1d needs 3D input, (B,C,L), adding C dimension
        insample_y = insample_y.unsqueeze(1)
        insample_y = self.pooling_layer(insample_y)
        insample_y = insample_y.squeeze(1)

        # Flatten MLP inputs [B, L+H, C] -> [B, (L+H)*C]
        # Contatenate [ Y_t, | X_{t-L},..., X_{t} | F_{t-L},..., F_{t+H} | S
        batch_size = len(insample_y)
        if self.hist_input_size > 0:
            hist_exog = hist_exog.permute(0,2,1) # [B, L, C] -> [B, C, L]
            hist_exog = self.pooling_layer(hist_exog)
            hist_exog = hist_exog.permute(0,2,1) # [B, C, L] -> [B, L, C]
            insample_y = torch.cat(( insample_y, hist_exog.reshape(batch_siz

        if self.futr_input_size > 0:
            futr_exog = futr_exog.permute(0,2,1) # [B, L, C] -> [B, C, L]
            futr_exog = self.pooling_layer(futr_exog)
            futr_exog = futr_exog.permute(0,2,1) # [B, C, L] -> [B, L, C]
            insample_y = torch.cat(( insample_y, futr_exog.reshape(batch_siz

        if self.stat_input_size > 0:
            insample_y = torch.cat(( insample_y, stat_exog.reshape(batch_siz

        # Compute local projection weights and projection
        theta = self.layers(insample_y)
```

Loading [MathJax]/extensions/Safe.js

```
        backcast, forecast = self.basis(theta)
        return backcast, forecast
```

In [17]:

```
#| export
class NHITS(BaseWindows):
    """ NHITS

    The Neural Hierarchical Interpolation for Time Series (NHITS), is an MLP
    neural architecture with backward and forward residual links. NHITS tack
    memory complexity challenges, by locally specializing its sequential pre
    the signals frequencies with hierarchical interpolation and pooling.

    **Parameters:**<br>
    `h`: int, Forecast horizon. <br>
    `input_size`: int, autorregresive inputs size, y=[1,2,3,4] input_size=2
    `stat_exog_list`: str list, static exogenous columns.<br>
    `hist_exog_list`: str list, historic exogenous columns.<br>
    `futr_exog_list`: str list, future exogenous columns.<br>
    `exclude_insample_y`: bool=False, the model skips the autoregressive fea
    `activation`: str, activation from ['ReLU', 'Softplus', 'Tanh', 'SELU',
    `stack_types`: List[str], stacks list in the form N * ['identity'], to b
    `n_blocks`: List[int], Number of blocks for each stack. Note that len(n_
    `mlp_units`: List[List[int]], Structure of hidden layers for each stack
    `n_freq_downsample`: List[int], list with the stack's coefficients (inve
    `interpolation_mode`: str='linear', interpolation basis from ['linear',
    `n_pool_kernel_size`: List[int], list with the size of the windows to ta
    `pooling_mode`: str, input pooling module from ['MaxPool1d', 'AvgPool1d'
    `dropout_prob_theta`: float, Float between (0, 1). Dropout for NHITS bas
    `loss`: PyTorch module, instantiated train loss class from [losses colle
    `valid_loss`: PyTorch module=`loss`, instantiated valid loss class from
    `max_steps`: int=1000, maximum number of training steps.<br>
    `learning_rate`: float=1e-3, Learning rate between (0, 1).<br>
    `num_lr_decays`: int=-1, Number of learning rate decays, evenly distribu
    `early_stop_patience_steps`: int=-1, Number of validation iterations bef
    `val_check_steps`: int=100, Number of training steps between every valid
    `batch_size`: int=32, number of different series in each batch.<br>
    `valid_batch_size`: int=None, number of different series in each validat
    `windows_batch_size`: int=1024, number of windows to sample in each trai
    `inference_windows_batch_size`: int=-1, number of windows to sample in e
    `start_padding_enabled`: bool=False, if True, the model will pad the tim
    `step_size`: int=1, step size between each window of temporal data.<br>
    `scaler_type`: str='identity', type of scaler for temporal inputs normal
    `random_seed`: int, random_seed for pytorch initializer and numpy genera
    `num_workers_loader`: int=os.cpu_count(), workers to be used by `TimeSer
    `drop_last_loader`: bool=False, if True `TimeSeriesDataLoader` drops las
    `alias`: str, optional,  Custom name of the model.<br>
    `**trainer_kwargs`: int,  keyword trainer arguments inherited from [PyTo

    **References:**<br>
    -[Cristian Challu, Kin G. Olivares, Boris N. Oreshkin, Federico Garza,
    Max Mergenthaler-Canseco, Artur Dubrawski (2023). "NHITS: Neural Hierarc
    Accepted at the Thirty-Seventh AAAI Conference on Artificial Intelligenc
    """
    # Class attributes
    SAMPLING_TYPE = 'windows'
```

Loading [MathJax]/extensions/Safe.js

```
    def __init__(self,
                 h,
                 input_size,
                 futr_exog_list = None,
                 hist_exog_list = None,
                 stat_exog_list = None,
                 exclude_insample_y = False,
                 stack_types: list = ['identity', 'identity', 'identity'],
                 n_blocks: list = [1, 1, 1],
                 mlp_units: list = 3 * [[512, 512]],
                 n_pool_kernel_size: list = [2, 2, 1],
                 n_freq_downsample: list = [4, 2, 1],
                 pooling_mode: str = 'MaxPool1d',
                 interpolation_mode: str = 'linear',
                 dropout_prob_theta = 0.,
                 activation = 'ReLU',
                 loss = MAE(),
                 valid_loss = None,
                 max_steps: int = 1000,
                 learning_rate: float = 1e-3,
                 num_lr_decays: int = 3,
                 early_stop_patience_steps: int =-1,
                 val_check_steps: int = 100,
                 batch_size: int = 32,
                 valid_batch_size: Optional[int] = None,
                 windows_batch_size: int = 1024,
                 inference_windows_batch_size: int = -1,
                 start_padding_enabled = False,
                 step_size: int = 1,
                 scaler_type: str = 'identity',
                 random_seed: int = 1,
                 num_workers_loader = 0,
                 drop_last_loader = False,
                 **trainer_kwargs):

        # Inherit BaseWindows class
        super(NHITS, self).__init__(h=h,
                                    input_size=input_size,
                                    futr_exog_list=futr_exog_list,
                                    hist_exog_list=hist_exog_list,
                                    stat_exog_list=stat_exog_list,
                                    exclude_insample_y = exclude_insample_y,
                                    loss=loss,
                                    valid_loss=valid_loss,
                                    max_steps=max_steps,
                                    learning_rate=learning_rate,
                                    num_lr_decays=num_lr_decays,
                                    early_stop_patience_steps=early_stop_pat
                                    val_check_steps=val_check_steps,
                                    batch_size=batch_size,
                                    windows_batch_size=windows_batch_size,
                                    valid_batch_size=valid_batch_size,
                                    inference_windows_batch_size=inference_w
                                    start_padding_enabled=start_padding_enab
                                    step_size=step_size,
                                    scaler_type=scaler_type,
```

Loading [MathJax]/extensions/Safe.js

```
                                    num_workers_loader=num_workers_loader,
                                    drop_last_loader=drop_last_loader,
                                    random_seed=random_seed,
                                    **trainer_kwargs)

        # Architecture
        self.futr_input_size = len(self.futr_exog_list)
        self.hist_input_size = len(self.hist_exog_list)
        self.stat_input_size = len(self.stat_exog_list)

        blocks = self.create_stack(h=h,
                                   input_size=input_size,
                                   stack_types=stack_types,
                                   futr_input_size=self.futr_input_size,
                                   hist_input_size=self.hist_input_size,
                                   stat_input_size=self.stat_input_size,
                                   n_blocks=n_blocks,
                                   mlp_units=mlp_units,
                                   n_pool_kernel_size=n_pool_kernel_size,
                                   n_freq_downsample=n_freq_downsample,
                                   pooling_mode=pooling_mode,
                                   interpolation_mode=interpolation_mode,
                                   dropout_prob_theta=dropout_prob_theta,
                                   activation=activation)
        self.blocks = torch.nn.ModuleList(blocks)

    def create_stack(self,
                     h,
                     input_size,
                     stack_types,
                     n_blocks,
                     mlp_units,
                     n_pool_kernel_size,
                     n_freq_downsample,
                     pooling_mode,
                     interpolation_mode,
                     dropout_prob_theta,
                     activation,
                     futr_input_size, hist_input_size, stat_input_size):

        block_list = []
        for i in range(len(stack_types)):
            for block_id in range(n_blocks[i]):

                assert stack_types[i] == 'identity', f'Block type {stack_typ

                n_theta = (input_size + self.loss.outputsize_multiplier*max(
                basis = _IdentityBasis(backcast_size=input_size, forecast_si
                                       out_features=self.loss.outputsize_mul
                                       interpolation_mode=interpolation_mode

                nbeats_block = NHITSBlock(h=h,
                                          input_size=input_size,
                                          futr_input_size=futr_input_size,
                                          hist_input_size=hist_input_size,
                                          stat_input_size=stat_input_size,
```

Loading [MathJax]/extensions/Safe.js

```
                                            n_theta=n_theta,
                                            mlp_units=mlp_units,
                                            n_pool_kernel_size=n_pool_kernel_s
                                            pooling_mode=pooling_mode,
                                            basis=basis,
                                            dropout_prob=dropout_prob_theta,
                                            activation=activation)

                # Select type of evaluation and apply it to all layers of bl
                block_list.append(nbeats_block)

        return block_list

    def forward(self, windows_batch):

        # Parse windows_batch
        insample_y    = windows_batch['insample_y']
        insample_mask = windows_batch['insample_mask']
        futr_exog     = windows_batch['futr_exog']
        hist_exog     = windows_batch['hist_exog']
        stat_exog     = windows_batch['stat_exog']

        # insample
        residuals = insample_y.flip(dims=(-1,)) #backcast init
        insample_mask = insample_mask.flip(dims=(-1,))

        forecast = insample_y[:, -1:, None] # Level with Naive1
        block_forecasts = [ forecast.repeat(1, self.h, 1) ]
        for i, block in enumerate(self.blocks):
            backcast, block_forecast = block(insample_y=residuals, futr_exog
                                             hist_exog=hist_exog, stat_exog=
            residuals = (residuals - backcast) * insample_mask
            forecast = forecast + block_forecast

            if self.decompose_forecast:
                block_forecasts.append(block_forecast)

        # Adapting output's domain
        forecast = self.loss.domain_map(forecast)

        if self.decompose_forecast:
            # (n_batch, n_blocks, h, output_size)
            block_forecasts = torch.stack(block_forecasts)
            block_forecasts = block_forecasts.permute(1,0,2,3)
            block_forecasts = block_forecasts.squeeze(-1) # univariate outpu
            return block_forecasts
        else:
            return forecast
```

In [8]:
```
show_doc(NHITS)
```

Loading [MathJax]/extensions/Safe.js

Out[8]:

NHITS

```
 NHITS (h, input_size, futr_exog_list=None, hist_exog_list
=None,
        stat_exog_list=None, exclude_insample_y=False,
        stack_types:list=['identity', 'identity', 'identit
y'],
        n_blocks:list=[1, 1, 1], mlp_units:list=[[512, 51
2], [512, 512],
        [512, 512]], n_pool_kernel_size:list=[2, 2, 1],
        n_freq_downsample:list=[4, 2, 1], pooling_mode:str
='MaxPool1d',
        interpolation_mode:str='linear', dropout_prob_thet
a=0.0,
        activation='ReLU', loss=MAE(), valid_loss=None,
        max_steps:int=1000, learning_rate:float=0.001,
        num_lr_decays:int=3, early_stop_patience_steps:int
=-1,
        val_check_steps:int=100, batch_size:int=32,
        valid_batch_size:Optional[int]=None, windows_batch
_size:int=1024,
        inference_windows_batch_size:int=-1, start_padding
_enabled=False,
        step_size:int=1, scaler_type:str='identity', rando
m_seed:int=1,
        num_workers_loader=0, drop_last_loader=False, **tr
ainer_kwargs)
```

NHITS

The Neural Hierarchical Interpolation for Time Series (NHITS), is an MLP-based deep neural architecture with backward and forward residual links. NHITS tackles volatility and memory complexity challenges, by locally specializing its sequential predictions into the signals frequencies with hierarchical interpolation and pooling.

Parameters:
`h`: int, Forecast horizon.
`input_size`: int, autorregresive inputs size, y=[1,2,3,4] input_size=2 -> y_[t-2:t]=[1,2].
`stat_exog_list`: str list, static exogenous columns.
`hist_exog_list`: str list, historic exogenous columns.

Loading [MathJax]/extensions/Safe.js

`futr_exog_list` : str list, future exogenous columns.
`exclude_insample_y` : bool=False, the model skips the autoregressive features y[t-input_size:t] if True.
`activation` : str, activation from ['ReLU', 'Softplus', 'Tanh', 'SELU', 'LeakyReLU', 'PReLU', 'Sigmoid'].
`stack_types` : List[str], stacks list in the form N * ['identity'], to be deprecated in favor of `n_stacks`. Note that len(stack_types)=len(n_freq_downsample)=len(n_pool_kernel_size).
`n_blocks` : List[int], Number of blocks for each stack. Note that len(n_blocks) = len(stack_types).
`mlp_units` : List[List[int]], Structure of hidden layers for each stack type. Each internal list should contain the number of units of each hidden layer. Note that len(n_hidden) = len(stack_types).
`n_freq_downsample` : List[int], list with the stack's coefficients (inverse expressivity ratios). Note that len(stack_types)=len(n_freq_downsample)=len(n_pool_kernel_size).
`interpolation_mode` : str='linear', interpolation basis from ['linear', 'nearest', 'cubic'].
`n_pool_kernel_size` : List[int], list with the size of the windows to take a max/avg over. Note that len(stack_types)=len(n_freq_downsample)=len(n_pool_kernel_size).
`pooling_mode` : str, input pooling module from ['MaxPool1d', 'AvgPool1d'].
`dropout_prob_theta` : float, Float between (0, 1). Dropout for NHITS basis.
`loss` : PyTorch module, instantiated train loss class from losses collection.
`valid_loss` : PyTorch module=`loss`, instantiated valid loss class from losses collection.
`max_steps` : int=1000, maximum number of training steps.
`learning_rate` : float=1e-3, Learning rate between (0, 1).
`num_lr_decays` : int=-1, Number of learning rate decays, evenly distributed across max_steps.
`early_stop_patience_steps` : int=-1, Number of validation iterations before early stopping.
`val_check_steps` : int=100, Number of training steps between every validation loss check.
`batch_size` : int=32, number of different series in each batch.
`valid_batch_size` : int=None, number of different series in each validation and test batch, if None uses batch_size.
`windows_batch_size` : int=1024, number of windows to sample in each training batch, default uses all.
`inference_windows_batch_size` : int=-1, number of windows to sample in each inference batch, -1 uses all.
`start_padding_enabled` : bool=False, if True, the model will pad the time series

Loading [MathJax]/extensions/Safe.js

with zeros at the beginning, by input size.
`step_size`: int=1, step size between each window of temporal data.
`scaler_type`: str='identity', type of scaler for temporal inputs normalization see temporal scalers.
`random_seed`: int, random_seed for pytorch initializer and numpy generators.
`num_workers_loader`: int=os.cpu_count(), workers to be used by `TimeSeriesDataLoader`.
`drop_last_loader`: bool=False, if True `TimeSeriesDataLoader` drops last non-full batch.
`alias`: str, optional, Custom name of the model.
`**trainer_kwargs`: int, keyword trainer arguments inherited from PyTorch Lighning's trainer.

References:
-Cristian Challu, Kin G. Olivares, Boris N. Oreshkin, Federico Garza, Max Mergenthaler-Canseco, Artur Dubrawski (2023). "NHITS: Neural Hierarchical Interpolation for Time Series Forecasting". Accepted at the Thirty-Seventh AAAI Conference on Artificial Intelligence.

In [11]:
```
show_doc(NHITS.fit, name='NHITS.fit')
```

Loading [MathJax]/extensions/Safe.js

Out[11]:

NHITS.fit

> NHITS.fit (dataset, val_size=0, test_size=0, random_seed=None)

Fit.

The `fit` method, optimizes the neural network's weights using the initialization parameters (`learning_rate`, `windows_batch_size`, ...) and the `loss` function as defined during the initialization. Within `fit` we use a PyTorch Lightning `Trainer` that inherits the initialization's `self.trainer_kwargs`, to customize its inputs, see PL's trainer arguments.

The method is designed to be compatible with SKLearn-like classes and in particular to be compatible with the StatsForecast library.

By default the `model` is not saving training checkpoints to protect disk memory, to get them change `enable_checkpointing=True` in `__init__`.

Parameters:
`dataset`: NeuralForecast's `TimeSeriesDataset`, see documentation.
`val_size`: int, validation size for temporal cross-validation.
`random_seed`: int=None, random_seed for pytorch initializer and numpy generators, overwrites model.**init**'s.
`test_size`: int, test size for temporal cross-validation.

In [12]:
```
show_doc(NHITS.predict, name='NHITS.predict')
```

Loading [MathJax]/extensions/Safe.js

Out[12]:

NHITS.predict

```
 NHITS.predict (dataset, test_size=None, step_size=1, rand
om_seed=None,
                **data_module_kwargs)
```

Predict.

Neural network prediction with PL's `Trainer` execution of `predict_step`.

Parameters:
`dataset` : NeuralForecast's `TimeSeriesDataset` , see documentation.
`test_size` : int=None, test size for temporal cross-validation.
`step_size` : int=1, Step size between each window.
`random_seed` : int=None, random_seed for pytorch initializer and numpy generators, overwrites model.**init**'s.
`**data_module_kwargs` : PL's TimeSeriesDataModule args, see documentation.

In [13]:
```
#| hide
import logging
import warnings
logging.getLogger("pytorch_lightning").setLevel(logging.ERROR)
warnings.filterwarnings("ignore")
```

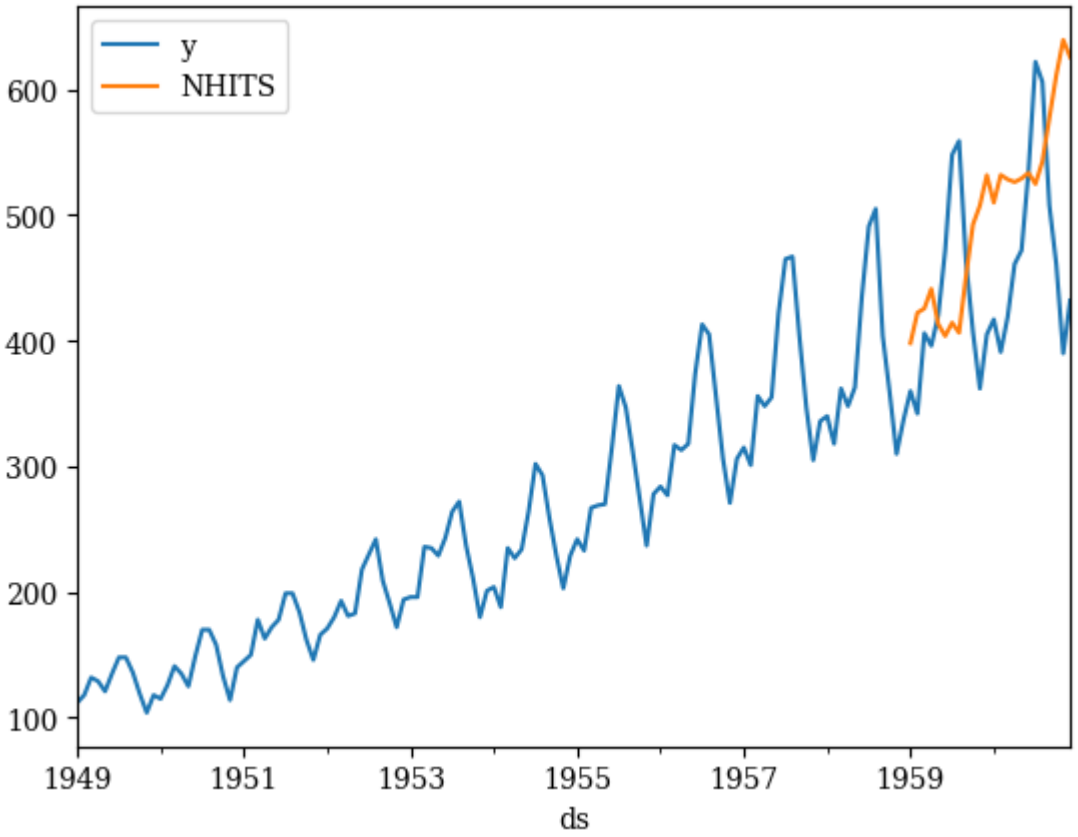
In [14]:
```
#| hide
import pandas as pd
import matplotlib.pyplot as plt

import pytorch_lightning as pl

from neuralforecast.utils import AirPassengersDF as Y_df
from neuralforecast.tsdataset import TimeSeriesDataset, TimeSeriesLoader

Y_train_df = Y_df[Y_df.ds<Y_df['ds'].values[-24]] # 132 train
Y_test_df = Y_df[Y_df.ds>=Y_df['ds'].values[-24]] # 12 test

dataset, *_ = TimeSeriesDataset.from_df(df = Y_train_df)
model = NHITS(h=24,
              input_size=24*2,
              max_steps=1,
              windows_batch_size=None,
              n_freq_downsample=[12,4,1],
              pooling_mode='MaxPool1d')
model.fit(dataset=dataset)
y_hat = model.predict(dataset=dataset)
Y_test_df['NHITS'] = y_hat

pd.concat([Y_train_df, Y_test_df]).drop('unique_id', axis=1).set_index('ds')
```

Loading [MathJax]/extensions/Safe.js

```
Seed set to 1
2023-11-02 17:57:13.906713: I tensorflow/core/util/port.cc:111] oneDNN custo
m operations are on. You may see slightly different numerical results due to
floating-point round-off errors from different computation orders. To turn t
hem off, set the environment variable `TF_ENABLE_ONEDNN_OPTS=0`.
2023-11-02 17:57:14.034658: I tensorflow/tsl/cuda/cudart_stub.cc:28] Could n
ot find cuda drivers on your machine, GPU will not be used.
2023-11-02 17:57:14.550773: E tensorflow/compiler/xla/stream_executor/cuda/c
uda_dnn.cc:9342] Unable to register cuDNN factory: Attempting to register fa
ctory for plugin cuDNN when one has already been registered
2023-11-02 17:57:14.550840: E tensorflow/compiler/xla/stream_executor/cuda/c
uda_fft.cc:609] Unable to register cuFFT factory: Attempting to register fac
tory for plugin cuFFT when one has already been registered
2023-11-02 17:57:14.554706: E tensorflow/compiler/xla/stream_executor/cuda/c
uda_blas.cc:1518] Unable to register cuBLAS factory: Attempting to register
factory for plugin cuBLAS when one has already been registered
2023-11-02 17:57:14.901983: I tensorflow/core/platform/cpu_feature_guard.cc:
182] This TensorFlow binary is optimized to use available CPU instructions i
n performance-critical operations.
To enable the following instructions: AVX2 AVX_VNNI FMA, in other operation
s, rebuild TensorFlow with the appropriate compiler flags.
2023-11-02 17:57:17.367922: W tensorflow/compiler/tf2tensorrt/utils/py_util
s.cc:38] TF-TRT Warning: Could not find TensorRT
Sanity Checking: |
| 0/? [00:00…
Training: |
| 0/? [00:00…
Validation: |
| 0/? [00:00…
Predicting: |
| 0/? [00:00…
```

Out[14]: <Axes: xlabel='ds'>

Loading [MathJax]/extensions/Safe.js

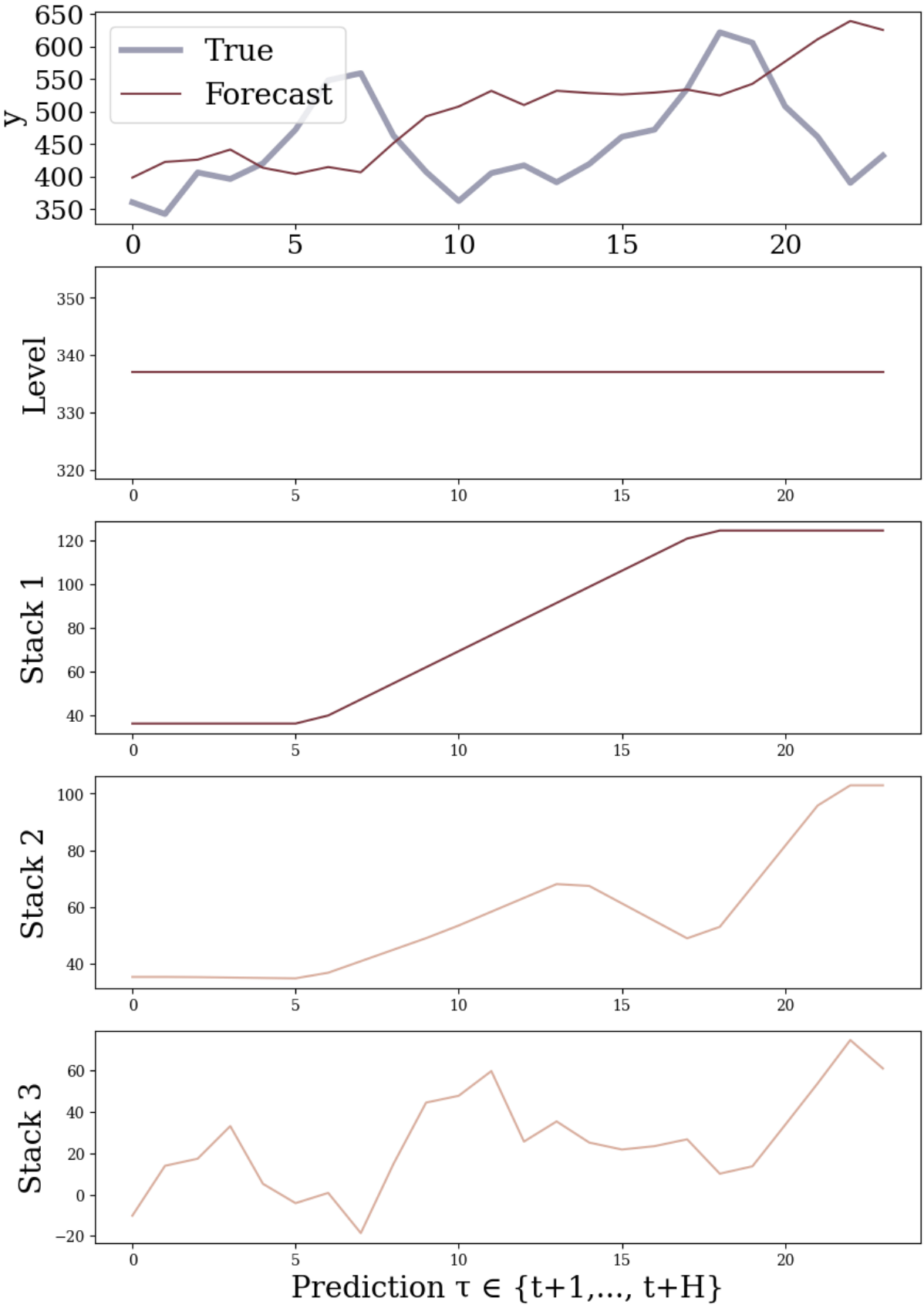


In [15]:
```
#| hide
# qualitative decomposition evaluation
y_hat = model.decompose(dataset=dataset)

fig, ax = plt.subplots(5, 1, figsize=(10, 15))

ax[0].plot(Y_test_df['y'].values, label='True', color="#9C9DB2", linewidth=4
ax[0].plot(y_hat.sum(axis=1).flatten(), label='Forecast', color="#7B3841")
ax[0].legend(prop={'size': 20})
for label in (ax[0].get_xticklabels() + ax[0].get_yticklabels()):
    label.set_fontsize(18)
ax[0].set_ylabel('y', fontsize=20)

ax[1].plot(y_hat[0,0], label='level', color="#7B3841")
ax[1].set_ylabel('Level', fontsize=20)

ax[2].plot(y_hat[0,1], label='stack1', color="#7B3841")
ax[2].set_ylabel('Stack 1', fontsize=20)

ax[3].plot(y_hat[0,2], label='stack2', color="#D9AE9E")
ax[3].set_ylabel('Stack 2', fontsize=20)

ax[4].plot(y_hat[0,3], label='stack3', color="#D9AE9E")
ax[4].set_ylabel('Stack 3', fontsize=20)

ax[4].set_xlabel('Prediction \u03C4 \u2208 {t+1,..., t+H}', fontsize=20)
```

```
Predicting: |
| 0/? [00:00…
```

Loading [MathJax]/extensions/Safe.js

Out[15]: Text(0.5, 0, 'Prediction τ ∈ {t+1,..., t+H}')



Usage Example

Loading [MathJax]/extensions/Safe.js

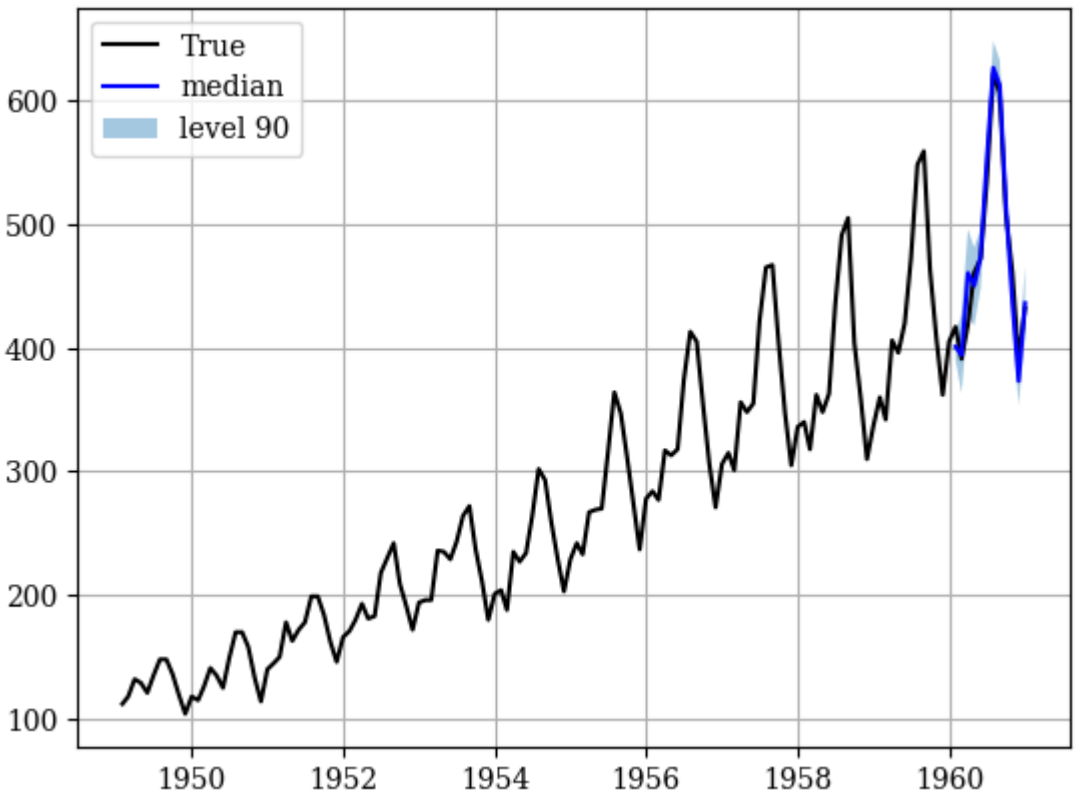
In [16]:

```
#| eval: false
import numpy as np
import pandas as pd
import pytorch_lightning as pl
import matplotlib.pyplot as plt

from neuralforecast import NeuralForecast
from neuralforecast.models import NHITS
from neuralforecast.losses.pytorch import MQLoss, DistributionLoss, PMM, GMM
from neuralforecast.tsdataset import TimeSeriesDataset
from neuralforecast.utils import AirPassengers, AirPassengersPanel, AirPasse


Y_train_df = AirPassengersPanel[AirPassengersPanel.ds<AirPassengersPanel['ds
Y_test_df = AirPassengersPanel[AirPassengersPanel.ds>=AirPassengersPanel['ds

model = NHITS(h=12,
              input_size=24,
              loss=DistributionLoss(distribution='StudentT', level=[80, 90],
              #loss=DistributionLoss(distribution='Normal', level=[80, 90],
              #loss=DistributionLoss(distribution='Poisson', level=[80, 90],
              #loss=DistributionLoss(distribution='Tweedie', level=[80, 90],
              #loss=DistributionLoss(distribution='NegativeBinomial', level=
              #loss=NBMM(n_components=2, level=[80,90]),
              #loss=GMM(n_components=2, level=[80,90]),
              #loss=PMM(n_components=1, level=[80,90]),
              stat_exog_list=['airline1'],
              futr_exog_list=['trend'],
              n_freq_downsample=[2, 1, 1],
              scaler_type='robust',
              max_steps=200,
              early_stop_patience_steps=2,
              inference_windows_batch_size=1,
              val_check_steps=10,
              learning_rate=1e-3)

fcst = NeuralForecast(models=[model], freq='M')
fcst.fit(df=Y_train_df, static_df=AirPassengersStatic, val_size=12)
forecasts = fcst.predict(futr_df=Y_test_df)

# Plot quantile predictions
Y_hat_df = forecasts.reset_index(drop=False).drop(columns=['unique_id','ds']
plot_df = pd.concat([Y_test_df, Y_hat_df], axis=1)
plot_df = pd.concat([Y_train_df, plot_df])

plot_df = plot_df[plot_df.unique_id=='Airline1'].drop('unique_id', axis=1)
plt.plot(plot_df['ds'], plot_df['y'], c='black', label='True')
plt.plot(plot_df['ds'], plot_df['NHITS-median'], c='blue', label='median')
plt.fill_between(x=plot_df['ds'][-12:],
                 y1=plot_df['NHITS-lo-90'][-12:].values,
                 y2=plot_df['NHITS-hi-90'][-12:].values,
                 alpha=0.4, label='level 90')
plt.legend()
plt.grid()
plt.plot()
```

Loading [MathJax]/extensions/Safe.js

```
Seed set to 1
Sanity Checking: |
| 0/? [00:00…
Training: |
| 0/? [00:00…
Validation: |
| 0/? [00:00…
Validation: |
| 0/? [00:00…
Validation: |
| 0/? [00:00…
Validation: |
| 0/? [00:00…
Validation: |
| 0/? [00:00…
Predicting: |
| 0/? [00:00…
```

Out[16]: []



In [17]:
```
from neuralforecast.losses.numpy import mae, mse

y_true = Y_test_df.y.values
y_hat = Y_hat_df['NHITS-median'].values


print('MAE: ', mae(y_hat, y_true))
print('MSE: ', mse(y_hat, y_true))
```

```
MAE:  12.984341939290365
MSE:  284.7534357710586
```

Loading [MathJax]/extensions/Safe.js

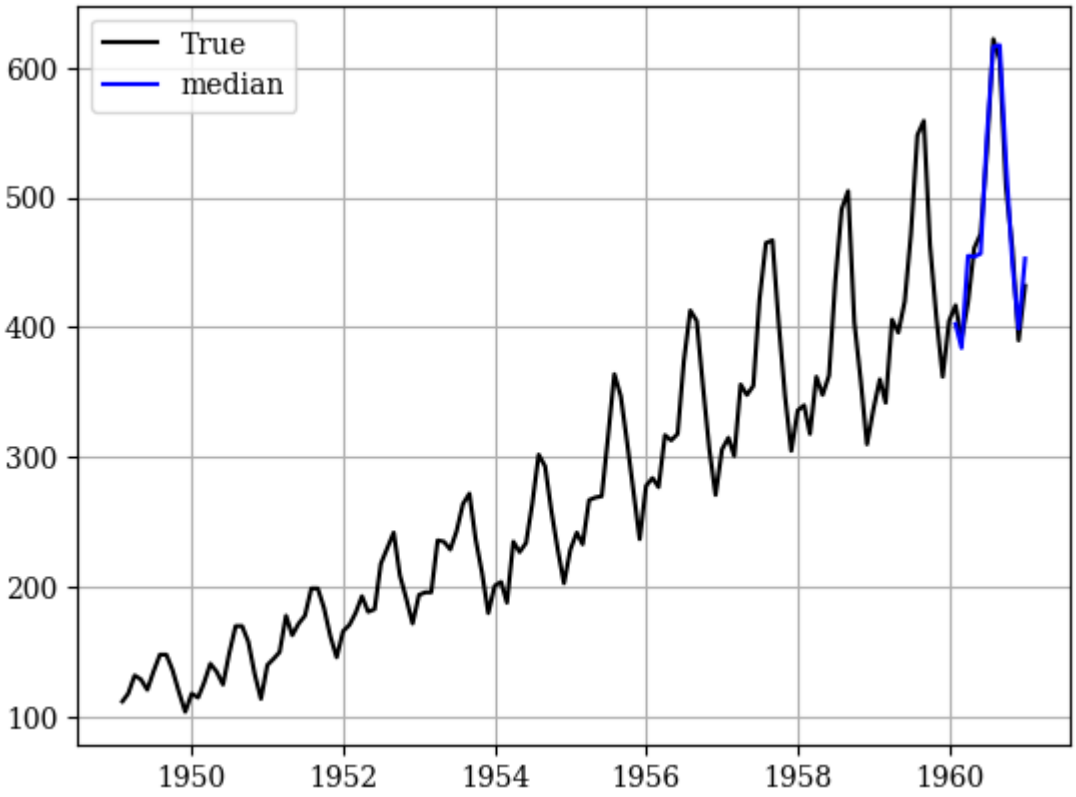
In [18]:
```
#| eval: false
import numpy as np
import pandas as pd
import pytorch_lightning as pl
import matplotlib.pyplot as plt

from neuralforecast import NeuralForecast
from neuralforecast.models import NHITS
from neuralforecast.losses.pytorch import DistributionLoss, HuberLoss, MAE
from neuralforecast.tsdataset import TimeSeriesDataset
from neuralforecast.utils import AirPassengers, AirPassengersPanel, AirPasse

#AirPassengersPanel['y'] = 1 * (AirPassengersPanel['trend'] % 12) < 2
Y_train_df = AirPassengersPanel[AirPassengersPanel.ds<AirPassengersPanel['ds
Y_test_df = AirPassengersPanel[AirPassengersPanel.ds>=AirPassengersPanel['ds

model = NHITS(h=12,
              input_size=24,
              #loss=DistributionLoss(distribution='StudentT', level=[80, 90]
              loss=HuberLoss(delta=0.5),
              valid_loss=MAE(),
              stat_exog_list=['airline1'],
              scaler_type='robust',
              max_steps=200,
              early_stop_patience_steps=2,
              val_check_steps=10,
              learning_rate=1e-3)

fcst = NeuralForecast(models=[model], freq='M')
fcst.fit(df=Y_train_df, static_df=AirPassengersStatic, val_size=12)
forecasts = fcst.predict(futr_df=Y_test_df)

# Plot quantile predictions
Y_hat_df = forecasts.reset_index(drop=False).drop(columns=['unique_id','ds']
plot_df = pd.concat([Y_test_df, Y_hat_df], axis=1)
plot_df = pd.concat([Y_train_df, plot_df])

plot_df = plot_df[plot_df.unique_id=='Airline1'].drop('unique_id', axis=1)
plt.plot(plot_df['ds'], plot_df['y'], c='black', label='True')
plt.plot(plot_df['ds'], plot_df['NHITS'], c='blue', label='median')
# plt.plot(plot_df['ds'], plot_df['NHITS-median'], c='blue', label='median')
# plt.fill_between(x=plot_df['ds'][-12:],
#                  y1=plot_df['NHITS-lo-90'][-12:].values,
#                  y2=plot_df['NHITS-hi-90'][-12:].values,
#                  alpha=0.4, label='level 90')
plt.legend()
plt.grid()
plt.plot()
```

```
Seed set to 1
Sanity Checking: |
| 0/? [00:00…
Training: |
| 0/? [00:00…
```

Loading [MathJax]/extensions/Safe.js

```
Validation: |
| 0/? [00:00…
Validation: |
| 0/? [00:00…
Validation: |
| 0/? [00:00…
Predicting: |
| 0/? [00:00…
```

Out[18]: []



In [19]:
```
from neuralforecast.losses.numpy import mae, mse

y_true = Y_test_df.y.values
y_hat = Y_hat_df['NHITS'].values


print('MAE: ', mae(y_hat, y_true))
print('MSE: ', mse(y_hat, y_true))
```

```
MAE:  14.19897206624349
MSE:  270.2152161945899
```

AutoNHITS Implementation

Exchange rate

In [1]:
```
from ray import tune
import pandas as pd
```

Loading [MathJax]/extensions/Safe.js

```
from neuralforecast.core import NeuralForecast
from neuralforecast.auto import AutoNHITS
```

In [2]:
```
Y_df = pd.read_csv("raw_data/df_Exchange.csv")

Y_df['ds'] = pd.to_datetime(Y_df['ds'])

# For this excercise we are going to take 20% of the DataSet
n_time = len(Y_df.ds.unique())
val_size = int(.1 * n_time)
test_size = int(.2 * n_time)

Y_df.groupby('unique_id').head(2)
```

Out[2]:

	unique_id	ds	y
0	0	1990-01-01	0.606785
1	0	1990-01-02	0.570900
7588	1	1990-01-01	-0.361671
7589	1	1990-01-02	-0.367639
15176	2	1990-01-01	0.735367
15177	2	1990-01-02	0.729629
22764	3	1990-01-01	-1.164373
22765	3	1990-01-02	-1.170907
30352	4	1990-01-01	2.851890
30353	4	1990-01-02	2.851890
37940	5	1990-01-01	-1.861369
37941	5	1990-01-02	-1.838665
45528	6	1990-01-01	-1.820047
45529	6	1990-01-02	-1.847258
53116	OT	1990-01-01	-0.124081
53117	OT	1990-01-02	-0.113588

In [3]:
```
horizon = 96 # 24hrs = 4 * 15 min.

# Use your own config or AutoNHITS.default_config
nhits_config = {
       "learning_rate": tune.choice([1e-3]),
       "max_steps": tune.choice([100]),
       "input_size": tune.choice([5 * horizon]),
       "batch_size": tune.choice([8]),
       "windows_batch_size": tune.choice([256]),
       "n_pool_kernel_size": tune.choice([[2, 2, 2], [16, 8, 1]]),
       "n_freq_downsample": tune.choice([[168, 24, 1], [24, 12, 1], [1, 1, 1
       "activation": tune.choice(['ReLU']),
```

Loading [MathJax]/extensions/Safe.js

```
        "n_blocks":  tune.choice([[1, 1, 1]]),
        "mlp_units":  tune.choice([[[512, 512], [512, 512], [512, 512]]]),
        "interpolation_mode": tune.choice(['linear']),
        "val_check_steps": tune.choice([10]),
        "random_seed": tune.randint(1, 10),
    }
```

In [4]:
```
models = [AutoNHITS(h=horizon,
                    config=nhits_config,
                    num_samples=5)]
```

In [5]:
```
nf = NeuralForecast(
    models=models,
    freq='D')

Y_hat_df = nf.cross_validation(df=Y_df, val_size=val_size,
                               test_size=test_size, n_windows=None)
```

```
(_train_tune pid=6155) Seed set to 6
(_train_tune pid=6155) 2023-11-04 14:21:28.456271: I tensorflow/core/util/port.cc:111] oneDNN custom operations are on. You may see slightly different numerical results due to floating-point round-off errors from different computation orders. To turn them off, set the environment variable `TF_ENABLE_ONEDNN_OPTS=0`.
(_train_tune pid=6155) 2023-11-04 14:21:28.458635: I tensorflow/tsl/cuda/cudart_stub.cc:28] Could not find cuda drivers on your machine, GPU will not be used.
(_train_tune pid=6155) 2023-11-04 14:21:28.494230: E tensorflow/compiler/xla/stream_executor/cuda/cuda_dnn.cc:9342] Unable to register cuDNN factory: Attempting to register factory for plugin cuDNN when one has already been registered
(_train_tune pid=6155) 2023-11-04 14:21:28.494266: E tensorflow/compiler/xla/stream_executor/cuda/cuda_fft.cc:609] Unable to register cuFFT factory: Attempting to register factory for plugin cuFFT when one has already been registered
(_train_tune pid=6155) 2023-11-04 14:21:28.494295: E tensorflow/compiler/xla/stream_executor/cuda/cuda_blas.cc:1518] Unable to register cuBLAS factory: Attempting to register factory for plugin cuBLAS when one has already been registered
(_train_tune pid=6155) 2023-11-04 14:21:28.502893: I tensorflow/core/platform/cpu_feature_guard.cc:182] This TensorFlow binary is optimized to use available CPU instructions in performance-critical operations.
(_train_tune pid=6155) To enable the following instructions: AVX2 AVX_VNNI FMA, in other operations, rebuild TensorFlow with the appropriate compiler flags.
(_train_tune pid=6155) 2023-11-04 14:21:29.276967: W tensorflow/compiler/tf2tensorrt/utils/py_utils.cc:38] TF-TRT Warning: Could not find TensorRT
```

Loading [MathJax]/extensions/Safe.js

```
Sanity Checking DataLoader 0:   0%|          | 0/1 [00:00<?, ?it/s]
Sanity Checking DataLoader 0: 100%|██████████| 1/1 [00:00<00:00,  3.56it/s]
Epoch 0:   0%|          | 0/1 [00:00<?, ?it/s]
Epoch 1:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.20
5, train_loss_epoch=0.205]
Epoch 1: 100%|██████████| 1/1 [00:00<00:00, 10.63it/s, v_num=0, train_loss_s
tep=0.232, train_loss_epoch=0.205]
Epoch 2:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.23
2, train_loss_epoch=0.232]
Epoch 3:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.23
5, train_loss_epoch=0.235]
Epoch 4:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.20
7, train_loss_epoch=0.207]
Epoch 5:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.20
7, train_loss_epoch=0.207]
Epoch 6:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.22
8, train_loss_epoch=0.228]
Epoch 7:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.24
8, train_loss_epoch=0.248]
Epoch 8:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.21
7, train_loss_epoch=0.217]
Epoch 9:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.20
7, train_loss_epoch=0.207]
Epoch 9: 100%|██████████| 1/1 [00:00<00:00,  9.96it/s, v_num=0, train_loss_s
tep=0.229, train_loss_epoch=0.207]
Validation: |          | 0/? [00:00<?, ?it/s]
Validation:   0%|          | 0/1 [00:00<?, ?it/s]
Validation DataLoader 0:   0%|          | 0/1 [00:00<?, ?it/s]
(_train_tune pid=6155)
Validation DataLoader 0: 100%|██████████| 1/1 [00:00<00:00,  3.74it/s]
Epoch 10:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
29, train_loss_epoch=0.229, valid_loss=0.264]
Epoch 11:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
07, train_loss_epoch=0.207, valid_loss=0.264]
Epoch 12:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
04, train_loss_epoch=0.204, valid_loss=0.264]
Epoch 13:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
28, train_loss_epoch=0.228, valid_loss=0.264]
Epoch 14:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
49, train_loss_epoch=0.249, valid_loss=0.264]
Epoch 15:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
07, train_loss_epoch=0.207, valid_loss=0.264]
Epoch 16:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.1
84, train_loss_epoch=0.184, valid_loss=0.264]
Epoch 17:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
06, train_loss_epoch=0.206, valid_loss=0.264]
Epoch 18:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
00, train_loss_epoch=0.200, valid_loss=0.264]
Epoch 19:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
18, train_loss_epoch=0.218, valid_loss=0.264]
Epoch 19: 100%|██████████| 1/1 [00:00<00:00,  7.69it/s, v_num=0, train_loss_
step=0.205, train_loss_epoch=0.218, valid_loss=0.264]
Validation: |          | 0/? [00:00<?, ?it/s]
Validation:   0%|          | 0/1 [00:00<?, ?it/s]
Validation DataLoader 0:   0%|          | 0/1 [00:00<?, ?it/s]
(_train_tune pid=6155)
```

Loading [MathJax]/extensions/Safe.js

```
Validation DataLoader 0: 100%|██████████| 1/1 [00:00<00:00,  3.43it/s]
Epoch 20:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
05, train_loss_epoch=0.205, valid_loss=0.298]
Epoch 21:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
11, train_loss_epoch=0.211, valid_loss=0.298]
Epoch 22:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
15, train_loss_epoch=0.215, valid_loss=0.298]
Epoch 23:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
13, train_loss_epoch=0.213, valid_loss=0.298]
Epoch 24:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
08, train_loss_epoch=0.208, valid_loss=0.298]
Epoch 25:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
14, train_loss_epoch=0.214, valid_loss=0.298]
Epoch 26:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
15, train_loss_epoch=0.215, valid_loss=0.298]
Epoch 27:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
07, train_loss_epoch=0.207, valid_loss=0.298]
Epoch 28:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
01, train_loss_epoch=0.201, valid_loss=0.298]
Epoch 29:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
22, train_loss_epoch=0.222, valid_loss=0.298]
Epoch 29: 100%|██████████| 1/1 [00:00<00:00,  9.48it/s, v_num=0, train_loss_
step=0.222, train_loss_epoch=0.222, valid_loss=0.298]
Validation: |          | 0/? [00:00<?, ?it/s]
Validation:   0%|          | 0/1 [00:00<?, ?it/s]
Validation DataLoader 0:   0%|          | 0/1 [00:00<?, ?it/s]
(_train_tune pid=6155)
Validation DataLoader 0: 100%|██████████| 1/1 [00:00<00:00,  3.52it/s]
Epoch 30:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
22, train_loss_epoch=0.222, valid_loss=0.360]
Epoch 31:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
15, train_loss_epoch=0.215, valid_loss=0.360]
Epoch 32:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
05, train_loss_epoch=0.205, valid_loss=0.360]
Epoch 33:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
07, train_loss_epoch=0.207, valid_loss=0.360]
Epoch 34:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
26, train_loss_epoch=0.226, valid_loss=0.360]
Epoch 35:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
07, train_loss_epoch=0.207, valid_loss=0.360]
Epoch 36:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.1
98, train_loss_epoch=0.198, valid_loss=0.360]
Epoch 37:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
06, train_loss_epoch=0.206, valid_loss=0.360]
Epoch 38:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
11, train_loss_epoch=0.211, valid_loss=0.360]
Epoch 39:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
08, train_loss_epoch=0.208, valid_loss=0.360]
Epoch 39: 100%|██████████| 1/1 [00:00<00:00,  5.96it/s, v_num=0, train_loss_
step=0.193, train_loss_epoch=0.208, valid_loss=0.360]
Validation: |          | 0/? [00:00<?, ?it/s]
Validation:   0%|          | 0/1 [00:00<?, ?it/s]
Validation DataLoader 0:   0%|          | 0/1 [00:00<?, ?it/s]
(_train_tune pid=6155)
Validation DataLoader 0: 100%|██████████| 1/1 [00:00<00:00,  3.19it/s]
Epoch 40:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.1
```

Loading [MathJax]/extensions/Safe.js

```
93, train_loss_epoch=0.193, valid_loss=0.266]
Epoch 41:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
07, train_loss_epoch=0.207, valid_loss=0.266]
Epoch 42:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
11, train_loss_epoch=0.211, valid_loss=0.266]
Epoch 43:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
05, train_loss_epoch=0.205, valid_loss=0.266]
Epoch 44:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.1
95, train_loss_epoch=0.195, valid_loss=0.266]
Epoch 45:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.1
97, train_loss_epoch=0.197, valid_loss=0.266]
Epoch 46:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
16, train_loss_epoch=0.216, valid_loss=0.266]
Epoch 47:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.1
92, train_loss_epoch=0.192, valid_loss=0.266]
Epoch 48:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
11, train_loss_epoch=0.211, valid_loss=0.266]
Epoch 49:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
16, train_loss_epoch=0.216, valid_loss=0.266]
Epoch 49: 100%|██████████| 1/1 [00:00<00:00,  7.50it/s, v_num=0, train_loss_
step=0.199, train_loss_epoch=0.216, valid_loss=0.266]
Validation: |          | 0/? [00:00<?, ?it/s]
Validation:   0%|          | 0/1 [00:00<?, ?it/s]
Validation DataLoader 0:   0%|          | 0/1 [00:00<?, ?it/s]
(_train_tune pid=6155)
Validation DataLoader 0: 100%|██████████| 1/1 [00:00<00:00,  2.88it/s]
Epoch 50:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.1
99, train_loss_epoch=0.199, valid_loss=0.272]
Epoch 51:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.1
97, train_loss_epoch=0.197, valid_loss=0.272]
Epoch 52:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.1
99, train_loss_epoch=0.199, valid_loss=0.272]
Epoch 53:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.1
97, train_loss_epoch=0.197, valid_loss=0.272]
Epoch 54:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.1
89, train_loss_epoch=0.189, valid_loss=0.272]
Epoch 55:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
09, train_loss_epoch=0.209, valid_loss=0.272]
Epoch 56:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.1
95, train_loss_epoch=0.195, valid_loss=0.272]
Epoch 57:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
05, train_loss_epoch=0.205, valid_loss=0.272]
Epoch 58:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
16, train_loss_epoch=0.216, valid_loss=0.272]
Epoch 59:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.1
98, train_loss_epoch=0.198, valid_loss=0.272]
Epoch 59: 100%|██████████| 1/1 [00:00<00:00,  8.26it/s, v_num=0, train_loss_
step=0.199, train_loss_epoch=0.198, valid_loss=0.272]
Validation: |          | 0/? [00:00<?, ?it/s]
Validation:   0%|          | 0/1 [00:00<?, ?it/s]
Validation DataLoader 0:   0%|          | 0/1 [00:00<?, ?it/s]
(_train_tune pid=6155)
Validation DataLoader 0: 100%|██████████| 1/1 [00:00<00:00,  3.13it/s]
Epoch 60:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.1
99, train_loss_epoch=0.199, valid_loss=0.305]
Epoch 61:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
```

Loading [MathJax]/extensions/Safe.js

```
01, train_loss_epoch=0.201, valid_loss=0.305]
Epoch 62:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.1
86, train_loss_epoch=0.186, valid_loss=0.305]
Epoch 63:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.1
97, train_loss_epoch=0.197, valid_loss=0.305]
Epoch 64:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.1
95, train_loss_epoch=0.195, valid_loss=0.305]
Epoch 65:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.1
89, train_loss_epoch=0.189, valid_loss=0.305]
Epoch 66:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
00, train_loss_epoch=0.200, valid_loss=0.305]
Epoch 67:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.1
96, train_loss_epoch=0.196, valid_loss=0.305]
Epoch 68:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
35, train_loss_epoch=0.235, valid_loss=0.305]
Epoch 69: 100%|██████████| 1/1 [00:00<00:00, 13.34it/s, v_num=0, train_loss_
step=0.183, train_loss_epoch=0.193, valid_loss=0.305]
Validation: |          | 0/? [00:00<?, ?it/s]
Validation:   0%|          | 0/1 [00:00<?, ?it/s]
Validation DataLoader 0:   0%|          | 0/1 [00:00<?, ?it/s]
(_train_tune pid=6155)
Validation DataLoader 0: 100%|██████████| 1/1 [00:00<00:00,  4.84it/s]
Epoch 70:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.1
83, train_loss_epoch=0.183, valid_loss=0.277]
Epoch 71:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.1
93, train_loss_epoch=0.193, valid_loss=0.277]
Epoch 72:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.1
94, train_loss_epoch=0.194, valid_loss=0.277]
Epoch 73:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.1
93, train_loss_epoch=0.193, valid_loss=0.277]
Epoch 74:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.1
86, train_loss_epoch=0.186, valid_loss=0.277]
Epoch 76:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
22, train_loss_epoch=0.222, valid_loss=0.277]
Epoch 77:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
09, train_loss_epoch=0.209, valid_loss=0.277]
Epoch 78:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
07, train_loss_epoch=0.207, valid_loss=0.277]
Epoch 79: 100%|██████████| 1/1 [00:00<00:00, 13.51it/s, v_num=0, train_loss_
step=0.206, train_loss_epoch=0.206, valid_loss=0.277]
Validation: |          | 0/? [00:00<?, ?it/s]
Validation:   0%|          | 0/1 [00:00<?, ?it/s]
Validation DataLoader 0:   0%|          | 0/1 [00:00<?, ?it/s]
(_train_tune pid=6155)
Validation DataLoader 0: 100%|██████████| 1/1 [00:00<00:00,  4.99it/s]
Epoch 80:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
06, train_loss_epoch=0.206, valid_loss=0.361]
Epoch 81:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.1
84, train_loss_epoch=0.184, valid_loss=0.361]
Epoch 82:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.1
96, train_loss_epoch=0.196, valid_loss=0.361]
Epoch 83:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
16, train_loss_epoch=0.216, valid_loss=0.361]
Epoch 84:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
03, train_loss_epoch=0.203, valid_loss=0.361]
Epoch 85:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.1
```

Loading [MathJax]/extensions/Safe.js

```
92, train_loss_epoch=0.192, valid_loss=0.361]
Epoch 87:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
02, train_loss_epoch=0.202, valid_loss=0.361]
Epoch 88:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
07, train_loss_epoch=0.207, valid_loss=0.361]
Epoch 89:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.1
86, train_loss_epoch=0.186, valid_loss=0.361]
Epoch 89: 100%|██████████| 1/1 [00:00<00:00, 11.39it/s, v_num=0, train_loss_
step=0.204, train_loss_epoch=0.186, valid_loss=0.361]
Validation: |          | 0/? [00:00<?, ?it/s]
Validation:   0%|          | 0/1 [00:00<?, ?it/s]
Validation DataLoader 0:   0%|          | 0/1 [00:00<?, ?it/s]
(_train_tune pid=6155)
Validation DataLoader 0: 100%|██████████| 1/1 [00:00<00:00,  4.67it/s]
Epoch 90:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
04, train_loss_epoch=0.204, valid_loss=0.286]
Epoch 91:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
06, train_loss_epoch=0.206, valid_loss=0.286]
Epoch 92:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
10, train_loss_epoch=0.210, valid_loss=0.286]
Epoch 93:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
05, train_loss_epoch=0.205, valid_loss=0.286]
Epoch 94:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
00, train_loss_epoch=0.200, valid_loss=0.286]
Epoch 96:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.1
80, train_loss_epoch=0.180, valid_loss=0.286]
Epoch 97:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
08, train_loss_epoch=0.208, valid_loss=0.286]
Epoch 98:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
01, train_loss_epoch=0.201, valid_loss=0.286]
Epoch 99:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
14, train_loss_epoch=0.214, valid_loss=0.286]
Epoch 99: 100%|██████████| 1/1 [00:00<00:00, 10.27it/s, v_num=0, train_loss_
step=0.194, train_loss_epoch=0.214, valid_loss=0.286]
Validation: |          | 0/? [00:00<?, ?it/s]
Validation:   0%|          | 0/1 [00:00<?, ?it/s]
Validation DataLoader 0:   0%|          | 0/1 [00:00<?, ?it/s]
(_train_tune pid=6155)
Validation DataLoader 0: 100%|██████████| 1/1 [00:00<00:00,  4.27it/s]
Epoch 99: 100%|██████████| 1/1 [00:00<00:00,  2.87it/s, v_num=0, train_loss_
step=0.194, train_loss_epoch=0.194, valid_loss=0.340]
Sanity Checking: |          | 0/? [00:00<?, ?it/s]
Sanity Checking DataLoader 0:   0%|          | 0/1 [00:00<?, ?it/s]
(_train_tune pid=6155) Seed set to 7
```

Loading [MathJax]/extensions/Safe.js

```
Epoch 0:   0%|          | 0/1 [00:00<?, ?it/s]
Epoch 1:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.23
5, train_loss_epoch=0.235]
Epoch 2:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.25
6, train_loss_epoch=0.256]
Epoch 3:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.22
8, train_loss_epoch=0.228]
Epoch 4:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.20
8, train_loss_epoch=0.208]
Epoch 5:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.22
2, train_loss_epoch=0.222]
Epoch 6:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.22
8, train_loss_epoch=0.228]
Epoch 7:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.20
6, train_loss_epoch=0.206]
Epoch 8:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.21
6, train_loss_epoch=0.216]
Epoch 9:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.20
5, train_loss_epoch=0.205]
Epoch 9: 100%|██████████| 1/1 [00:00<00:00,  9.74it/s, v_num=0, train_loss_s
tep=0.219, train_loss_epoch=0.205]
Validation: |          | 0/? [00:00<?, ?it/s]
Validation:   0%|          | 0/1 [00:00<?, ?it/s]
Validation DataLoader 0:   0%|          | 0/1 [00:00<?, ?it/s]
(_train_tune pid=6155)
Validation DataLoader 0: 100%|██████████| 1/1 [00:00<00:00,  4.03it/s]
Epoch 10:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
19, train_loss_epoch=0.219, valid_loss=0.246]
Epoch 11:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
14, train_loss_epoch=0.214, valid_loss=0.246]
Epoch 12:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
21, train_loss_epoch=0.221, valid_loss=0.246]
Epoch 13:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.1
97, train_loss_epoch=0.197, valid_loss=0.246]
Epoch 14:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
13, train_loss_epoch=0.213, valid_loss=0.246]
Epoch 15:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.1
97, train_loss_epoch=0.197, valid_loss=0.246]
Epoch 16:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
20, train_loss_epoch=0.220, valid_loss=0.246]
Epoch 17:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
09, train_loss_epoch=0.209, valid_loss=0.246]
Epoch 17: 100%|██████████| 1/1 [00:00<00:00, 10.53it/s, v_num=0, train_loss_
step=0.234, train_loss_epoch=0.209, valid_loss=0.246]
Epoch 18:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
34, train_loss_epoch=0.234, valid_loss=0.246]
Epoch 19:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
16, train_loss_epoch=0.216, valid_loss=0.246]
Epoch 19: 100%|██████████| 1/1 [00:00<00:00, 10.87it/s, v_num=0, train_loss_
step=0.210, train_loss_epoch=0.216, valid_loss=0.246]
Validation: |          | 0/? [00:00<?, ?it/s]
Validation:   0%|          | 0/1 [00:00<?, ?it/s]
Validation DataLoader 0:   0%|          | 0/1 [00:00<?, ?it/s]
(_train_tune pid=6155)
Validation DataLoader 0: 100%|██████████| 1/1 [00:00<00:00,  4.85it/s]
Epoch 20:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
```

Loading [MathJax]/extensions/Safe.js

```
10, train_loss_epoch=0.210, valid_loss=0.291]
Epoch 21:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
09, train_loss_epoch=0.209, valid_loss=0.291]
Epoch 22:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
05, train_loss_epoch=0.205, valid_loss=0.291]
Epoch 23:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
18, train_loss_epoch=0.218, valid_loss=0.291]
Epoch 24:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
19, train_loss_epoch=0.219, valid_loss=0.291]
Epoch 25:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
21, train_loss_epoch=0.221, valid_loss=0.291]
Epoch 27:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
14, train_loss_epoch=0.214, valid_loss=0.291]
Epoch 28:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
23, train_loss_epoch=0.223, valid_loss=0.291]
Epoch 29:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
01, train_loss_epoch=0.201, valid_loss=0.291]
Epoch 29: 100%|██████████| 1/1 [00:00<00:00, 11.55it/s, v_num=0, train_loss_
step=0.212, train_loss_epoch=0.201, valid_loss=0.291]
Validation: |          | 0/? [00:00<?, ?it/s]
Validation:   0%|          | 0/1 [00:00<?, ?it/s]
Validation DataLoader 0:   0%|          | 0/1 [00:00<?, ?it/s]
(_train_tune pid=6155)
Validation DataLoader 0: 100%|██████████| 1/1 [00:00<00:00,  4.78it/s]
Epoch 30:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
12, train_loss_epoch=0.212, valid_loss=0.291]
Epoch 31:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
07, train_loss_epoch=0.207, valid_loss=0.291]
Epoch 32:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
00, train_loss_epoch=0.200, valid_loss=0.291]
Epoch 33:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
01, train_loss_epoch=0.201, valid_loss=0.291]
Epoch 34:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
18, train_loss_epoch=0.218, valid_loss=0.291]
Epoch 35:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
02, train_loss_epoch=0.202, valid_loss=0.291]
Epoch 36:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.1
98, train_loss_epoch=0.198, valid_loss=0.291]
Epoch 37:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.1
99, train_loss_epoch=0.199, valid_loss=0.291]
Epoch 39:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
05, train_loss_epoch=0.205, valid_loss=0.291]
Epoch 39: 100%|██████████| 1/1 [00:00<00:00,  9.51it/s, v_num=0, train_loss_
step=0.199, train_loss_epoch=0.205, valid_loss=0.291]
Validation: |          | 0/? [00:00<?, ?it/s]
Validation:   0%|          | 0/1 [00:00<?, ?it/s]
Validation DataLoader 0:   0%|          | 0/1 [00:00<?, ?it/s]
(_train_tune pid=6155)
Validation DataLoader 0: 100%|██████████| 1/1 [00:00<00:00,  4.11it/s]
Epoch 40:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.1
99, train_loss_epoch=0.199, valid_loss=0.343]
Epoch 41:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
09, train_loss_epoch=0.209, valid_loss=0.343]
Epoch 42:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
19, train_loss_epoch=0.219, valid_loss=0.343]
Epoch 43:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
```

Loading [MathJax]/extensions/Safe.js

```
00, train_loss_epoch=0.200, valid_loss=0.343]
Epoch 45:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.1
99, train_loss_epoch=0.199, valid_loss=0.343]
Epoch 46:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.1
96, train_loss_epoch=0.196, valid_loss=0.343]
Epoch 47:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.1
96, train_loss_epoch=0.196, valid_loss=0.343]
Epoch 48:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.1
87, train_loss_epoch=0.187, valid_loss=0.343]
Epoch 49: 100%|██████████| 1/1 [00:00<00:00, 12.69it/s, v_num=0, train_loss_
step=0.219, train_loss_epoch=0.188, valid_loss=0.343]
Validation: |          | 0/? [00:00<?, ?it/s]
Validation:   0%|          | 0/1 [00:00<?, ?it/s]
Validation DataLoader 0:   0%|          | 0/1 [00:00<?, ?it/s]
(_train_tune pid=6155)
Validation DataLoader 0: 100%|██████████| 1/1 [00:00<00:00,  4.35it/s]
Epoch 49: 100%|██████████| 1/1 [00:00<00:00,  3.16it/s, v_num=0, train_loss_
step=0.219, train_loss_epoch=0.188, valid_loss=0.343]
Epoch 49: 100%|██████████| 1/1 [00:00<00:00,  3.10it/s, v_num=0, train_loss_
step=0.219, train_loss_epoch=0.219, valid_loss=0.343]
Epoch 50:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
19, train_loss_epoch=0.219, valid_loss=0.343]
Epoch 51:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
08, train_loss_epoch=0.208, valid_loss=0.343]
Epoch 52:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
01, train_loss_epoch=0.201, valid_loss=0.343]
Epoch 53:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
04, train_loss_epoch=0.204, valid_loss=0.343]
Epoch 54:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
06, train_loss_epoch=0.206, valid_loss=0.343]
Epoch 55:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
11, train_loss_epoch=0.211, valid_loss=0.343]
Epoch 56:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.1
93, train_loss_epoch=0.193, valid_loss=0.343]
Epoch 57:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
00, train_loss_epoch=0.200, valid_loss=0.343]
Epoch 58:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
06, train_loss_epoch=0.206, valid_loss=0.343]
Epoch 59:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.1
88, train_loss_epoch=0.188, valid_loss=0.343]
Epoch 59: 100%|██████████| 1/1 [00:00<00:00,  8.69it/s, v_num=0, train_loss_
step=0.185, train_loss_epoch=0.188, valid_loss=0.343]
Validation: |          | 0/? [00:00<?, ?it/s]
Validation:   0%|          | 0/1 [00:00<?, ?it/s]
Validation DataLoader 0:   0%|          | 0/1 [00:00<?, ?it/s]
(_train_tune pid=6155)
Validation DataLoader 0: 100%|██████████| 1/1 [00:00<00:00,  3.62it/s]
Epoch 60:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.1
85, train_loss_epoch=0.185, valid_loss=0.286]
Epoch 61:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
15, train_loss_epoch=0.215, valid_loss=0.286]
Epoch 62:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.1
99, train_loss_epoch=0.199, valid_loss=0.286]
Epoch 63:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.1
99, train_loss_epoch=0.199, valid_loss=0.286]
Epoch 64:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
```

Loading [MathJax]/extensions/Safe.js

```
02, train_loss_epoch=0.202, valid_loss=0.286]
Epoch 65:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
08, train_loss_epoch=0.208, valid_loss=0.286]
Epoch 66:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
09, train_loss_epoch=0.209, valid_loss=0.286]
Epoch 67:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.1
80, train_loss_epoch=0.180, valid_loss=0.286]
Epoch 67: 100%|██████████| 1/1 [00:00<00:00,  8.93it/s, v_num=0, train_loss_
step=0.207, train_loss_epoch=0.180, valid_loss=0.286]
Epoch 68:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
07, train_loss_epoch=0.207, valid_loss=0.286]
Epoch 68: 100%|██████████| 1/1 [00:00<00:00,  9.49it/s, v_num=0, train_loss_
step=0.195, train_loss_epoch=0.195, valid_loss=0.286]
Epoch 69:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.1
95, train_loss_epoch=0.195, valid_loss=0.286]
Epoch 69: 100%|██████████| 1/1 [00:00<00:00, 10.00it/s, v_num=0, train_loss_
step=0.193, train_loss_epoch=0.195, valid_loss=0.286]
Validation: |          | 0/? [00:00<?, ?it/s]
Validation:   0%|          | 0/1 [00:00<?, ?it/s]
Validation DataLoader 0:   0%|          | 0/1 [00:00<?, ?it/s]
(_train_tune pid=6155)
Validation DataLoader 0: 100%|██████████| 1/1 [00:00<00:00,  3.43it/s]
Epoch 69: 100%|██████████| 1/1 [00:00<00:00,  2.42it/s, v_num=0, train_loss_
step=0.193, train_loss_epoch=0.193, valid_loss=0.282]
Epoch 70:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.1
93, train_loss_epoch=0.193, valid_loss=0.282]
Epoch 71:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
11, train_loss_epoch=0.211, valid_loss=0.282]
Epoch 72:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.1
89, train_loss_epoch=0.189, valid_loss=0.282]
Epoch 73:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
00, train_loss_epoch=0.200, valid_loss=0.282]
Epoch 74:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
01, train_loss_epoch=0.201, valid_loss=0.282]
Epoch 75:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.1
90, train_loss_epoch=0.190, valid_loss=0.282]
Epoch 76:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.1
89, train_loss_epoch=0.189, valid_loss=0.282]
Epoch 77:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.1
89, train_loss_epoch=0.189, valid_loss=0.282]
Epoch 78:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.1
90, train_loss_epoch=0.190, valid_loss=0.282]
Epoch 79:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.1
82, train_loss_epoch=0.182, valid_loss=0.282]
Epoch 79: 100%|██████████| 1/1 [00:00<00:00, 11.43it/s, v_num=0, train_loss_
step=0.199, train_loss_epoch=0.182, valid_loss=0.282]
Validation: |          | 0/? [00:00<?, ?it/s]
Validation:   0%|          | 0/1 [00:00<?, ?it/s]
Validation DataLoader 0:   0%|          | 0/1 [00:00<?, ?it/s]
(_train_tune pid=6155)
Validation DataLoader 0: 100%|██████████| 1/1 [00:00<00:00,  4.79it/s]
Epoch 80:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.1
99, train_loss_epoch=0.199, valid_loss=0.320]
Epoch 81:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.1
89, train_loss_epoch=0.189, valid_loss=0.320]
Epoch 82:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
```

Loading [MathJax]/extensions/Safe.js

```
01, train_loss_epoch=0.201, valid_loss=0.320]
Epoch 83:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
03, train_loss_epoch=0.203, valid_loss=0.320]
Epoch 84:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
03, train_loss_epoch=0.203, valid_loss=0.320]
Epoch 85:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
20, train_loss_epoch=0.220, valid_loss=0.320]
Epoch 86:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.1
61, train_loss_epoch=0.161, valid_loss=0.320]
Epoch 87:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.1
74, train_loss_epoch=0.174, valid_loss=0.320]
Epoch 88:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.1
92, train_loss_epoch=0.192, valid_loss=0.320]
Epoch 89:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.1
88, train_loss_epoch=0.188, valid_loss=0.320]
Epoch 89: 100%|██████████| 1/1 [00:00<00:00, 11.96it/s, v_num=0, train_loss_
step=0.174, train_loss_epoch=0.188, valid_loss=0.320]
Validation: |          | 0/? [00:00<?, ?it/s]
Validation:   0%|          | 0/1 [00:00<?, ?it/s]
Validation DataLoader 0:   0%|          | 0/1 [00:00<?, ?it/s]
(_train_tune pid=6155)
Validation DataLoader 0: 100%|██████████| 1/1 [00:00<00:00,  4.76it/s]
Epoch 90:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.1
74, train_loss_epoch=0.174, valid_loss=0.333]
Epoch 91:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
10, train_loss_epoch=0.210, valid_loss=0.333]
Epoch 93:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.1
86, train_loss_epoch=0.186, valid_loss=0.333]
Epoch 94:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.1
89, train_loss_epoch=0.189, valid_loss=0.333]
Epoch 95:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.1
88, train_loss_epoch=0.188, valid_loss=0.333]
Epoch 96:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.1
87, train_loss_epoch=0.187, valid_loss=0.333]
Epoch 97: 100%|██████████| 1/1 [00:00<00:00, 12.33it/s, v_num=0, train_loss_
step=0.190, train_loss_epoch=0.190, valid_loss=0.333]
Epoch 98:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.1
90, train_loss_epoch=0.190, valid_loss=0.333]
Epoch 99:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
11, train_loss_epoch=0.211, valid_loss=0.333]
Epoch 99: 100%|██████████| 1/1 [00:00<00:00, 12.28it/s, v_num=0, train_loss_
step=0.191, train_loss_epoch=0.211, valid_loss=0.333]
Validation: |          | 0/? [00:00<?, ?it/s]
Validation:   0%|          | 0/1 [00:00<?, ?it/s]
Validation DataLoader 0:   0%|          | 0/1 [00:00<?, ?it/s]
(_train_tune pid=6155)
Validation DataLoader 0: 100%|██████████| 1/1 [00:00<00:00,  4.76it/s]
Epoch 99: 100%|██████████| 1/1 [00:00<00:00,  3.28it/s, v_num=0, train_loss_
step=0.191, train_loss_epoch=0.191, valid_loss=0.291]
(_train_tune pid=6155) Seed set to 9
```

Loading [MathJax]/extensions/Safe.js

```
Sanity Checking DataLoader 0:   0%|          | 0/1 [00:00<?, ?it/s]
Epoch 0:   0%|          | 0/1 [00:00<?, ?it/s]
Epoch 1:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.22
5, train_loss_epoch=0.225]
Epoch 2:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.23
1, train_loss_epoch=0.231]
Epoch 3:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.21
0, train_loss_epoch=0.210]
Epoch 4:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.23
3, train_loss_epoch=0.233]
Epoch 5:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.21
3, train_loss_epoch=0.213]
Epoch 7:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.20
5, train_loss_epoch=0.205]
Epoch 8:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.24
1, train_loss_epoch=0.241]
Epoch 9:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.19
6, train_loss_epoch=0.196]
Epoch 9: 100%|██████████| 1/1 [00:00<00:00,  9.22it/s, v_num=0, train_loss_s
tep=0.216, train_loss_epoch=0.196]
Validation: |          | 0/? [00:00<?, ?it/s]
Validation:   0%|          | 0/1 [00:00<?, ?it/s]
Validation DataLoader 0:   0%|          | 0/1 [00:00<?, ?it/s]
(_train_tune pid=6155)
Validation DataLoader 0: 100%|██████████| 1/1 [00:00<00:00,  4.73it/s]
Epoch 10:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
16, train_loss_epoch=0.216, valid_loss=0.283]
Epoch 11:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
09, train_loss_epoch=0.209, valid_loss=0.283]
Epoch 12:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
32, train_loss_epoch=0.232, valid_loss=0.283]
Epoch 13:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
15, train_loss_epoch=0.215, valid_loss=0.283]
Epoch 14:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
16, train_loss_epoch=0.216, valid_loss=0.283]
Epoch 15: 100%|██████████| 1/1 [00:00<00:00, 11.22it/s, v_num=0, train_loss_
step=0.203, train_loss_epoch=0.203, valid_loss=0.283]
Epoch 15:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
03, train_loss_epoch=0.203, valid_loss=0.283]
Epoch 16:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
03, train_loss_epoch=0.203, valid_loss=0.283]
Epoch 17:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
19, train_loss_epoch=0.219, valid_loss=0.283]
Epoch 18:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
42, train_loss_epoch=0.242, valid_loss=0.283]
Epoch 19:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
08, train_loss_epoch=0.208, valid_loss=0.283]
Epoch 19: 100%|██████████| 1/1 [00:00<00:00, 10.91it/s, v_num=0, train_loss_
step=0.214, train_loss_epoch=0.208, valid_loss=0.283]
Validation: |          | 0/? [00:00<?, ?it/s]
Validation:   0%|          | 0/1 [00:00<?, ?it/s]
Validation DataLoader 0:   0%|          | 0/1 [00:00<?, ?it/s]
(_train_tune pid=6155)
Validation DataLoader 0: 100%|██████████| 1/1 [00:00<00:00,  4.72it/s]
Epoch 20:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
14, train_loss_epoch=0.214, valid_loss=0.344]
```

Loading [MathJax]/extensions/Safe.js

```
Epoch 21:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
03, train_loss_epoch=0.203, valid_loss=0.344]
Epoch 22:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
31, train_loss_epoch=0.231, valid_loss=0.344]
Epoch 23:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
01, train_loss_epoch=0.201, valid_loss=0.344]
Epoch 25:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
13, train_loss_epoch=0.213, valid_loss=0.344]
Epoch 26:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.1
91, train_loss_epoch=0.191, valid_loss=0.344]
Epoch 27:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
14, train_loss_epoch=0.214, valid_loss=0.344]
Epoch 28:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
18, train_loss_epoch=0.218, valid_loss=0.344]
Epoch 29: 100%|██████████| 1/1 [00:00<00:00, 11.85it/s, v_num=0, train_loss_
step=0.222, train_loss_epoch=0.205, valid_loss=0.344]
(_train_tune pid=6155)
Validation: |          | 0/? [00:00<?, ?it/s]
(_train_tune pid=6155)
Validation:   0%|          | 0/1 [00:00<?, ?it/s]
Validation DataLoader 0:   0%|          | 0/1 [00:00<?, ?it/s]
(_train_tune pid=6155)
Validation DataLoader 0: 100%|██████████| 1/1 [00:00<00:00,  4.50it/s]
Epoch 30:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
22, train_loss_epoch=0.222, valid_loss=0.272]
Epoch 31:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
01, train_loss_epoch=0.201, valid_loss=0.272]
Epoch 32:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
09, train_loss_epoch=0.209, valid_loss=0.272]
Epoch 33:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
12, train_loss_epoch=0.212, valid_loss=0.272]
Epoch 34:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
39, train_loss_epoch=0.239, valid_loss=0.272]
Epoch 36:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
27, train_loss_epoch=0.227, valid_loss=0.272]
Epoch 37:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
18, train_loss_epoch=0.218, valid_loss=0.272]
Epoch 38:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
18, train_loss_epoch=0.218, valid_loss=0.272]
Epoch 39:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.1
79, train_loss_epoch=0.179, valid_loss=0.272]
Epoch 39: 100%|██████████| 1/1 [00:00<00:00, 10.65it/s, v_num=0, train_loss_
step=0.186, train_loss_epoch=0.179, valid_loss=0.272]
Validation: |          | 0/? [00:00<?, ?it/s]
Validation:   0%|          | 0/1 [00:00<?, ?it/s]
Validation DataLoader 0:   0%|          | 0/1 [00:00<?, ?it/s]
(_train_tune pid=6155)
Validation DataLoader 0: 100%|██████████| 1/1 [00:00<00:00,  4.72it/s]
Epoch 40:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.1
86, train_loss_epoch=0.186, valid_loss=0.292]
Epoch 41:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
04, train_loss_epoch=0.204, valid_loss=0.292]
Epoch 42:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
19, train_loss_epoch=0.219, valid_loss=0.292]
Epoch 43:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
17, train_loss_epoch=0.217, valid_loss=0.292]
```

Loading [MathJax]/extensions/Safe.js

```
Epoch 44:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
02, train_loss_epoch=0.202, valid_loss=0.292]
Epoch 45:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
07, train_loss_epoch=0.207, valid_loss=0.292]
Epoch 46:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
02, train_loss_epoch=0.202, valid_loss=0.292]
Epoch 47:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
02, train_loss_epoch=0.202, valid_loss=0.292]
Epoch 48:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
11, train_loss_epoch=0.211, valid_loss=0.292]
Epoch 49:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
10, train_loss_epoch=0.210, valid_loss=0.292]
Epoch 49: 100%|██████████| 1/1 [00:00<00:00, 10.62it/s, v_num=0, train_loss_
step=0.210, train_loss_epoch=0.210, valid_loss=0.292]
Validation: |          | 0/? [00:00<?, ?it/s]
Validation:   0%|          | 0/1 [00:00<?, ?it/s]
Validation DataLoader 0:   0%|          | 0/1 [00:00<?, ?it/s]
(_train_tune pid=6155)
Validation DataLoader 0: 100%|██████████| 1/1 [00:00<00:00,  4.35it/s]
Epoch 50:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
10, train_loss_epoch=0.210, valid_loss=0.302]
Epoch 51:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
02, train_loss_epoch=0.202, valid_loss=0.302]
Epoch 52:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
13, train_loss_epoch=0.213, valid_loss=0.302]
Epoch 53:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.1
89, train_loss_epoch=0.189, valid_loss=0.302]
Epoch 54:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
15, train_loss_epoch=0.215, valid_loss=0.302]
Epoch 55:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.1
89, train_loss_epoch=0.189, valid_loss=0.302]
Epoch 56:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.1
81, train_loss_epoch=0.181, valid_loss=0.302]
Epoch 57:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.1
90, train_loss_epoch=0.190, valid_loss=0.302]
Epoch 58: 100%|██████████| 1/1 [00:00<00:00, 10.87it/s, v_num=0, train_loss_
step=0.200, train_loss_epoch=0.200, valid_loss=0.302]
Epoch 59:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
00, train_loss_epoch=0.200, valid_loss=0.302]
Epoch 59: 100%|██████████| 1/1 [00:00<00:00, 11.33it/s, v_num=0, train_loss_
step=0.183, train_loss_epoch=0.200, valid_loss=0.302]
Validation: |          | 0/? [00:00<?, ?it/s]
Validation:   0%|          | 0/1 [00:00<?, ?it/s]
Validation DataLoader 0:   0%|          | 0/1 [00:00<?, ?it/s]
(_train_tune pid=6155)
Validation DataLoader 0: 100%|██████████| 1/1 [00:00<00:00,  4.97it/s]
Epoch 60:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.1
83, train_loss_epoch=0.183, valid_loss=0.280]
Epoch 61:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
03, train_loss_epoch=0.203, valid_loss=0.280]
Epoch 62:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.1
71, train_loss_epoch=0.171, valid_loss=0.280]
Epoch 63:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
09, train_loss_epoch=0.209, valid_loss=0.280]
Epoch 64:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
03, train_loss_epoch=0.203, valid_loss=0.280]
```

Loading [MathJax]/extensions/Safe.js

```
Epoch 65:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
21, train_loss_epoch=0.221, valid_loss=0.280]
Epoch 66:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
06, train_loss_epoch=0.206, valid_loss=0.280]
Epoch 67:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.1
95, train_loss_epoch=0.195, valid_loss=0.280]
Epoch 69:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
04, train_loss_epoch=0.204, valid_loss=0.280]
Epoch 69: 100%|██████████| 1/1 [00:00<00:00, 11.47it/s, v_num=0, train_loss_
step=0.187, train_loss_epoch=0.204, valid_loss=0.280]
Validation: |          | 0/? [00:00<?, ?it/s]
Validation:   0%|          | 0/1 [00:00<?, ?it/s]
Validation DataLoader 0:   0%|          | 0/1 [00:00<?, ?it/s]
(_train_tune pid=6155)
Validation DataLoader 0: 100%|██████████| 1/1 [00:00<00:00,  4.66it/s]
Epoch 70:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.1
87, train_loss_epoch=0.187, valid_loss=0.266]
Epoch 71:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
03, train_loss_epoch=0.203, valid_loss=0.266]
Epoch 72:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.1
91, train_loss_epoch=0.191, valid_loss=0.266]
Epoch 73:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.1
84, train_loss_epoch=0.184, valid_loss=0.266]
Epoch 74:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
01, train_loss_epoch=0.201, valid_loss=0.266]
Epoch 75:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.1
90, train_loss_epoch=0.190, valid_loss=0.266]
Epoch 76:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.1
89, train_loss_epoch=0.189, valid_loss=0.266]
Epoch 78:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.1
88, train_loss_epoch=0.188, valid_loss=0.266]
Epoch 79:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
23, train_loss_epoch=0.223, valid_loss=0.266]
Epoch 79: 100%|██████████| 1/1 [00:00<00:00, 12.20it/s, v_num=0, train_loss_
step=0.194, train_loss_epoch=0.223, valid_loss=0.266]
Validation: |          | 0/? [00:00<?, ?it/s]
Validation:   0%|          | 0/1 [00:00<?, ?it/s]
Validation DataLoader 0:   0%|          | 0/1 [00:00<?, ?it/s]
(_train_tune pid=6155)
Validation DataLoader 0: 100%|██████████| 1/1 [00:00<00:00,  4.63it/s]
Epoch 80:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.1
94, train_loss_epoch=0.194, valid_loss=0.295]
Epoch 81:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.1
94, train_loss_epoch=0.194, valid_loss=0.295]
Epoch 82:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.1
84, train_loss_epoch=0.184, valid_loss=0.295]
Epoch 83: 100%|██████████| 1/1 [00:00<00:00, 12.15it/s, v_num=0, train_loss_
step=0.214, train_loss_epoch=0.214, valid_loss=0.295]
Epoch 84:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
14, train_loss_epoch=0.214, valid_loss=0.295]
Epoch 85:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
13, train_loss_epoch=0.213, valid_loss=0.295]
Epoch 86:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.1
73, train_loss_epoch=0.173, valid_loss=0.295]
Epoch 87:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
13, train_loss_epoch=0.213, valid_loss=0.295]
```

Loading [MathJax]/extensions/Safe.js

```
Epoch 88:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.1
75, train_loss_epoch=0.175, valid_loss=0.295]
Epoch 89: 100%|██████████| 1/1 [00:00<00:00, 11.09it/s, v_num=0, train_loss_
step=0.202, train_loss_epoch=0.218, valid_loss=0.295]
(_train_tune pid=6155)
Validation: |          | 0/? [00:00<?, ?it/s]
Validation:   0%|          | 0/1 [00:00<?, ?it/s]
Validation DataLoader 0:   0%|          | 0/1 [00:00<?, ?it/s]
(_train_tune pid=6155)
Validation DataLoader 0: 100%|██████████| 1/1 [00:00<00:00,  4.58it/s]
Epoch 90:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
02, train_loss_epoch=0.202, valid_loss=0.324]
Epoch 91: 100%|██████████| 1/1 [00:00<00:00, 10.74it/s, v_num=0, train_loss_
step=0.203, train_loss_epoch=0.198, valid_loss=0.324]
Epoch 91: 100%|██████████| 1/1 [00:00<00:00, 10.68it/s, v_num=0, train_loss_
step=0.203, train_loss_epoch=0.203, valid_loss=0.324]
Epoch 92:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
03, train_loss_epoch=0.203, valid_loss=0.324]
Epoch 93:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
13, train_loss_epoch=0.213, valid_loss=0.324]
Epoch 94:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.1
87, train_loss_epoch=0.187, valid_loss=0.324]
Epoch 95:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.1
83, train_loss_epoch=0.183, valid_loss=0.324]
Epoch 96:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.1
75, train_loss_epoch=0.175, valid_loss=0.324]
Epoch 97:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.1
69, train_loss_epoch=0.169, valid_loss=0.324]
Epoch 98:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.1
95, train_loss_epoch=0.195, valid_loss=0.324]
Epoch 98: 100%|██████████| 1/1 [00:00<00:00, 10.91it/s, v_num=0, train_loss_
step=0.217, train_loss_epoch=0.217, valid_loss=0.324]
Epoch 99:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
17, train_loss_epoch=0.217, valid_loss=0.324]
Epoch 99: 100%|██████████| 1/1 [00:00<00:00,  6.34it/s, v_num=0, train_loss_
step=0.191, train_loss_epoch=0.217, valid_loss=0.324]
Validation: |          | 0/? [00:00<?, ?it/s]
Validation:   0%|          | 0/1 [00:00<?, ?it/s]
Validation DataLoader 0:   0%|          | 0/1 [00:00<?, ?it/s]
(_train_tune pid=6155)
Validation DataLoader 0: 100%|██████████| 1/1 [00:00<00:00,  4.43it/s]
Epoch 99: 100%|██████████| 1/1 [00:00<00:00,  2.52it/s, v_num=0, train_loss_
step=0.191, train_loss_epoch=0.191, valid_loss=0.316]
Sanity Checking DataLoader 0:   0%|          | 0/1 [00:00<?, ?it/s]
(_train_tune pid=6155) Seed set to 3
```

Loading [MathJax]/extensions/Safe.js

```
Epoch 0:   0%|          | 0/1 [00:00<?, ?it/s]
Epoch 1:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.21
1, train_loss_epoch=0.211]
Epoch 2:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.23
6, train_loss_epoch=0.236]
Epoch 3:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.22
7, train_loss_epoch=0.227]
Epoch 4:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.22
5, train_loss_epoch=0.225]
Epoch 5:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.22
3, train_loss_epoch=0.223]
Epoch 6:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.21
4, train_loss_epoch=0.214]
Epoch 8:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.20
6, train_loss_epoch=0.206]
Epoch 9:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.22
7, train_loss_epoch=0.227]
Epoch 9: 100%|██████████| 1/1 [00:00<00:00, 11.78it/s, v_num=0, train_loss_s
tep=0.219, train_loss_epoch=0.227]
Validation: |          | 0/? [00:00<?, ?it/s]
Validation:   0%|          | 0/1 [00:00<?, ?it/s]
Validation DataLoader 0:   0%|          | 0/1 [00:00<?, ?it/s]
(_train_tune pid=6155)
Validation DataLoader 0: 100%|██████████| 1/1 [00:00<00:00,  4.68it/s]
Epoch 10:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
19, train_loss_epoch=0.219, valid_loss=0.272]
Epoch 11:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
19, train_loss_epoch=0.219, valid_loss=0.272]
Epoch 12:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
18, train_loss_epoch=0.218, valid_loss=0.272]
Epoch 13:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
11, train_loss_epoch=0.211, valid_loss=0.272]
Epoch 14:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
09, train_loss_epoch=0.209, valid_loss=0.272]
Epoch 15:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
21, train_loss_epoch=0.221, valid_loss=0.272]
Epoch 16:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
32, train_loss_epoch=0.232, valid_loss=0.272]
Epoch 17:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
18, train_loss_epoch=0.218, valid_loss=0.272]
Epoch 18:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
12, train_loss_epoch=0.212, valid_loss=0.272]
Epoch 19:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
26, train_loss_epoch=0.226, valid_loss=0.272]
Epoch 19: 100%|██████████| 1/1 [00:00<00:00,  9.78it/s, v_num=0, train_loss_
step=0.210, train_loss_epoch=0.226, valid_loss=0.272]
Validation: |          | 0/? [00:00<?, ?it/s]
Validation:   0%|          | 0/1 [00:00<?, ?it/s]
Validation DataLoader 0:   0%|          | 0/1 [00:00<?, ?it/s]
(_train_tune pid=6155)
Validation DataLoader 0: 100%|██████████| 1/1 [00:00<00:00,  3.88it/s]
Epoch 20:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
10, train_loss_epoch=0.210, valid_loss=0.271]
Epoch 21:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
04, train_loss_epoch=0.204, valid_loss=0.271]
Epoch 22:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
```

Loading [MathJax]/extensions/Safe.js

```
32, train_loss_epoch=0.232, valid_loss=0.271]
Epoch 23:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
18, train_loss_epoch=0.218, valid_loss=0.271]
Epoch 24:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
34, train_loss_epoch=0.234, valid_loss=0.271]
Epoch 25:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
17, train_loss_epoch=0.217, valid_loss=0.271]
Epoch 27:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.1
96, train_loss_epoch=0.196, valid_loss=0.271]
Epoch 28:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
28, train_loss_epoch=0.228, valid_loss=0.271]
Epoch 29:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.1
85, train_loss_epoch=0.185, valid_loss=0.271]
Epoch 29: 100%|██████████| 1/1 [00:00<00:00, 11.96it/s, v_num=0, train_loss_
step=0.209, train_loss_epoch=0.185, valid_loss=0.271]
Validation: |          | 0/? [00:00<?, ?it/s]
Validation:   0%|          | 0/1 [00:00<?, ?it/s]
Validation DataLoader 0:   0%|          | 0/1 [00:00<?, ?it/s]
(_train_tune pid=6155)
Validation DataLoader 0: 100%|██████████| 1/1 [00:00<00:00,  4.14it/s]
Epoch 30:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
09, train_loss_epoch=0.209, valid_loss=0.274]
Epoch 31:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
05, train_loss_epoch=0.205, valid_loss=0.274]
Epoch 32:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
01, train_loss_epoch=0.201, valid_loss=0.274]
Epoch 34:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
12, train_loss_epoch=0.212, valid_loss=0.274]
Epoch 35:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.1
88, train_loss_epoch=0.188, valid_loss=0.274]
Epoch 36:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.1
86, train_loss_epoch=0.186, valid_loss=0.274]
Epoch 37:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
14, train_loss_epoch=0.214, valid_loss=0.274]
Epoch 38:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
04, train_loss_epoch=0.204, valid_loss=0.274]
Epoch 39:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
17, train_loss_epoch=0.217, valid_loss=0.274]
Epoch 39: 100%|██████████| 1/1 [00:00<00:00, 10.77it/s, v_num=0, train_loss_
step=0.217, train_loss_epoch=0.217, valid_loss=0.274]
Validation: |          | 0/? [00:00<?, ?it/s]
Validation:   0%|          | 0/1 [00:00<?, ?it/s]
Validation DataLoader 0:   0%|          | 0/1 [00:00<?, ?it/s]
(_train_tune pid=6155)
Validation DataLoader 0: 100%|██████████| 1/1 [00:00<00:00,  4.41it/s]
Epoch 40:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
17, train_loss_epoch=0.217, valid_loss=0.310]
Epoch 41:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
02, train_loss_epoch=0.202, valid_loss=0.310]
Epoch 42:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.1
95, train_loss_epoch=0.195, valid_loss=0.310]
Epoch 43:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.1
95, train_loss_epoch=0.195, valid_loss=0.310]
Epoch 44:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.1
92, train_loss_epoch=0.192, valid_loss=0.310]
Epoch 46:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.1
```

Loading [MathJax]/extensions/Safe.js

```
94, train_loss_epoch=0.194, valid_loss=0.310]
Epoch 47:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
14, train_loss_epoch=0.214, valid_loss=0.310]
Epoch 48:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.1
89, train_loss_epoch=0.189, valid_loss=0.310]
Epoch 49:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.1
99, train_loss_epoch=0.199, valid_loss=0.310]
Epoch 49: 100%|██████████| 1/1 [00:00<00:00, 12.12it/s, v_num=0, train_loss_
step=0.203, train_loss_epoch=0.199, valid_loss=0.310]
Validation: |          | 0/? [00:00<?, ?it/s]
Validation:   0%|          | 0/1 [00:00<?, ?it/s]
Validation DataLoader 0:   0%|          | 0/1 [00:00<?, ?it/s]
(_train_tune pid=6155)
Validation DataLoader 0: 100%|██████████| 1/1 [00:00<00:00,  5.03it/s]
Epoch 50:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
03, train_loss_epoch=0.203, valid_loss=0.324]
Epoch 51:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
18, train_loss_epoch=0.218, valid_loss=0.324]
Epoch 52:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.1
98, train_loss_epoch=0.198, valid_loss=0.324]
Epoch 53:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.1
90, train_loss_epoch=0.190, valid_loss=0.324]
Epoch 54:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
12, train_loss_epoch=0.212, valid_loss=0.324]
Epoch 56:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
05, train_loss_epoch=0.205, valid_loss=0.324]
Epoch 57:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.1
76, train_loss_epoch=0.176, valid_loss=0.324]
Epoch 59:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.1
94, train_loss_epoch=0.194, valid_loss=0.324]
Epoch 59: 100%|██████████| 1/1 [00:00<00:00, 11.38it/s, v_num=0, train_loss_
step=0.190, train_loss_epoch=0.194, valid_loss=0.324]
Validation: |          | 0/? [00:00<?, ?it/s]
Validation:   0%|          | 0/1 [00:00<?, ?it/s]
Validation DataLoader 0:   0%|          | 0/1 [00:00<?, ?it/s]
(_train_tune pid=6155)
Validation DataLoader 0: 100%|██████████| 1/1 [00:00<00:00,  4.69it/s]
Epoch 60:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.1
90, train_loss_epoch=0.190, valid_loss=0.320]
Epoch 61:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.1
98, train_loss_epoch=0.198, valid_loss=0.320]
Epoch 62:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.1
88, train_loss_epoch=0.188, valid_loss=0.320]
Epoch 63:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.1
86, train_loss_epoch=0.186, valid_loss=0.320]
Epoch 65:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.1
91, train_loss_epoch=0.191, valid_loss=0.320]
Epoch 66:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
08, train_loss_epoch=0.208, valid_loss=0.320]
Epoch 67:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.1
79, train_loss_epoch=0.179, valid_loss=0.320]
Epoch 68:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
01, train_loss_epoch=0.201, valid_loss=0.320]
Epoch 69:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
01, train_loss_epoch=0.201, valid_loss=0.320]
Epoch 69: 100%|██████████| 1/1 [00:00<00:00, 11.41it/s, v_num=0, train_loss_
```

Loading [MathJax]/extensions/Safe.js

```
step=0.196, train_loss_epoch=0.201, valid_loss=0.320]
Validation: |          | 0/? [00:00<?, ?it/s]
Validation:   0%|          | 0/1 [00:00<?, ?it/s]
Validation DataLoader 0:   0%|          | 0/1 [00:00<?, ?it/s]
(_train_tune pid=6155)
Validation DataLoader 0: 100%|██████████| 1/1 [00:00<00:00,  5.07it/s]
Epoch 70:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.1
96, train_loss_epoch=0.196, valid_loss=0.304]
Epoch 71:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.1
92, train_loss_epoch=0.192, valid_loss=0.304]
Epoch 72:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.1
91, train_loss_epoch=0.191, valid_loss=0.304]
Epoch 74:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.1
97, train_loss_epoch=0.197, valid_loss=0.304]
Epoch 75:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.1
80, train_loss_epoch=0.180, valid_loss=0.304]
Epoch 76:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
06, train_loss_epoch=0.206, valid_loss=0.304]
Epoch 77:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.1
76, train_loss_epoch=0.176, valid_loss=0.304]
Epoch 78:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.1
84, train_loss_epoch=0.184, valid_loss=0.304]
Epoch 79: 100%|██████████| 1/1 [00:00<00:00, 12.66it/s, v_num=0, train_loss_
step=0.193, train_loss_epoch=0.165, valid_loss=0.304]
Validation: |          | 0/? [00:00<?, ?it/s]
Validation:   0%|          | 0/1 [00:00<?, ?it/s]
Validation DataLoader 0:   0%|          | 0/1 [00:00<?, ?it/s]
(_train_tune pid=6155)
Validation DataLoader 0: 100%|██████████| 1/1 [00:00<00:00,  4.93it/s]
Epoch 79: 100%|██████████| 1/1 [00:00<00:00,  3.49it/s, v_num=0, train_loss_
step=0.193, train_loss_epoch=0.165, valid_loss=0.296]
Epoch 81:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.1
86, train_loss_epoch=0.186, valid_loss=0.296]
Epoch 82:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.1
86, train_loss_epoch=0.186, valid_loss=0.296]
Epoch 83:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.1
82, train_loss_epoch=0.182, valid_loss=0.296]
Epoch 84:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.1
77, train_loss_epoch=0.177, valid_loss=0.296]
Epoch 85:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.1
89, train_loss_epoch=0.189, valid_loss=0.296]
Epoch 86:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.1
72, train_loss_epoch=0.172, valid_loss=0.296]
Epoch 87:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.1
74, train_loss_epoch=0.174, valid_loss=0.296]
Epoch 88:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.1
85, train_loss_epoch=0.185, valid_loss=0.296]
Epoch 89:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.1
83, train_loss_epoch=0.183, valid_loss=0.296]
Epoch 89: 100%|██████████| 1/1 [00:00<00:00, 11.53it/s, v_num=0, train_loss_
step=0.166, train_loss_epoch=0.183, valid_loss=0.296]
Validation: |          | 0/? [00:00<?, ?it/s]
Validation:   0%|          | 0/1 [00:00<?, ?it/s]
Validation DataLoader 0:   0%|          | 0/1 [00:00<?, ?it/s]
(_train_tune pid=6155)
Validation DataLoader 0: 100%|██████████| 1/1 [00:00<00:00,  4.86it/s]
```

Loading [MathJax]/extensions/Safe.js

```
Epoch 90:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.1
66, train_loss_epoch=0.166, valid_loss=0.338]
Epoch 91:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.1
86, train_loss_epoch=0.186, valid_loss=0.338]
Epoch 92:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.1
80, train_loss_epoch=0.180, valid_loss=0.338]
Epoch 93:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.1
89, train_loss_epoch=0.189, valid_loss=0.338]
Epoch 94:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.1
76, train_loss_epoch=0.176, valid_loss=0.338]
Epoch 95:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.1
81, train_loss_epoch=0.181, valid_loss=0.338]
Epoch 96:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
04, train_loss_epoch=0.204, valid_loss=0.338]
Epoch 97:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.1
82, train_loss_epoch=0.182, valid_loss=0.338]
Epoch 98:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.1
88, train_loss_epoch=0.188, valid_loss=0.338]
Epoch 99:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.1
97, train_loss_epoch=0.197, valid_loss=0.338]
Epoch 99: 100%|██████████| 1/1 [00:00<00:00, 11.36it/s, v_num=0, train_loss_
step=0.190, train_loss_epoch=0.197, valid_loss=0.338]
Validation: |          | 0/? [00:00<?, ?it/s]
Validation:   0%|          | 0/1 [00:00<?, ?it/s]
Validation DataLoader 0:   0%|          | 0/1 [00:00<?, ?it/s]
(_train_tune pid=6155)
Validation DataLoader 0: 100%|██████████| 1/1 [00:00<00:00,  4.43it/s]
Epoch 99: 100%|██████████| 1/1 [00:00<00:00,  3.05it/s, v_num=0, train_loss_
step=0.190, train_loss_epoch=0.190, valid_loss=0.370]
(_train_tune pid=6155) Seed set to 6
```

Loading [MathJax]/extensions/Safe.js

```
Sanity Checking DataLoader 0:   0%|          | 0/1 [00:00<?, ?it/s]
Epoch 0:   0%|          | 0/1 [00:00<?, ?it/s]
Epoch 1:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.20
4, train_loss_epoch=0.204]
Epoch 2:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.23
3, train_loss_epoch=0.233]
Epoch 3:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.23
7, train_loss_epoch=0.237]
Epoch 4:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.20
7, train_loss_epoch=0.207]
Epoch 5:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.20
7, train_loss_epoch=0.207]
Epoch 6:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.22
9, train_loss_epoch=0.229]
Epoch 7:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.24
8, train_loss_epoch=0.248]
Epoch 9:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.20
7, train_loss_epoch=0.207]
Epoch 9: 100%|██████████| 1/1 [00:00<00:00, 10.77it/s, v_num=0, train_loss_s
tep=0.229, train_loss_epoch=0.207]
Validation: |          | 0/? [00:00<?, ?it/s]
Validation:   0%|          | 0/1 [00:00<?, ?it/s]
Validation DataLoader 0:   0%|          | 0/1 [00:00<?, ?it/s]
(_train_tune pid=6155)
Validation DataLoader 0: 100%|██████████| 1/1 [00:00<00:00,  4.50it/s]
Epoch 11:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
07, train_loss_epoch=0.207, valid_loss=0.264]
Epoch 12:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
04, train_loss_epoch=0.204, valid_loss=0.264]
Epoch 13:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
28, train_loss_epoch=0.228, valid_loss=0.264]
Epoch 14:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
49, train_loss_epoch=0.249, valid_loss=0.264]
Epoch 16:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.1
82, train_loss_epoch=0.182, valid_loss=0.264]
Epoch 17:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
05, train_loss_epoch=0.205, valid_loss=0.264]
Epoch 18:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.1
99, train_loss_epoch=0.199, valid_loss=0.264]
Epoch 19:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
18, train_loss_epoch=0.218, valid_loss=0.264]
Epoch 19: 100%|██████████| 1/1 [00:00<00:00, 12.24it/s, v_num=0, train_loss_
step=0.204, train_loss_epoch=0.218, valid_loss=0.264]
Validation: |          | 0/? [00:00<?, ?it/s]
Validation:   0%|          | 0/1 [00:00<?, ?it/s]
Validation DataLoader 0:   0%|          | 0/1 [00:00<?, ?it/s]
(_train_tune pid=6155)
Validation DataLoader 0: 100%|██████████| 1/1 [00:00<00:00,  4.53it/s]
Epoch 20:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
04, train_loss_epoch=0.204, valid_loss=0.309]
Epoch 21:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
11, train_loss_epoch=0.211, valid_loss=0.309]
Epoch 22:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
14, train_loss_epoch=0.214, valid_loss=0.309]
Epoch 23:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
12, train_loss_epoch=0.212, valid_loss=0.309]
```

Loading [MathJax]/extensions/Safe.js

```
Epoch 24:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
08, train_loss_epoch=0.208, valid_loss=0.309]
Epoch 25:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
14, train_loss_epoch=0.214, valid_loss=0.309]
Epoch 26:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
15, train_loss_epoch=0.215, valid_loss=0.309]
Epoch 28:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.1
97, train_loss_epoch=0.197, valid_loss=0.309]
Epoch 29:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
19, train_loss_epoch=0.219, valid_loss=0.309]
Epoch 29: 100%|██████████| 1/1 [00:00<00:00, 11.52it/s, v_num=0, train_loss_
step=0.219, train_loss_epoch=0.219, valid_loss=0.309]
Validation: |          | 0/? [00:00<?, ?it/s]
Validation:   0%|          | 0/1 [00:00<?, ?it/s]
Validation DataLoader 0:   0%|          | 0/1 [00:00<?, ?it/s]
(_train_tune pid=6155)
Validation DataLoader 0: 100%|██████████| 1/1 [00:00<00:00,  4.83it/s]
Epoch 30:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
19, train_loss_epoch=0.219, valid_loss=0.334]
Epoch 31:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
14, train_loss_epoch=0.214, valid_loss=0.334]
Epoch 32:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
03, train_loss_epoch=0.203, valid_loss=0.334]
Epoch 33:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
01, train_loss_epoch=0.201, valid_loss=0.334]
Epoch 34:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
24, train_loss_epoch=0.224, valid_loss=0.334]
Epoch 35:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
10, train_loss_epoch=0.210, valid_loss=0.334]
Epoch 36:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.1
97, train_loss_epoch=0.197, valid_loss=0.334]
Epoch 36: 100%|██████████| 1/1 [00:00<00:00,  7.97it/s, v_num=0, train_loss_
step=0.201, train_loss_epoch=0.201, valid_loss=0.334]
Epoch 37:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
01, train_loss_epoch=0.201, valid_loss=0.334]
Epoch 38:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
09, train_loss_epoch=0.209, valid_loss=0.334]
Epoch 39:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
08, train_loss_epoch=0.208, valid_loss=0.334]
Epoch 39: 100%|██████████| 1/1 [00:00<00:00, 12.33it/s, v_num=0, train_loss_
step=0.194, train_loss_epoch=0.208, valid_loss=0.334]
Validation: |          | 0/? [00:00<?, ?it/s]
Validation:   0%|          | 0/1 [00:00<?, ?it/s]
Validation DataLoader 0:   0%|          | 0/1 [00:00<?, ?it/s]
(_train_tune pid=6155)
Validation DataLoader 0: 100%|██████████| 1/1 [00:00<00:00,  4.45it/s]
Epoch 40:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.1
94, train_loss_epoch=0.194, valid_loss=0.265]
Epoch 41:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
03, train_loss_epoch=0.203, valid_loss=0.265]
Epoch 42:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
08, train_loss_epoch=0.208, valid_loss=0.265]
Epoch 43:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
03, train_loss_epoch=0.203, valid_loss=0.265]
Epoch 44: 100%|██████████| 1/1 [00:00<00:00, 11.14it/s, v_num=0, train_loss_
step=0.197, train_loss_epoch=0.197, valid_loss=0.265]
```

Loading [MathJax]/extensions/Safe.js

```
Epoch 45:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.1
97, train_loss_epoch=0.197, valid_loss=0.265]
Epoch 46:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
15, train_loss_epoch=0.215, valid_loss=0.265]
Epoch 47:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.1
92, train_loss_epoch=0.192, valid_loss=0.265]
Epoch 48:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
10, train_loss_epoch=0.210, valid_loss=0.265]
Epoch 49:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
16, train_loss_epoch=0.216, valid_loss=0.265]
Epoch 49: 100%|██████████| 1/1 [00:00<00:00, 11.18it/s, v_num=0, train_loss_
step=0.196, train_loss_epoch=0.216, valid_loss=0.265]
Validation: |          | 0/? [00:00<?, ?it/s]
Validation:   0%|          | 0/1 [00:00<?, ?it/s]
Validation DataLoader 0:   0%|          | 0/1 [00:00<?, ?it/s]
(_train_tune pid=6155)
Validation DataLoader 0: 100%|██████████| 1/1 [00:00<00:00,  4.45it/s]
Epoch 50:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.1
96, train_loss_epoch=0.196, valid_loss=0.279]
Epoch 51:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.1
95, train_loss_epoch=0.195, valid_loss=0.279]
Epoch 52:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.1
97, train_loss_epoch=0.197, valid_loss=0.279]
Epoch 53:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.1
93, train_loss_epoch=0.193, valid_loss=0.279]
Epoch 54:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.1
87, train_loss_epoch=0.187, valid_loss=0.279]
Epoch 55:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
05, train_loss_epoch=0.205, valid_loss=0.279]
Epoch 56:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.1
95, train_loss_epoch=0.195, valid_loss=0.279]
Epoch 58:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
08, train_loss_epoch=0.208, valid_loss=0.279]
Epoch 59:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.1
94, train_loss_epoch=0.194, valid_loss=0.279]
Epoch 59: 100%|██████████| 1/1 [00:00<00:00, 11.79it/s, v_num=0, train_loss_
step=0.198, train_loss_epoch=0.194, valid_loss=0.279]
Validation: |          | 0/? [00:00<?, ?it/s]
Validation:   0%|          | 0/1 [00:00<?, ?it/s]
Validation DataLoader 0:   0%|          | 0/1 [00:00<?, ?it/s]
(_train_tune pid=6155)
Validation DataLoader 0: 100%|██████████| 1/1 [00:00<00:00,  4.60it/s]
Epoch 60:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.1
98, train_loss_epoch=0.198, valid_loss=0.314]
Epoch 61:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.1
98, train_loss_epoch=0.198, valid_loss=0.314]
Epoch 62:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.1
84, train_loss_epoch=0.184, valid_loss=0.314]
Epoch 63:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.1
91, train_loss_epoch=0.191, valid_loss=0.314]
Epoch 64:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.1
92, train_loss_epoch=0.192, valid_loss=0.314]
Epoch 65:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.1
88, train_loss_epoch=0.188, valid_loss=0.314]
Epoch 67:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.1
92, train_loss_epoch=0.192, valid_loss=0.314]
```

Loading [MathJax]/extensions/Safe.js

```
Epoch 68:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
30, train_loss_epoch=0.230, valid_loss=0.314]
Epoch 69:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.1
84, train_loss_epoch=0.184, valid_loss=0.314]
Epoch 69: 100%|██████████| 1/1 [00:00<00:00, 10.85it/s, v_num=0, train_loss_
step=0.175, train_loss_epoch=0.184, valid_loss=0.314]
Validation: |          | 0/? [00:00<?, ?it/s]
Validation:   0%|          | 0/1 [00:00<?, ?it/s]
Validation DataLoader 0:   0%|          | 0/1 [00:00<?, ?it/s]
(_train_tune pid=6155)
Validation DataLoader 0: 100%|██████████| 1/1 [00:00<00:00,  4.82it/s]
Epoch 70:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.1
75, train_loss_epoch=0.175, valid_loss=0.275]
Epoch 72:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.1
88, train_loss_epoch=0.188, valid_loss=0.275]
Epoch 73:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.1
89, train_loss_epoch=0.189, valid_loss=0.275]
Epoch 74:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.1
77, train_loss_epoch=0.177, valid_loss=0.275]
Epoch 75:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.1
95, train_loss_epoch=0.195, valid_loss=0.275]
Epoch 76:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
14, train_loss_epoch=0.214, valid_loss=0.275]
Epoch 77:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
05, train_loss_epoch=0.205, valid_loss=0.275]
Epoch 79:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
01, train_loss_epoch=0.201, valid_loss=0.275]
Epoch 79: 100%|██████████| 1/1 [00:00<00:00, 10.51it/s, v_num=0, train_loss_
step=0.199, train_loss_epoch=0.201, valid_loss=0.275]
Validation: |          | 0/? [00:00<?, ?it/s]
Validation:   0%|          | 0/1 [00:00<?, ?it/s]
Validation DataLoader 0:   0%|          | 0/1 [00:00<?, ?it/s]
(_train_tune pid=6155)
Validation DataLoader 0: 100%|██████████| 1/1 [00:00<00:00,  4.51it/s]
Epoch 80:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.1
99, train_loss_epoch=0.199, valid_loss=0.379]
Epoch 82:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.1
90, train_loss_epoch=0.190, valid_loss=0.379]
Epoch 83:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
10, train_loss_epoch=0.210, valid_loss=0.379]
Epoch 84:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.1
96, train_loss_epoch=0.196, valid_loss=0.379]
Epoch 85:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.1
87, train_loss_epoch=0.187, valid_loss=0.379]
Epoch 86:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.1
92, train_loss_epoch=0.192, valid_loss=0.379]
Epoch 88:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
02, train_loss_epoch=0.202, valid_loss=0.379]
Epoch 89:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.1
84, train_loss_epoch=0.184, valid_loss=0.379]
Epoch 89: 100%|██████████| 1/1 [00:00<00:00, 12.59it/s, v_num=0, train_loss_
step=0.195, train_loss_epoch=0.184, valid_loss=0.379]
Validation: |          | 0/? [00:00<?, ?it/s]
Validation:   0%|          | 0/1 [00:00<?, ?it/s]
Validation DataLoader 0:   0%|          | 0/1 [00:00<?, ?it/s]
(_train_tune pid=6155)
```

Loading [MathJax]/extensions/Safe.js

```
Validation DataLoader 0: 100%|██████████| 1/1 [00:00<00:00,  4.21it/s]
Epoch 90:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.1
95, train_loss_epoch=0.195, valid_loss=0.320]
Epoch 91:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
03, train_loss_epoch=0.203, valid_loss=0.320]
Epoch 92:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
06, train_loss_epoch=0.206, valid_loss=0.320]
Epoch 93:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.1
98, train_loss_epoch=0.198, valid_loss=0.320]
Epoch 94: 100%|██████████| 1/1 [00:00<00:00, 11.47it/s, v_num=0, train_loss_
step=0.173, train_loss_epoch=0.173, valid_loss=0.320]
Epoch 95:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.1
73, train_loss_epoch=0.173, valid_loss=0.320]
Epoch 96:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.1
72, train_loss_epoch=0.172, valid_loss=0.320]
Epoch 97:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
04, train_loss_epoch=0.204, valid_loss=0.320]
Epoch 98:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.1
98, train_loss_epoch=0.198, valid_loss=0.320]
Epoch 99:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
07, train_loss_epoch=0.207, valid_loss=0.320]
Seed set to 7
Epoch 99: 100%|██████████| 1/1 [00:00<00:00, 10.58it/s, v_num=0, train_loss_
step=0.189, train_loss_epoch=0.207, valid_loss=0.320]
Validation: |          | 0/? [00:00<?, ?it/s]
Validation:   0%|          | 0/1 [00:00<?, ?it/s]
Validation DataLoader 0:   0%|          | 0/1 [00:00<?, ?it/s]
(_train_tune pid=6155)
Validation DataLoader 0: 100%|██████████| 1/1 [00:00<00:00,  4.67it/s]
Epoch 99: 100%|██████████| 1/1 [00:00<00:00,  3.10it/s, v_num=0, train_loss_
step=0.189, train_loss_epoch=0.189, valid_loss=0.355]
2023-11-04 14:22:36.024653: I tensorflow/core/util/port.cc:111] oneDNN custo
m operations are on. You may see slightly different numerical results due to
floating-point round-off errors from different computation orders. To turn t
hem off, set the environment variable `TF_ENABLE_ONEDNN_OPTS=0`.
2023-11-04 14:22:36.026315: I tensorflow/tsl/cuda/cudart_stub.cc:28] Could n
ot find cuda drivers on your machine, GPU will not be used.
2023-11-04 14:22:36.050932: E tensorflow/compiler/xla/stream_executor/cuda/c
uda_dnn.cc:9342] Unable to register cuDNN factory: Attempting to register fa
ctory for plugin cuDNN when one has already been registered
2023-11-04 14:22:36.050955: E tensorflow/compiler/xla/stream_executor/cuda/c
uda_fft.cc:609] Unable to register cuFFT factory: Attempting to register fac
tory for plugin cuFFT when one has already been registered
2023-11-04 14:22:36.050972: E tensorflow/compiler/xla/stream_executor/cuda/c
uda_blas.cc:1518] Unable to register cuBLAS factory: Attempting to register
factory for plugin cuBLAS when one has already been registered
2023-11-04 14:22:36.058040: I tensorflow/core/platform/cpu_feature_guard.cc:
182] This TensorFlow binary is optimized to use available CPU instructions i
n performance-critical operations.
To enable the following instructions: AVX2 AVX_VNNI FMA, in other operation
s, rebuild TensorFlow with the appropriate compiler flags.
2023-11-04 14:22:36.624180: W tensorflow/compiler/tf2tensorrt/utils/py_util
s.cc:38] TF-TRT Warning: Could not find TensorRT
Sanity Checking: |
| 0/? [00:00…
```

Loading [MathJax]/extensions/Safe.js

```
Training: |
| 0/? [00:00…
Validation: |
| 0/? [00:00…
Validation: |
| 0/? [00:00…
Validation: |
| 0/? [00:00…
Validation: |
| 0/? [00:00…
Validation: |
| 0/? [00:00…
Validation: |
| 0/? [00:00…
Validation: |
| 0/? [00:00…
Validation: |
| 0/? [00:00…
Validation: |
| 0/? [00:00…
Validation: |
| 0/? [00:00…
Validation: |
| 0/? [00:00…
Predicting: |
| 0/? [00:00…
```

In [5]:

```
nf = NeuralForecast(
    models=models,
    freq='D')

Y_hat_df = nf.cross_validation(df=Y_df, val_size=val_size,
                               test_size=test_size, n_windows=None)
```

Loading [MathJax]/extensions/Safe.js

```
(_train_tune pid=7588) Seed set to 8
(_train_tune pid=7588) 2023-11-02 03:18:33.048410: I tensorflow/core/util/port.cc:111] oneDNN custom operations are on. You may see slightly different numerical results due to floating-point round-off errors from different computation orders. To turn them off, set the environment variable `TF_ENABLE_ONEDNN_OPTS=0`.
(_train_tune pid=7588) 2023-11-02 03:18:33.081815: I tensorflow/tsl/cuda/cudart_stub.cc:28] Could not find cuda drivers on your machine, GPU will not be used.
(_train_tune pid=7588) 2023-11-02 03:18:33.234842: E tensorflow/compiler/xla/stream_executor/cuda/cuda_dnn.cc:9342] Unable to register cuDNN factory: Attempting to register factory for plugin cuDNN when one has already been registered
(_train_tune pid=7588) 2023-11-02 03:18:33.235007: E tensorflow/compiler/xla/stream_executor/cuda/cuda_fft.cc:609] Unable to register cuFFT factory: Attempting to register factory for plugin cuFFT when one has already been registered
(_train_tune pid=7588) 2023-11-02 03:18:33.235157: E tensorflow/compiler/xla/stream_executor/cuda/cuda_blas.cc:1518] Unable to register cuBLAS factory: Attempting to register factory for plugin cuBLAS when one has already been registered
(_train_tune pid=7588) 2023-11-02 03:18:33.357964: I tensorflow/core/platform/cpu_feature_guard.cc:182] This TensorFlow binary is optimized to use available CPU instructions in performance-critical operations.
(_train_tune pid=7588) To enable the following instructions: AVX2 AVX_VNNI FMA, in other operations, rebuild TensorFlow with the appropriate compiler flags.
(_train_tune pid=7588) 2023-11-02 03:18:37.752386: W tensorflow/compiler/tf2tensorrt/utils/py_utils.cc:38] TF-TRT Warning: Could not find TensorRT
```

Loading [MathJax]/extensions/Safe.js

```
Sanity Checking DataLoader 0:   0%|          | 0/1 [00:00<?, ?it/s]
Sanity Checking DataLoader 0: 100%|██████████| 1/1 [00:01<00:00,  0.88it/s]
Epoch 0:   0%|          | 0/1 [00:00<?, ?it/s]
Epoch 1:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.22
4, train_loss_epoch=0.224]
Epoch 2:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.21
5, train_loss_epoch=0.215]
Epoch 3:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.21
5, train_loss_epoch=0.215]
Epoch 4:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.19
4, train_loss_epoch=0.194]
Epoch 5:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.20
7, train_loss_epoch=0.207]
Epoch 6:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.20
9, train_loss_epoch=0.209]
Epoch 7:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.23
0, train_loss_epoch=0.230]
Epoch 8:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.21
0, train_loss_epoch=0.210]
Epoch 9:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.23
0, train_loss_epoch=0.230]
Epoch 10:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
21, train_loss_epoch=0.221]
Epoch 11:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
10, train_loss_epoch=0.210]
Epoch 11:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.1
99, train_loss_epoch=0.199]
Epoch 12:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.1
99, train_loss_epoch=0.199]
Epoch 13:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
18, train_loss_epoch=0.218]
Epoch 14:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
03, train_loss_epoch=0.203]
Epoch 14: 100%|██████████| 1/1 [00:00<00:00,  1.76it/s, v_num=0, train_loss_
step=0.203, train_loss_epoch=0.203]
Epoch 14: 100%|██████████| 1/1 [00:00<00:00,  1.76it/s, v_num=0, train_loss_
step=0.231, train_loss_epoch=0.231]
Epoch 15:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
31, train_loss_epoch=0.231]
Epoch 15:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
15, train_loss_epoch=0.215]
Epoch 16:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
15, train_loss_epoch=0.215]
Epoch 17:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
12, train_loss_epoch=0.212]
Epoch 18:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.1
83, train_loss_epoch=0.183]
Epoch 19:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.1
94, train_loss_epoch=0.194]
Epoch 20:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
28, train_loss_epoch=0.228]
Epoch 21:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
14, train_loss_epoch=0.214]
Epoch 22:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
07, train_loss_epoch=0.207]
Epoch 23:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
```

Loading [MathJax]/extensions/Safe.js

```
61, train_loss_epoch=0.261]
Epoch 24:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
16, train_loss_epoch=0.216]
Epoch 25:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
08, train_loss_epoch=0.208]
Epoch 26:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
04, train_loss_epoch=0.204]
Epoch 27:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.1
94, train_loss_epoch=0.194]
Epoch 28:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.1
92, train_loss_epoch=0.192]
Epoch 29:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
19, train_loss_epoch=0.219]
Epoch 30:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.1
98, train_loss_epoch=0.198]
Epoch 31:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.1
88, train_loss_epoch=0.188]
Epoch 32:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
02, train_loss_epoch=0.202]
Epoch 33:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
14, train_loss_epoch=0.214]
Epoch 34:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
01, train_loss_epoch=0.201]
Epoch 35:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.1
99, train_loss_epoch=0.199]
Epoch 36:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
17, train_loss_epoch=0.217]
Epoch 37:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
02, train_loss_epoch=0.202]
Epoch 38:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
21, train_loss_epoch=0.221]
Epoch 39:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
15, train_loss_epoch=0.215]
Epoch 40:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
36, train_loss_epoch=0.236]
Epoch 41:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
10, train_loss_epoch=0.210]
Epoch 42:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
10, train_loss_epoch=0.210]
Epoch 43:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
31, train_loss_epoch=0.231]
Epoch 44:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
04, train_loss_epoch=0.204]
Epoch 45:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
06, train_loss_epoch=0.206]
Epoch 46:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.1
96, train_loss_epoch=0.196]
Epoch 47:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.1
83, train_loss_epoch=0.183]
Epoch 48:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.1
89, train_loss_epoch=0.189]
Epoch 49:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.1
92, train_loss_epoch=0.192]
Epoch 50:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.1
97, train_loss_epoch=0.197]
Epoch 51:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.1
```

Loading [MathJax]/extensions/Safe.js

```
96, train_loss_epoch=0.196]
Epoch 52:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.1
87, train_loss_epoch=0.187]
Epoch 53:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.1
84, train_loss_epoch=0.184]
Epoch 54:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
00, train_loss_epoch=0.200]
Epoch 55:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
04, train_loss_epoch=0.204]
Epoch 56:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.1
96, train_loss_epoch=0.196]
Epoch 57:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
16, train_loss_epoch=0.216]
Epoch 58:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
33, train_loss_epoch=0.233]
Epoch 59:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.1
84, train_loss_epoch=0.184]
Epoch 60:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
37, train_loss_epoch=0.237]
Epoch 61:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
16, train_loss_epoch=0.216]
Epoch 62:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.1
83, train_loss_epoch=0.183]
Epoch 63:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
00, train_loss_epoch=0.200]
Epoch 64:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.1
94, train_loss_epoch=0.194]
Epoch 65:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
13, train_loss_epoch=0.213]
Epoch 66:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
18, train_loss_epoch=0.218]
Epoch 67:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.1
84, train_loss_epoch=0.184]
Epoch 68:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
09, train_loss_epoch=0.209]
Epoch 69:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
27, train_loss_epoch=0.227]
Epoch 70:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
11, train_loss_epoch=0.211]
Epoch 71:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
12, train_loss_epoch=0.212]
Epoch 72:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
20, train_loss_epoch=0.220]
Epoch 73:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
10, train_loss_epoch=0.210]
Epoch 74:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.1
89, train_loss_epoch=0.189]
Epoch 75:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.1
96, train_loss_epoch=0.196]
Epoch 76:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.1
84, train_loss_epoch=0.184]
Epoch 77:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.1
83, train_loss_epoch=0.183]
Epoch 78:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.1
99, train_loss_epoch=0.199]
Epoch 79:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.1
```

Loading [MathJax]/extensions/Safe.js

```
97, train_loss_epoch=0.197]
Epoch 80:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.1
84, train_loss_epoch=0.184]
Epoch 81:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.1
90, train_loss_epoch=0.190]
Epoch 82:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
06, train_loss_epoch=0.206]
Epoch 83:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.1
99, train_loss_epoch=0.199]
Epoch 84:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
00, train_loss_epoch=0.200]
Epoch 85:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.1
98, train_loss_epoch=0.198]
Epoch 86:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.1
95, train_loss_epoch=0.195]
Epoch 87:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
03, train_loss_epoch=0.203]
Epoch 88:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.1
81, train_loss_epoch=0.181]
Epoch 89:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
03, train_loss_epoch=0.203]
Epoch 90:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.1
99, train_loss_epoch=0.199]
Epoch 91:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
10, train_loss_epoch=0.210]
Epoch 92:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.1
82, train_loss_epoch=0.182]
Epoch 93:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.1
80, train_loss_epoch=0.180]
Epoch 94:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
06, train_loss_epoch=0.206]
Epoch 95:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.1
95, train_loss_epoch=0.195]
Epoch 96:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
00, train_loss_epoch=0.200]
Epoch 97:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
05, train_loss_epoch=0.205]
Epoch 98:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.1
98, train_loss_epoch=0.198]
Epoch 99:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.1
90, train_loss_epoch=0.190]
Epoch 99: 100%|██████████| 1/1 [00:00<00:00,  5.85it/s, v_num=0, train_loss_
step=0.199, train_loss_epoch=0.190]
Validation: |          | 0/? [00:00<?, ?it/s]
(_train_tune pid=7588)
Validation:   0%|          | 0/1 [00:00<?, ?it/s]
Validation DataLoader 0:   0%|          | 0/1 [00:00<?, ?it/s]
(_train_tune pid=7588)
Validation DataLoader 0: 100%|██████████| 1/1 [00:00<00:00,  1.45it/s]
Epoch 99: 100%|██████████| 1/1 [00:00<00:00,  1.09it/s, v_num=0, train_loss_
step=0.199, train_loss_epoch=0.199, valid_loss=0.339]
(_train_tune pid=7588) Seed set to 3
```

Loading [MathJax]/extensions/Safe.js

```
Sanity Checking DataLoader 0:   0%|          | 0/1 [00:00<?, ?it/s]
Epoch 0:   0%|          | 0/1 [00:00<?, ?it/s]
Epoch 1:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.21
0, train_loss_epoch=0.210]
Epoch 2:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.24
1, train_loss_epoch=0.241]
Epoch 2: 100%|██████████| 1/1 [00:00<00:00,  2.87it/s, v_num=0, train_loss_s
tep=0.241, train_loss_epoch=0.241]
Epoch 2: 100%|██████████| 1/1 [00:00<00:00,  2.86it/s, v_num=0, train_loss_s
tep=0.227, train_loss_epoch=0.241]
Epoch 3:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.22
7, train_loss_epoch=0.227]
Epoch 4:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.22
4, train_loss_epoch=0.224]
Epoch 5:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.22
3, train_loss_epoch=0.223]
Epoch 6:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.21
4, train_loss_epoch=0.214]
Epoch 7:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.24
9, train_loss_epoch=0.249]
Epoch 8:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.20
6, train_loss_epoch=0.206]
Epoch 9:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.22
7, train_loss_epoch=0.227]
Epoch 10:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
19, train_loss_epoch=0.219]
Epoch 11:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
02, train_loss_epoch=0.202]
Epoch 12:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
16, train_loss_epoch=0.216]
Epoch 13:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
27, train_loss_epoch=0.227]
Epoch 14:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
14, train_loss_epoch=0.214]
Epoch 15:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
25, train_loss_epoch=0.225]
Epoch 16:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
21, train_loss_epoch=0.221]
Epoch 17:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.1
98, train_loss_epoch=0.198]
Epoch 18:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
14, train_loss_epoch=0.214]
Epoch 19:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
00, train_loss_epoch=0.200]
Epoch 20:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
34, train_loss_epoch=0.234]
Epoch 20: 100%|██████████| 1/1 [00:00<00:00,  1.65it/s, v_num=0, train_loss_
step=0.209, train_loss_epoch=0.209]
Epoch 21:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
09, train_loss_epoch=0.209]
Epoch 22:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
17, train_loss_epoch=0.217]
Epoch 23:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
03, train_loss_epoch=0.203]
Epoch 24:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
21, train_loss_epoch=0.231]
```

Loading [MathJax]/extensions/Safe.js

```
Epoch 25:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
07, train_loss_epoch=0.207]
Epoch 26:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
14, train_loss_epoch=0.214]
Epoch 27:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.1
99, train_loss_epoch=0.199]
Epoch 28:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.1
78, train_loss_epoch=0.178]
Epoch 29:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
32, train_loss_epoch=0.232]
Epoch 30:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.1
90, train_loss_epoch=0.190]
Epoch 31:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
05, train_loss_epoch=0.205]
Epoch 32:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
14, train_loss_epoch=0.214]
Epoch 33:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
03, train_loss_epoch=0.203]
Epoch 34:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
16, train_loss_epoch=0.216]
Epoch 35:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.1
80, train_loss_epoch=0.180]
Epoch 36:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.1
91, train_loss_epoch=0.191]
Epoch 37:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
14, train_loss_epoch=0.214]
Epoch 37: 100%|██████████| 1/1 [00:00<00:00,  1.28it/s, v_num=0, train_loss_
step=0.207, train_loss_epoch=0.214]
Epoch 37: 100%|██████████| 1/1 [00:00<00:00,  1.28it/s, v_num=0, train_loss_
step=0.207, train_loss_epoch=0.207]
Epoch 38:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
07, train_loss_epoch=0.207]
Epoch 39:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
33, train_loss_epoch=0.233]
Epoch 40:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
27, train_loss_epoch=0.227]
Epoch 41:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.1
99, train_loss_epoch=0.199]
Epoch 42:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
04, train_loss_epoch=0.204]
Epoch 43:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
07, train_loss_epoch=0.207]
Epoch 44:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
25, train_loss_epoch=0.225]
Epoch 45:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
20, train_loss_epoch=0.220]
Epoch 46:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
11, train_loss_epoch=0.211]
Epoch 47:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
18, train_loss_epoch=0.218]
Epoch 48:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.1
99, train_loss_epoch=0.199]
Epoch 48: 100%|██████████| 1/1 [00:00<00:00,  2.56it/s, v_num=0, train_loss_
step=0.199, train_loss_epoch=0.199]
Epoch 49:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.1
99, train_loss_epoch=0.199]
```

Loading [MathJax]/extensions/Safe.js

```
Epoch 50:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
17, train_loss_epoch=0.217]
Epoch 51:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
16, train_loss_epoch=0.216]
Epoch 52:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
06, train_loss_epoch=0.206]
Epoch 53:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.1
96, train_loss_epoch=0.196]
Epoch 53: 100%|██████████| 1/1 [00:00<00:00,  5.17it/s, v_num=0, train_loss_
step=0.234, train_loss_epoch=0.196]
Epoch 54:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
34, train_loss_epoch=0.234]
Epoch 55:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
04, train_loss_epoch=0.204]
Epoch 56:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
01, train_loss_epoch=0.201]
Epoch 57:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
19, train_loss_epoch=0.219]
Epoch 58:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
15, train_loss_epoch=0.215]
Epoch 59:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
08, train_loss_epoch=0.208]
Epoch 60:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
13, train_loss_epoch=0.213]
Epoch 61:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
06, train_loss_epoch=0.206]
Epoch 62:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.1
76, train_loss_epoch=0.176]
Epoch 63:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.1
90, train_loss_epoch=0.190]
Epoch 63:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.1
69, train_loss_epoch=0.169]
Epoch 64:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.1
69, train_loss_epoch=0.169]
Epoch 65:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
09, train_loss_epoch=0.209]
Epoch 66:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
09, train_loss_epoch=0.209]
Epoch 66: 100%|██████████| 1/1 [00:00<00:00,  3.30it/s, v_num=0, train_loss_
step=0.184, train_loss_epoch=0.184]
Epoch 67:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.1
84, train_loss_epoch=0.184]
Epoch 68:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
26, train_loss_epoch=0.226]
Epoch 69:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
11, train_loss_epoch=0.211]
Epoch 70:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
10, train_loss_epoch=0.210]
Epoch 71:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
01, train_loss_epoch=0.201]
Epoch 71: 100%|██████████| 1/1 [00:00<00:00,  2.77it/s, v_num=0, train_loss_
step=0.185, train_loss_epoch=0.201]
Epoch 72:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.1
85, train_loss_epoch=0.185]
Epoch 73:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
15, train_loss_epoch=0.215]
```

Loading [MathJax]/extensions/Safe.js

```
Epoch 73:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.1
89, train_loss_epoch=0.189]
Epoch 74:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.1
89, train_loss_epoch=0.189]
Epoch 75:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
06, train_loss_epoch=0.206]
Epoch 76:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
13, train_loss_epoch=0.213]
Epoch 77:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
18, train_loss_epoch=0.218]
Epoch 78:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.1
93, train_loss_epoch=0.193]
Epoch 79:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.1
94, train_loss_epoch=0.194]
Epoch 80:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
13, train_loss_epoch=0.213]
Epoch 81:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
04, train_loss_epoch=0.204]
Epoch 82:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
03, train_loss_epoch=0.203]
Epoch 83:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
02, train_loss_epoch=0.202]
Epoch 84:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.1
81, train_loss_epoch=0.181]
Epoch 85:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
09, train_loss_epoch=0.209]
Epoch 86:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.1
88, train_loss_epoch=0.188]
Epoch 87:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.1
94, train_loss_epoch=0.194]
Epoch 88:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
05, train_loss_epoch=0.205]
Epoch 89:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
03, train_loss_epoch=0.203]
Epoch 90:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.1
74, train_loss_epoch=0.174]
Epoch 91:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.1
95, train_loss_epoch=0.195]
Epoch 92:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
01, train_loss_epoch=0.201]
Epoch 93:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.1
94, train_loss_epoch=0.194]
Epoch 94:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
07, train_loss_epoch=0.207]
Epoch 95:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
07, train_loss_epoch=0.207]
Epoch 96:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
07, train_loss_epoch=0.207]
Epoch 96: 100%|██████████| 1/1 [00:00<00:00,  4.17it/s, v_num=0, train_loss_
step=0.195, train_loss_epoch=0.195]
Epoch 96:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.1
95, train_loss_epoch=0.195]
Epoch 97:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.1
95, train_loss_epoch=0.195]
Epoch 98:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
09, train_loss_epoch=0.209]
```

Loading [MathJax]/extensions/Safe.js

```
Epoch 99:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.203, train_loss_epoch=0.203]
Epoch 99: 100%|██████████| 1/1 [00:00<00:00,  1.96it/s, v_num=0, train_loss_step=0.199, train_loss_epoch=0.203]
Validation: |          | 0/? [00:00<?, ?it/s]
Validation:   0%|          | 0/1 [00:00<?, ?it/s]
Validation DataLoader 0:   0%|          | 0/1 [00:00<?, ?it/s]
(_train_tune pid=7588)
Validation DataLoader 0: 100%|██████████| 1/1 [00:01<00:00,  0.92it/s]
Epoch 99: 100%|██████████| 1/1 [00:01<00:00,  0.62it/s, v_num=0, train_loss_step=0.199, train_loss_epoch=0.203, valid_loss=0.339]
Epoch 99: 100%|██████████| 1/1 [00:01<00:00,  0.61it/s, v_num=0, train_loss_step=0.199, train_loss_epoch=0.199, valid_loss=0.339]
(_train_tune pid=7588) Seed set to 6
```

Loading [MathJax]/extensions/Safe.js

```
Sanity Checking DataLoader 0:   0%|          | 0/1 [00:00<?, ?it/s]
Epoch 0:   0%|          | 0/1 [00:00<?, ?it/s]
Epoch 1:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.20
2, train_loss_epoch=0.202]
Epoch 2:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.30
3, train_loss_epoch=0.303]
Epoch 3:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.23
7, train_loss_epoch=0.237]
Epoch 4:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.21
1, train_loss_epoch=0.211]
Epoch 5:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.20
7, train_loss_epoch=0.207]
Epoch 6:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.22
7, train_loss_epoch=0.227]
Epoch 7:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.25
2, train_loss_epoch=0.252]
Epoch 7: 100%|██████████| 1/1 [00:00<00:00,  3.01it/s, v_num=0, train_loss_s
tep=0.217, train_loss_epoch=0.217]
Epoch 8:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.21
7, train_loss_epoch=0.217]
Epoch 9:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.20
9, train_loss_epoch=0.209]
Epoch 10:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
29, train_loss_epoch=0.229]
Epoch 11:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
14, train_loss_epoch=0.214]
Epoch 12:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
12, train_loss_epoch=0.212]
Epoch 13:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.1
90, train_loss_epoch=0.190]
Epoch 14:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
42, train_loss_epoch=0.242]
Epoch 15:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
29, train_loss_epoch=0.229]
Epoch 16:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
00, train_loss_epoch=0.200]
Epoch 17:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.1
96, train_loss_epoch=0.196]
Epoch 18:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
11, train_loss_epoch=0.211]
Epoch 19:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.1
98, train_loss_epoch=0.198]
Epoch 20:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
00, train_loss_epoch=0.200]
Epoch 21:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
13, train_loss_epoch=0.213]
Epoch 22:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
22, train_loss_epoch=0.222]
Epoch 23:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
11, train_loss_epoch=0.211]
Epoch 24:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
13, train_loss_epoch=0.213]
Epoch 25:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
32, train_loss_epoch=0.232]
Epoch 26:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.1
93, train_loss_epoch=0.193]
```

Loading [MathJax]/extensions/Safe.js

```
Epoch 27:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
17, train_loss_epoch=0.217]
Epoch 28:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
24, train_loss_epoch=0.224]
Epoch 29:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
00, train_loss_epoch=0.200]
Epoch 30:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.1
98, train_loss_epoch=0.198]
Epoch 31:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
07, train_loss_epoch=0.207]
Epoch 32:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.1
97, train_loss_epoch=0.197]
Epoch 33:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
26, train_loss_epoch=0.226]
Epoch 34:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
24, train_loss_epoch=0.224]
Epoch 35:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.1
89, train_loss_epoch=0.189]
Epoch 36:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.1
96, train_loss_epoch=0.196]
Epoch 37:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.1
85, train_loss_epoch=0.185]
Epoch 38:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
02, train_loss_epoch=0.202]
Epoch 39:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
17, train_loss_epoch=0.217]
Epoch 40:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.1
98, train_loss_epoch=0.198]
Epoch 41:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.1
93, train_loss_epoch=0.193]
Epoch 42:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
25, train_loss_epoch=0.225]
Epoch 42: 100%|██████████| 1/1 [00:00<00:00,  3.00it/s, v_num=0, train_loss_
step=0.205, train_loss_epoch=0.225]
Epoch 43:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
05, train_loss_epoch=0.205]
Epoch 44:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
16, train_loss_epoch=0.216]
Epoch 45:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.1
81, train_loss_epoch=0.181]
Epoch 46:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.1
99, train_loss_epoch=0.199]
Epoch 47:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.1
87, train_loss_epoch=0.187]
Epoch 48:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.1
87, train_loss_epoch=0.187]
Epoch 48: 100%|██████████| 1/1 [00:00<00:00,  2.68it/s, v_num=0, train_loss_
step=0.218, train_loss_epoch=0.218]
Epoch 49:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
18, train_loss_epoch=0.218]
Epoch 50:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.1
98, train_loss_epoch=0.198]
Epoch 51:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.1
98, train_loss_epoch=0.198]
Epoch 52:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.1
92, train_loss_epoch=0.192]
```

Loading [MathJax]/extensions/Safe.js

```
Epoch 53:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
07, train_loss_epoch=0.207]
Epoch 54:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.1
95, train_loss_epoch=0.195]
Epoch 55:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
27, train_loss_epoch=0.227]
Epoch 56:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.1
96, train_loss_epoch=0.196]
Epoch 57:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.1
83, train_loss_epoch=0.183]
Epoch 58:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
07, train_loss_epoch=0.207]
Epoch 59:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.1
96, train_loss_epoch=0.196]
Epoch 60:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
01, train_loss_epoch=0.201]
Epoch 61:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
04, train_loss_epoch=0.204]
Epoch 62:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.1
99, train_loss_epoch=0.199]
Epoch 63:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
02, train_loss_epoch=0.202]
Epoch 64:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
15, train_loss_epoch=0.215]
Epoch 65:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.1
95, train_loss_epoch=0.195]
Epoch 66:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
08, train_loss_epoch=0.208]
Epoch 67:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
04, train_loss_epoch=0.204]
Epoch 68:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
24, train_loss_epoch=0.224]
Epoch 69:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.1
92, train_loss_epoch=0.192]
Epoch 70:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.1
86, train_loss_epoch=0.186]
Epoch 71:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
05, train_loss_epoch=0.205]
Epoch 72:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.1
99, train_loss_epoch=0.199]
Epoch 73:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.1
80, train_loss_epoch=0.180]
Epoch 74:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.1
99, train_loss_epoch=0.199]
Epoch 75:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
01, train_loss_epoch=0.201]
Epoch 76:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
24, train_loss_epoch=0.224]
Epoch 77:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
08, train_loss_epoch=0.208]
Epoch 78:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
11, train_loss_epoch=0.211]
Epoch 79:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
24, train_loss_epoch=0.224]
Epoch 80:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
10, train_loss_epoch=0.210]
```

Loading [MathJax]/extensions/Safe.js

```
Epoch 81:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.1
98, train_loss_epoch=0.198]
Epoch 82:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.1
95, train_loss_epoch=0.195]
Epoch 83:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.1
96, train_loss_epoch=0.196]
Epoch 84:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
02, train_loss_epoch=0.202]
Epoch 84: 100%|██████████| 1/1 [00:00<00:00,  3.72it/s, v_num=0, train_loss_
step=0.182, train_loss_epoch=0.182]
Epoch 85:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.1
82, train_loss_epoch=0.182]
Epoch 86:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.1
95, train_loss_epoch=0.195]
Epoch 87:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.1
94, train_loss_epoch=0.194]
Epoch 88:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.1
99, train_loss_epoch=0.199]
Epoch 89:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
05, train_loss_epoch=0.205]
Epoch 90:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
09, train_loss_epoch=0.209]
Epoch 91:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.1
88, train_loss_epoch=0.188]
Epoch 92:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.1
94, train_loss_epoch=0.194]
Epoch 93:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
10, train_loss_epoch=0.210]
Epoch 94:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.1
82, train_loss_epoch=0.182]
Epoch 95:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.1
71, train_loss_epoch=0.171]
Epoch 95: 100%|██████████| 1/1 [00:00<00:00,  5.70it/s, v_num=0, train_loss_
step=0.206, train_loss_epoch=0.206]
Epoch 96:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
06, train_loss_epoch=0.206]
Epoch 97:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.1
96, train_loss_epoch=0.196]
Epoch 98:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
04, train_loss_epoch=0.204]
Epoch 99:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
00, train_loss_epoch=0.200]
Epoch 99: 100%|██████████| 1/1 [00:00<00:00,  1.92it/s, v_num=0, train_loss_
step=0.181, train_loss_epoch=0.200]
Validation: |          | 0/? [00:00<?, ?it/s]
(_train_tune pid=7588)
Validation:   0%|          | 0/1 [00:00<?, ?it/s]
Validation DataLoader 0:   0%|          | 0/1 [00:00<?, ?it/s]
(_train_tune pid=7588)
Validation DataLoader 0: 100%|██████████| 1/1 [00:00<00:00,  1.25it/s]
Epoch 99: 100%|██████████| 1/1 [00:01<00:00,  0.73it/s, v_num=0, train_loss_
step=0.181, train_loss_epoch=0.181, valid_loss=0.317]
(_train_tune pid=7588) Seed set to 4
```

Loading [MathJax]/extensions/Safe.js

```
Sanity Checking DataLoader 0:   0%|          | 0/1 [00:00<?, ?it/s]
Epoch 0:   0%|          | 0/1 [00:00<?, ?it/s]
Epoch 1:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.22
9, train_loss_epoch=0.229]
Epoch 2:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.24
6, train_loss_epoch=0.246]
Epoch 3:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.22
6, train_loss_epoch=0.226]
Epoch 4:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.22
1, train_loss_epoch=0.221]
Epoch 4: 100%|██████████| 1/1 [00:00<00:00,  2.14it/s, v_num=0, train_loss_s
tep=0.199, train_loss_epoch=0.221]
Epoch 4: 100%|██████████| 1/1 [00:00<00:00,  2.07it/s, v_num=0, train_loss_s
tep=0.199, train_loss_epoch=0.199]
Epoch 5:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.19
9, train_loss_epoch=0.199]
Epoch 6:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.21
4, train_loss_epoch=0.214]
Epoch 7:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.21
7, train_loss_epoch=0.217]
Epoch 8:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.20
9, train_loss_epoch=0.209]
Epoch 9:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.20
2, train_loss_epoch=0.202]
Epoch 10:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
04, train_loss_epoch=0.204]
Epoch 11:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
04, train_loss_epoch=0.204]
Epoch 12:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
14, train_loss_epoch=0.214]
Epoch 13:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
06, train_loss_epoch=0.206]
Epoch 14:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
02, train_loss_epoch=0.202]
Epoch 15:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
09, train_loss_epoch=0.209]
Epoch 16:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
20, train_loss_epoch=0.220]
Epoch 17:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.1
95, train_loss_epoch=0.195]
Epoch 18:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
01, train_loss_epoch=0.201]
Epoch 19:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
14, train_loss_epoch=0.214]
Epoch 19: 100%|██████████| 1/1 [00:00<00:00,  3.94it/s, v_num=0, train_loss_
step=0.204, train_loss_epoch=0.214]
Epoch 19: 100%|██████████| 1/1 [00:00<00:00,  3.92it/s, v_num=0, train_loss_
step=0.204, train_loss_epoch=0.204]
Epoch 20:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
04, train_loss_epoch=0.204]
Epoch 21:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
10, train_loss_epoch=0.210]
Epoch 22:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
15, train_loss_epoch=0.215]
Epoch 23:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
06, train_loss_epoch=0.206]
```

Loading [MathJax]/extensions/Safe.js

```
Epoch 24:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
11, train_loss_epoch=0.211]
Epoch 25:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
28, train_loss_epoch=0.228]
Epoch 26:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.1
96, train_loss_epoch=0.196]
Epoch 27:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
32, train_loss_epoch=0.232]
Epoch 28:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.1
99, train_loss_epoch=0.199]
Epoch 29:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
14, train_loss_epoch=0.214]
Epoch 30:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
09, train_loss_epoch=0.209]
Epoch 31:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
26, train_loss_epoch=0.226]
Epoch 32:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.1
93, train_loss_epoch=0.193]
Epoch 33:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
14, train_loss_epoch=0.214]
Epoch 33: 100%|██████████| 1/1 [00:00<00:00,  1.26it/s, v_num=0, train_loss_
step=0.214, train_loss_epoch=0.214]
Epoch 34:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
07, train_loss_epoch=0.207]
Epoch 35:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
14, train_loss_epoch=0.214]
Epoch 36:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
19, train_loss_epoch=0.219]
Epoch 37:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.1
89, train_loss_epoch=0.189]
Epoch 37: 100%|██████████| 1/1 [00:00<00:00,  2.84it/s, v_num=0, train_loss_
step=0.205, train_loss_epoch=0.189]
Epoch 38:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
05, train_loss_epoch=0.205]
Epoch 39:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.1
98, train_loss_epoch=0.198]
Epoch 40:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
09, train_loss_epoch=0.209]
Epoch 41:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
06, train_loss_epoch=0.206]
Epoch 42:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
13, train_loss_epoch=0.213]
Epoch 42: 100%|██████████| 1/1 [00:00<00:00,  2.21it/s, v_num=0, train_loss_
step=0.213, train_loss_epoch=0.213]
Epoch 43:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
13, train_loss_epoch=0.213]
Epoch 44:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
06, train_loss_epoch=0.206]
Epoch 45:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.1
94, train_loss_epoch=0.194]
Epoch 45: 100%|██████████| 1/1 [00:00<00:00,  2.98it/s, v_num=0, train_loss_
step=0.215, train_loss_epoch=0.215]
Epoch 46:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
15, train_loss_epoch=0.215]
Epoch 47:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.1
91, train_loss_epoch=0.191]
```

Loading [MathJax]/extensions/Safe.js

```
Epoch 48:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.1
92, train_loss_epoch=0.192]
Epoch 49:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.1
89, train_loss_epoch=0.189]
Epoch 50:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
12, train_loss_epoch=0.212]
Epoch 51:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
06, train_loss_epoch=0.206]
Epoch 52:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
04, train_loss_epoch=0.204]
Epoch 53:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
16, train_loss_epoch=0.216]
Epoch 53: 100%|██████████| 1/1 [00:00<00:00,  2.60it/s, v_num=0, train_loss_
step=0.196, train_loss_epoch=0.196]
Epoch 54:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.1
96, train_loss_epoch=0.196]
Epoch 55:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
00, train_loss_epoch=0.200]
Epoch 56:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.1
98, train_loss_epoch=0.198]
Epoch 57:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
07, train_loss_epoch=0.207]
Epoch 58:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
02, train_loss_epoch=0.202]
Epoch 59:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
27, train_loss_epoch=0.227]
Epoch 60:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
11, train_loss_epoch=0.211]
Epoch 61:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.1
96, train_loss_epoch=0.196]
Epoch 62:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
09, train_loss_epoch=0.209]
Epoch 63:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
11, train_loss_epoch=0.211]
Epoch 64:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.1
94, train_loss_epoch=0.194]
Epoch 65:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
01, train_loss_epoch=0.201]
Epoch 66:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
00, train_loss_epoch=0.200]
Epoch 67:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
12, train_loss_epoch=0.212]
Epoch 68:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
08, train_loss_epoch=0.208]
Epoch 69:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.1
98, train_loss_epoch=0.198]
Epoch 70:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
09, train_loss_epoch=0.209]
Epoch 71:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.1
94, train_loss_epoch=0.194]
Epoch 72:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.1
97, train_loss_epoch=0.197]
Epoch 72: 100%|██████████| 1/1 [00:00<00:00,  2.29it/s, v_num=0, train_loss_
step=0.194, train_loss_epoch=0.197]
Epoch 73:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.1
94, train_loss_epoch=0.194]
```

Loading [MathJax]/extensions/Safe.js

```
Epoch 74:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.1
99, train_loss_epoch=0.199]
Epoch 75:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.1
84, train_loss_epoch=0.184]
Epoch 76:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.1
81, train_loss_epoch=0.181]
Epoch 77:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.1
75, train_loss_epoch=0.175]
Epoch 78:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.1
95, train_loss_epoch=0.195]
Epoch 79:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
30, train_loss_epoch=0.230]
Epoch 80:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
08, train_loss_epoch=0.208]
Epoch 81:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.1
88, train_loss_epoch=0.188]
Epoch 82:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.1
76, train_loss_epoch=0.176]
Epoch 83:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.1
96, train_loss_epoch=0.196]
Epoch 84:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.1
83, train_loss_epoch=0.183]
Epoch 85:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.1
99, train_loss_epoch=0.199]
Epoch 86:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.1
79, train_loss_epoch=0.179]
Epoch 87:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.1
88, train_loss_epoch=0.188]
Epoch 88:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
00, train_loss_epoch=0.200]
Epoch 89:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.1
97, train_loss_epoch=0.197]
Epoch 90:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
11, train_loss_epoch=0.211]
Epoch 91:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.1
88, train_loss_epoch=0.188]
Epoch 92:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.1
89, train_loss_epoch=0.189]
Epoch 93:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
25, train_loss_epoch=0.225]
Epoch 94:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.1
79, train_loss_epoch=0.179]
Epoch 95:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.1
92, train_loss_epoch=0.192]
Epoch 96:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
00, train_loss_epoch=0.200]
Epoch 97:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
18, train_loss_epoch=0.218]
Epoch 98:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.1
82, train_loss_epoch=0.182]
Epoch 99:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
15, train_loss_epoch=0.215]
Epoch 99: 100%|██████████| 1/1 [00:00<00:00,  3.86it/s, v_num=0, train_loss_
step=0.212, train_loss_epoch=0.215]
Validation: |          | 0/? [00:00<?, ?it/s]
Validation:   0%|          | 0/1 [00:00<?, ?it/s]
```

Loading [MathJax]/extensions/Safe.js

```
Validation DataLoader 0:   0%|          | 0/1 [00:00<?, ?it/s]
(_train_tune pid=7588)
Validation DataLoader 0: 100%|██████████| 1/1 [00:00<00:00,  1.22it/s]
Epoch 99: 100%|██████████| 1/1 [00:01<00:00,  0.89it/s, v_num=0, train_loss_
step=0.212, train_loss_epoch=0.212, valid_loss=0.331]
(_train_tune pid=7588) Seed set to 3
```

Loading [MathJax]/extensions/Safe.js

```
Sanity Checking DataLoader 0:   0%|          | 0/1 [00:00<?, ?it/s]
Epoch 0:   0%|          | 0/1 [00:00<?, ?it/s]
Epoch 1:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.20
8, train_loss_epoch=0.208]
Epoch 2:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.50
7, train_loss_epoch=0.507]
Epoch 3:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.23
5, train_loss_epoch=0.235]
Epoch 4:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.25
0, train_loss_epoch=0.250]
Epoch 5:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.23
0, train_loss_epoch=0.230]
Epoch 6:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.21
3, train_loss_epoch=0.213]
Epoch 7:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.25
2, train_loss_epoch=0.252]
Epoch 8:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.20
7, train_loss_epoch=0.207]
Epoch 9:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.22
7, train_loss_epoch=0.227]
Epoch 10:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
20, train_loss_epoch=0.220]
Epoch 11:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
02, train_loss_epoch=0.202]
Epoch 12:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
17, train_loss_epoch=0.217]
Epoch 13:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
27, train_loss_epoch=0.227]
Epoch 14:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
15, train_loss_epoch=0.215]
Epoch 15:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
27, train_loss_epoch=0.227]
Epoch 16:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
20, train_loss_epoch=0.220]
Epoch 17:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
00, train_loss_epoch=0.200]
Epoch 18:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
13, train_loss_epoch=0.213]
Epoch 19:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
00, train_loss_epoch=0.200]
Epoch 20:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
32, train_loss_epoch=0.232]
Epoch 21:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
09, train_loss_epoch=0.209]
Epoch 22:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
19, train_loss_epoch=0.219]
Epoch 23:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
01, train_loss_epoch=0.201]
Epoch 24:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
31, train_loss_epoch=0.231]
Epoch 25:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
06, train_loss_epoch=0.206]
Epoch 25: 100%|██████████| 1/1 [00:00<00:00,  1.69it/s, v_num=0, train_loss_
step=0.214, train_loss_epoch=0.214]
Epoch 26:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
14, train_loss_epoch=0.214]
```

Loading [MathJax]/extensions/Safe.js

```
Epoch 27:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.1
97, train_loss_epoch=0.197]
Epoch 28:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.1
75, train_loss_epoch=0.175]
Epoch 29:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
28, train_loss_epoch=0.228]
Epoch 30:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.1
86, train_loss_epoch=0.186]
Epoch 31:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
06, train_loss_epoch=0.206]
Epoch 32:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
12, train_loss_epoch=0.212]
Epoch 33:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.1
94, train_loss_epoch=0.194]
Epoch 34:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
14, train_loss_epoch=0.214]
Epoch 35:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.1
79, train_loss_epoch=0.179]
Epoch 36:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.1
88, train_loss_epoch=0.188]
Epoch 37:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
09, train_loss_epoch=0.209]
Epoch 38:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
07, train_loss_epoch=0.207]
Epoch 39:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
29, train_loss_epoch=0.229]
Epoch 40:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
21, train_loss_epoch=0.221]
Epoch 41:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.1
94, train_loss_epoch=0.194]
Epoch 42:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
00, train_loss_epoch=0.200]
Epoch 43:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
02, train_loss_epoch=0.202]
Epoch 44:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
16, train_loss_epoch=0.216]
Epoch 45:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
14, train_loss_epoch=0.214]
Epoch 46:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
03, train_loss_epoch=0.203]
Epoch 47:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
16, train_loss_epoch=0.216]
Epoch 48:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.1
91, train_loss_epoch=0.191]
Epoch 49:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.1
91, train_loss_epoch=0.191]
Epoch 50:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
11, train_loss_epoch=0.211]
Epoch 51:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
12, train_loss_epoch=0.212]
Epoch 52:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.1
99, train_loss_epoch=0.199]
Epoch 53:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.1
86, train_loss_epoch=0.186]
Epoch 54:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
26, train_loss_epoch=0.226]
```

Loading [MathJax]/extensions/Safe.js

```
Epoch 55:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.1
96, train_loss_epoch=0.196]
Epoch 56:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.1
96, train_loss_epoch=0.196]
Epoch 57:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
12, train_loss_epoch=0.212]
Epoch 58:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
10, train_loss_epoch=0.210]
Epoch 59:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
05, train_loss_epoch=0.205]
Epoch 60:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
04, train_loss_epoch=0.204]
Epoch 60: 100%|██████████| 1/1 [00:00<00:00,  2.54it/s, v_num=0, train_loss_
step=0.200, train_loss_epoch=0.204]
Epoch 60: 100%|██████████| 1/1 [00:00<00:00,  2.52it/s, v_num=0, train_loss_
step=0.200, train_loss_epoch=0.200]
Epoch 61:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
00, train_loss_epoch=0.200]
Epoch 62:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.1
71, train_loss_epoch=0.171]
Epoch 63:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.1
86, train_loss_epoch=0.186]
Epoch 64:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.1
63, train_loss_epoch=0.163]
Epoch 65:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
07, train_loss_epoch=0.207]
Epoch 66:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
03, train_loss_epoch=0.203]
Epoch 67:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.1
79, train_loss_epoch=0.179]
Epoch 68:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
19, train_loss_epoch=0.219]
Epoch 69:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
05, train_loss_epoch=0.205]
Epoch 70:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
06, train_loss_epoch=0.206]
Epoch 71:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.1
96, train_loss_epoch=0.196]
Epoch 72:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.1
81, train_loss_epoch=0.181]
Epoch 73:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
10, train_loss_epoch=0.210]
Epoch 74:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.1
86, train_loss_epoch=0.186]
Epoch 74: 100%|██████████| 1/1 [00:00<00:00,  1.56it/s, v_num=0, train_loss_
step=0.197, train_loss_epoch=0.186]
Epoch 75:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.1
97, train_loss_epoch=0.197]
Epoch 76:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
04, train_loss_epoch=0.204]
Epoch 77:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
12, train_loss_epoch=0.212]
Epoch 78:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.1
87, train_loss_epoch=0.187]
Epoch 79:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.1
90, train_loss_epoch=0.190]
```

Loading [MathJax]/extensions/Safe.js

```
Epoch 80:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
07, train_loss_epoch=0.207]
Epoch 81:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.1
97, train_loss_epoch=0.197]
Epoch 82:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.1
98, train_loss_epoch=0.198]
Epoch 83:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.1
98, train_loss_epoch=0.198]
Epoch 84:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.1
78, train_loss_epoch=0.178]
Epoch 85:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
03, train_loss_epoch=0.203]
Epoch 86:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.1
83, train_loss_epoch=0.183]
Epoch 87:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.1
89, train_loss_epoch=0.189]
Epoch 88:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.1
98, train_loss_epoch=0.198]
Epoch 89:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.1
97, train_loss_epoch=0.197]
Epoch 90:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.1
69, train_loss_epoch=0.169]
Epoch 91:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.1
88, train_loss_epoch=0.188]
Epoch 92:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.1
96, train_loss_epoch=0.196]
Epoch 93:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.1
89, train_loss_epoch=0.189]
Epoch 94:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
04, train_loss_epoch=0.204]
Epoch 95:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
03, train_loss_epoch=0.203]
Epoch 96:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
04, train_loss_epoch=0.204]
Epoch 96:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.1
91, train_loss_epoch=0.191]
Epoch 97:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.1
91, train_loss_epoch=0.191]
Epoch 98:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
04, train_loss_epoch=0.204]
Epoch 99:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.1
95, train_loss_epoch=0.195]
Epoch 99: 100%|██████████| 1/1 [00:00<00:00,  3.83it/s, v_num=0, train_loss_
step=0.194, train_loss_epoch=0.195]
Validation: |          | 0/? [00:00<?, ?it/s]
Validation:   0%|          | 0/1 [00:00<?, ?it/s]
Validation DataLoader 0:   0%|          | 0/1 [00:00<?, ?it/s]
Seed set to 6
(_train_tune pid=7588)
Validation DataLoader 0: 100%|██████████| 1/1 [00:00<00:00,  1.80it/s]
Epoch 99: 100%|██████████| 1/1 [00:00<00:00,  1.09it/s, v_num=0, train_loss_
step=0.194, train_loss_epoch=0.194, valid_loss=0.340]
```

Loading [MathJax]/extensions/Safe.js

```
2023-11-02 03:22:11.024096: I tensorflow/core/util/port.cc:111] oneDNN custo
m operations are on. You may see slightly different numerical results due to
floating-point round-off errors from different computation orders. To turn t
hem off, set the environment variable `TF_ENABLE_ONEDNN_OPTS=0`.
2023-11-02 03:22:11.040501: I tensorflow/tsl/cuda/cudart_stub.cc:28] Could n
ot find cuda drivers on your machine, GPU will not be used.
2023-11-02 03:22:11.252872: E tensorflow/compiler/xla/stream_executor/cuda/c
uda_dnn.cc:9342] Unable to register cuDNN factory: Attempting to register fa
ctory for plugin cuDNN when one has already been registered
2023-11-02 03:22:11.252943: E tensorflow/compiler/xla/stream_executor/cuda/c
uda_fft.cc:609] Unable to register cuFFT factory: Attempting to register fac
tory for plugin cuFFT when one has already been registered
2023-11-02 03:22:11.252993: E tensorflow/compiler/xla/stream_executor/cuda/c
uda_blas.cc:1518] Unable to register cuBLAS factory: Attempting to register
factory for plugin cuBLAS when one has already been registered
2023-11-02 03:22:11.283614: I tensorflow/core/platform/cpu_feature_guard.cc:
182] This TensorFlow binary is optimized to use available CPU instructions i
n performance-critical operations.
To enable the following instructions: AVX2 AVX_VNNI FMA, in other operation
s, rebuild TensorFlow with the appropriate compiler flags.
2023-11-02 03:22:15.280598: W tensorflow/compiler/tf2tensorrt/utils/py_util
s.cc:38] TF-TRT Warning: Could not find TensorRT
Sanity Checking: |
| 0/? [00:00…
Training: |
| 0/? [00:00…
Validation: |
| 0/? [00:00…
Predicting: |
| 0/? [00:00…
```

In [8]:
```
Y_hat_df.head()
```

Out[8]:

	unique_id	ds	cutoff	AutoNHITS	y
0	0	2006-08-16	2006-08-15	2.945293	2.948076
1	0	2006-08-17	2006-08-15	2.936055	3.049320
2	0	2006-08-18	2006-08-15	2.945046	3.064168
3	0	2006-08-19	2006-08-15	2.964904	3.005783
4	0	2006-08-20	2006-08-15	2.908318	3.010031

In [7]:
```
Y_hat_df.to_csv('results/Exchange_rate/AutoNHITS.csv')
```

In [6]:
```
from neuralforecast.losses.numpy import mae, mse

print('MAE: ', mae(Y_hat_df['y'], Y_hat_df['AutoNHITS']))
print('MSE: ', mse(Y_hat_df['y'], Y_hat_df['AutoNHITS']))
```

```
MAE:  0.2089021863763513
MSE:  0.08673494391505497
```

Loading [MathJax]/extensions/Safe.js

Ettm2

In [26]:
```
import pandas as pd
from neuralforecast.core import NeuralForecast

Y_df = pd.read_csv("raw_data/df_Ettm2.csv")

Y_df['ds'] = pd.to_datetime(Y_df['ds'])

# For this excercise we are going to take 20% of the DataSet
n_time = len(Y_df.ds.unique())
val_size = int(.2 * n_time)
test_size = int(.2 * n_time)

Y_df.groupby('unique_id').head(2)
```

Out[26]:

	unique_id	ds	y
0	HUFL	2016-07-01 00:00:00	-0.041413
1	HUFL	2016-07-01 00:15:00	-0.185467
57600	HULL	2016-07-01 00:00:00	0.040104
57601	HULL	2016-07-01 00:15:00	-0.214450
115200	LUFL	2016-07-01 00:00:00	0.695804
115201	LUFL	2016-07-01 00:15:00	0.434685
172800	LULL	2016-07-01 00:00:00	0.434430
172801	LULL	2016-07-01 00:15:00	0.428168
230400	MUFL	2016-07-01 00:00:00	-0.599211
230401	MUFL	2016-07-01 00:15:00	-0.658068
288000	MULL	2016-07-01 00:00:00	-0.393536
288001	MULL	2016-07-01 00:15:00	-0.659338
345600	OT	2016-07-01 00:00:00	1.018032
345601	OT	2016-07-01 00:15:00	0.980124

In [30]:
```
horizon = 96 # 24hrs = 4 * 15 min.

# Use your own config or AutoNHITS.default_config
nhits_config = {
       "learning_rate": tune.choice([1e-3]),
       "max_steps": tune.choice([1000]),
       "input_size": tune.choice([5 * horizon]),
       "batch_size": tune.choice([7]),
       "windows_batch_size": tune.choice([256]),
       "n_pool_kernel_size": tune.choice([[2, 2, 2], [16, 8, 1]]),
       "n_freq_downsample": tune.choice([[168, 24, 1], [24, 12, 1], [1, 1, 1
       "activation": tune.choice(['ReLU']),
```

Loading [MathJax]/extensions/Safe.js

```
        "n_blocks":  tune.choice([[1, 1, 1]]),
        "mlp_units":  tune.choice([[[512, 512], [512, 512], [512, 512]]]),
        "interpolation_mode": tune.choice(['linear']),
        "val_check_steps": tune.choice([100]),
        "random_seed": tune.randint(1, 10),
    }

models = [AutoNHITS(h=horizon,
                    config=nhits_config,
                    num_samples=5)]

%%capture
nf = NeuralForecast(
    models=models,
    freq='15min')

Y_hat_df = nf.cross_validation(df=Y_df, val_size=val_size,
                               test_size=test_size, n_windows=None)
```

```
Seed set to 1
Sanity Checking: |
| 0/? [00:00…
Training: |
| 0/? [00:00…
Predicting: |
| 0/? [00:00…
```

In [1]:
```
from neuralforecast.losses.numpy import mae, mse

print('MAE: ', 0.26096806135482414)
print('MSE: ', 0.18279484416711375)
```

```
MAE:  0.26096806135482414
MSE:  0.18279484416711375
```

In [ ]:
```
Y_hat_df.to_csv('results/Ettm2/NHITS.csv')
```

Weather

In [9]:
```
import pandas as pd
from neuralforecast.core import NeuralForecast

Y_df = pd.read_csv("raw_data/df_Weather.csv")

Y_df['ds'] = pd.to_datetime(Y_df['ds'])

# For this excercise we are going to take 20% of the DataSet
n_time = len(Y_df.ds.unique())
val_size = int(.1 * n_time)
test_size = int(.2 * n_time)

Y_df.groupby('unique_id').head(2)
```

Loading [MathJax]/extensions/Safe.js

Out[9]:

	unique_id	ds	y
0	H2OC (mmol/mol)	2020-01-01 00:10:00	-0.999107
1	H2OC (mmol/mol)	2020-01-01 00:20:00	-1.008072
52695	OT	2020-01-01 00:10:00	0.044395
52696	OT	2020-01-01 00:20:00	0.044134
105390	PAR (µmol/m²/s)	2020-01-01 00:10:00	-0.679493
105391	PAR (µmol/m²/s)	2020-01-01 00:20:00	-0.679493
158085	SWDR (W/m²)	2020-01-01 00:10:00	-0.672767
158086	SWDR (W/m²)	2020-01-01 00:20:00	-0.672767
210780	T (degC)	2020-01-01 00:10:00	-1.459980
210781	T (degC)	2020-01-01 00:20:00	-1.454798
263475	Tdew (degC)	2020-01-01 00:10:00	-1.052596
263476	Tdew (degC)	2020-01-01 00:20:00	-1.069612
316170	Tlog (degC)	2020-01-01 00:10:00	-1.424132
316171	Tlog (degC)	2020-01-01 00:20:00	-1.416612
368865	Tpot (K)	2020-01-01 00:10:00	-1.607935
368866	Tpot (K)	2020-01-01 00:20:00	-1.602882
421560	VPact (mbar)	2020-01-01 00:10:00	-0.979132
421561	VPact (mbar)	2020-01-01 00:20:00	-0.990506
474255	VPdef (mbar)	2020-01-01 00:10:00	-0.838497
474256	VPdef (mbar)	2020-01-01 00:20:00	-0.828332
526950	VPmax (mbar)	2020-01-01 00:10:00	-1.141181
526951	VPmax (mbar)	2020-01-01 00:20:00	-1.138714
579645	max. PAR (µmol/m²/s)	2020-01-01 00:10:00	-0.588296
579646	max. PAR (µmol/m²/s)	2020-01-01 00:20:00	-0.588296
632340	max. wv (m/s)	2020-01-01 00:10:00	-0.832381
632341	max. wv (m/s)	2020-01-01 00:20:00	-1.125140
685035	p (mbar)	2020-01-01 00:10:00	2.114257
685036	p (mbar)	2020-01-01 00:20:00	2.099194
737730	rain (mm)	2020-01-01 00:10:00	-0.093506
737731	rain (mm)	2020-01-01 00:20:00	-0.093506
790425	raining (s)	2020-01-01 00:10:00	-0.221050
790426	raining (s)	2020-01-01 00:20:00	-0.221050
843120	rh (%)	2020-01-01 00:10:00	0.990128

Loading [MathJax]/extensions/Safe.js

	unique_id	ds	y
843121	rh (%)	2020-01-01 00:20:00	0.942141
895815	rho (g/m**3)	2020-01-01 00:10:00	1.940406
895816	rho (g/m**3)	2020-01-01 00:20:00	1.932788
948510	sh (g/kg)	2020-01-01 00:10:00	-0.998513
948511	sh (g/kg)	2020-01-01 00:20:00	-1.009228
1001205	wd (deg)	2020-01-01 00:10:00	0.555571
1001206	wd (deg)	2020-01-01 00:20:00	0.354339
1053900	wv (m/s)	2020-01-01 00:10:00	-0.017801
1053901	wv (m/s)	2020-01-01 00:20:00	-0.029125

In [3]:
```
horizon = 96 # 24hrs = 4 * 15 min.

# Use your own config or AutoNHITS.default_config
nhits_config = {
       "learning_rate": tune.choice([1e-3]),
       "max_steps": tune.choice([100]),
       "input_size": tune.choice([horizon]),
       "batch_size": tune.choice([21]),
       "windows_batch_size": tune.choice([256]),
       "n_pool_kernel_size": tune.choice([[2, 2, 2], [16, 8, 1]]),
       "n_freq_downsample": tune.choice([[168, 24, 1], [24, 12, 1], [1, 1, 1
       "activation": tune.choice(['ReLU']),
       "n_blocks":  tune.choice([[1, 1, 1]]),
       "mlp_units":  tune.choice([[[512, 512], [512, 512], [512, 512]]]),
       "interpolation_mode": tune.choice(['linear']),
       "val_check_steps": tune.choice([10]),
       "random_seed": tune.randint(1, 10),
    }

models = [AutoNHITS(h=horizon,
                    config=nhits_config,
                    num_samples=5)]

nf = NeuralForecast(
    models=models,
    freq='10min')

Y_hat_df = nf.cross_validation(df=Y_df, val_size=val_size,
                               test_size=test_size, n_windows=None)
```

Loading [MathJax]/extensions/Safe.js

```
(_train_tune pid=14309) Seed set to 7
(_train_tune pid=14309) 2023-11-02 17:50:05.859747: I tensorflow/core/util/port.cc:111] oneDNN custom operations are on. You may see slightly different numerical results due to floating-point round-off errors from different computation orders. To turn them off, set the environment variable `TF_ENABLE_ONEDNN_OPTS=0`.
(_train_tune pid=14309) 2023-11-02 17:50:05.898412: I tensorflow/tsl/cuda/cudart_stub.cc:28] Could not find cuda drivers on your machine, GPU will not be used.
(_train_tune pid=14309) 2023-11-02 17:50:05.952885: E tensorflow/compiler/xla/stream_executor/cuda/cuda_dnn.cc:9342] Unable to register cuDNN factory: Attempting to register factory for plugin cuDNN when one has already been registered
(_train_tune pid=14309) 2023-11-02 17:50:05.952943: E tensorflow/compiler/xla/stream_executor/cuda/cuda_fft.cc:609] Unable to register cuFFT factory: Attempting to register factory for plugin cuFFT when one has already been registered
(_train_tune pid=14309) 2023-11-02 17:50:05.952970: E tensorflow/compiler/xla/stream_executor/cuda/cuda_blas.cc:1518] Unable to register cuBLAS factory: Attempting to register factory for plugin cuBLAS when one has already been registered
(_train_tune pid=14309) 2023-11-02 17:50:05.961992: I tensorflow/core/platform/cpu_feature_guard.cc:182] This TensorFlow binary is optimized to use available CPU instructions in performance-critical operations.
(_train_tune pid=14309) To enable the following instructions: AVX2 AVX_VNNI FMA, in other operations, rebuild TensorFlow with the appropriate compiler flags.
(_train_tune pid=14309) 2023-11-02 17:50:08.061134: W tensorflow/compiler/tf2tensorrt/utils/py_utils.cc:38] TF-TRT Warning: Could not find TensorRT
```

Loading [MathJax]/extensions/Safe.js

```
Sanity Checking: |          | 0/? [00:00<?, ?it/s]
Sanity Checking DataLoader 0:   0%|          | 0/1 [00:00<?, ?it/s]
Sanity Checking DataLoader 0: 100%|██████████| 1/1 [00:05<00:00,  0.18it/s]
Epoch 0:   0%|          | 0/1 [00:00<?, ?it/s]
Epoch 1:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.57
6, train_loss_epoch=0.576]
Epoch 2:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.48
8, train_loss_epoch=0.488]
Epoch 3:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.62
8, train_loss_epoch=0.628]
Epoch 4:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.50
1, train_loss_epoch=0.501]
Epoch 5:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.43
6, train_loss_epoch=0.436]
Epoch 6:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.42
7, train_loss_epoch=0.427]
Epoch 7:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.47
6, train_loss_epoch=0.476]
Epoch 8:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.35
4, train_loss_epoch=0.354]
Epoch 9:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.40
9, train_loss_epoch=0.409]
Epoch 9: 100%|██████████| 1/1 [00:01<00:00,  0.93it/s, v_num=0, train_loss_s
tep=0.351, train_loss_epoch=0.409]
Validation: |          | 0/? [00:00<?, ?it/s]
Validation:   0%|          | 0/1 [00:00<?, ?it/s]
Validation DataLoader 0:   0%|          | 0/1 [00:00<?, ?it/s]
(_train_tune pid=14309)
Validation DataLoader 0: 100%|██████████| 1/1 [00:07<00:00,  0.14it/s]
Epoch 9: 100%|██████████| 1/1 [00:08<00:00,  0.12it/s, v_num=0, train_loss_s
tep=0.351, train_loss_epoch=0.409, valid_loss=0.359]
Epoch 10:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.3
51, train_loss_epoch=0.351, valid_loss=0.359]
Epoch 11:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.3
81, train_loss_epoch=0.381, valid_loss=0.359]
Epoch 12:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.3
75, train_loss_epoch=0.375, valid_loss=0.359]
Epoch 13:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.3
34, train_loss_epoch=0.334, valid_loss=0.359]
Epoch 14:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.3
99, train_loss_epoch=0.399, valid_loss=0.359]
Epoch 15:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.3
58, train_loss_epoch=0.358, valid_loss=0.359]
Epoch 16:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.3
47, train_loss_epoch=0.347, valid_loss=0.359]
Epoch 17:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.3
76, train_loss_epoch=0.376, valid_loss=0.359]
Epoch 18:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.4
10, train_loss_epoch=0.410, valid_loss=0.359]
Epoch 19:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.3
48, train_loss_epoch=0.348, valid_loss=0.359]
Epoch 19: 100%|██████████| 1/1 [00:00<00:00,  1.46it/s, v_num=0, train_loss_
step=0.366, train_loss_epoch=0.348, valid_loss=0.359]
Validation: |          | 0/? [00:00<?, ?it/s]
Validation:   0%|          | 0/1 [00:00<?, ?it/s]
Validation DataLoader 0:   0%|          | 0/1 [00:00<?, ?it/s]
```

Loading [MathJax]/extensions/Safe.js

```
(_train_tune pid=14309)
Validation DataLoader 0: 100%|██████████| 1/1 [00:06<00:00,  0.15it/s]
Epoch 20:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.3
66, train_loss_epoch=0.366, valid_loss=0.325]
Epoch 21:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.3
42, train_loss_epoch=0.342, valid_loss=0.325]
Epoch 22:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.3
39, train_loss_epoch=0.339, valid_loss=0.325]
Epoch 23:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.3
90, train_loss_epoch=0.390, valid_loss=0.325]
Epoch 24:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.3
33, train_loss_epoch=0.333, valid_loss=0.325]
Epoch 25:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.3
34, train_loss_epoch=0.334, valid_loss=0.325]
Epoch 26:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.3
56, train_loss_epoch=0.356, valid_loss=0.325]
Epoch 27:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.3
65, train_loss_epoch=0.365, valid_loss=0.325]
Epoch 28:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.3
29, train_loss_epoch=0.329, valid_loss=0.325]
Epoch 29:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.3
56, train_loss_epoch=0.356, valid_loss=0.325]
Epoch 29: 100%|██████████| 1/1 [00:00<00:00,  1.36it/s, v_num=0, train_loss_
step=0.366, train_loss_epoch=0.356, valid_loss=0.325]
Validation: |          | 0/? [00:00<?, ?it/s]
Validation:   0%|          | 0/1 [00:00<?, ?it/s]
Validation DataLoader 0:   0%|          | 0/1 [00:00<?, ?it/s]
(_train_tune pid=14309)
Validation DataLoader 0: 100%|██████████| 1/1 [00:06<00:00,  0.16it/s]
Epoch 30:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.3
66, train_loss_epoch=0.366, valid_loss=0.314]
Epoch 31:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.3
43, train_loss_epoch=0.343, valid_loss=0.314]
Epoch 32:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.3
19, train_loss_epoch=0.319, valid_loss=0.314]
Epoch 33:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.3
49, train_loss_epoch=0.349, valid_loss=0.314]
Epoch 34:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.2
99, train_loss_epoch=0.299, valid_loss=0.314]
Epoch 35:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.3
29, train_loss_epoch=0.329, valid_loss=0.314]
Epoch 36:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.3
15, train_loss_epoch=0.315, valid_loss=0.314]
Epoch 37:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.3
18, train_loss_epoch=0.318, valid_loss=0.314]
Epoch 38:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.3
21, train_loss_epoch=0.321, valid_loss=0.314]
Epoch 39:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.3
18, train_loss_epoch=0.318, valid_loss=0.314]
Epoch 39: 100%|██████████| 1/1 [00:00<00:00,  1.21it/s, v_num=0, train_loss_
step=0.346, train_loss_epoch=0.318, valid_loss=0.314]
Validation: |          | 0/? [00:00<?, ?it/s]
Validation:   0%|          | 0/1 [00:00<?, ?it/s]
Validation DataLoader 0:   0%|          | 0/1 [00:00<?, ?it/s]
(_train_tune pid=14309)
Validation DataLoader 0: 100%|██████████| 1/1 [00:06<00:00,  0.16it/s]
```

Loading [MathJax]/extensions/Safe.js

```
Epoch 40:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.3
46, train_loss_epoch=0.346, valid_loss=0.305]
Epoch 41:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.3
26, train_loss_epoch=0.326, valid_loss=0.305]
Epoch 42:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.3
32, train_loss_epoch=0.332, valid_loss=0.305]
Epoch 43:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.3
29, train_loss_epoch=0.329, valid_loss=0.305]
Epoch 44:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.3
52, train_loss_epoch=0.352, valid_loss=0.305]
Epoch 45:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.3
16, train_loss_epoch=0.316, valid_loss=0.305]
Epoch 46:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.3
16, train_loss_epoch=0.316, valid_loss=0.305]
Epoch 47:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.3
28, train_loss_epoch=0.328, valid_loss=0.305]
Epoch 48:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.3
10, train_loss_epoch=0.310, valid_loss=0.305]
Epoch 49:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.3
41, train_loss_epoch=0.341, valid_loss=0.305]
Epoch 49: 100%|██████████| 1/1 [00:00<00:00,  1.49it/s, v_num=0, train_loss_
step=0.324, train_loss_epoch=0.341, valid_loss=0.305]
Validation: |          | 0/? [00:00<?, ?it/s]
Validation:   0%|          | 0/1 [00:00<?, ?it/s]
Validation DataLoader 0:   0%|          | 0/1 [00:00<?, ?it/s]
(_train_tune pid=14309)
Validation DataLoader 0: 100%|██████████| 1/1 [00:05<00:00,  0.17it/s]
Epoch 49: 100%|██████████| 1/1 [00:06<00:00,  0.15it/s, v_num=0, train_loss_
step=0.324, train_loss_epoch=0.324, valid_loss=0.301]
(_train_tune pid=14309) Seed set to 7
```

Loading [MathJax]/extensions/Safe.js

```
Sanity Checking: |          | 0/? [00:00<?, ?it/s]
Sanity Checking DataLoader 0:   0%|          | 0/1 [00:00<?, ?it/s]
Epoch 0:   0%|          | 0/1 [00:00<?, ?it/s]
Epoch 1:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.57
6, train_loss_epoch=0.576]
Epoch 1: 100%|██████████| 1/1 [00:01<00:00,  0.95it/s, v_num=0, train_loss_s
tep=0.549, train_loss_epoch=0.549]
Epoch 2:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.54
9, train_loss_epoch=0.549]
Epoch 3:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.61
6, train_loss_epoch=0.616]
Epoch 4:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.48
0, train_loss_epoch=0.480]
Epoch 5:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.43
3, train_loss_epoch=0.433]
Epoch 6:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.42
2, train_loss_epoch=0.422]
Epoch 7:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.43
1, train_loss_epoch=0.431]
Epoch 8:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.35
7, train_loss_epoch=0.357]
Epoch 9:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.40
7, train_loss_epoch=0.407]
Epoch 9: 100%|██████████| 1/1 [00:00<00:00,  1.10it/s, v_num=0, train_loss_s
tep=0.361, train_loss_epoch=0.407]
Validation: |          | 0/? [00:00<?, ?it/s]
Validation:   0%|          | 0/1 [00:00<?, ?it/s]
Validation DataLoader 0:   0%|          | 0/1 [00:00<?, ?it/s]
(_train_tune pid=14309)
Validation DataLoader 0: 100%|██████████| 1/1 [00:06<00:00,  0.15it/s]
Epoch 10:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.3
61, train_loss_epoch=0.361, valid_loss=0.364]
Epoch 11:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.3
87, train_loss_epoch=0.387, valid_loss=0.364]
Epoch 12:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.3
74, train_loss_epoch=0.374, valid_loss=0.364]
Epoch 13:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.3
20, train_loss_epoch=0.320, valid_loss=0.364]
Epoch 14:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.3
92, train_loss_epoch=0.392, valid_loss=0.364]
Epoch 15:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.3
61, train_loss_epoch=0.361, valid_loss=0.364]
Epoch 16:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.3
68, train_loss_epoch=0.368, valid_loss=0.364]
Epoch 17:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.3
73, train_loss_epoch=0.373, valid_loss=0.364]
Epoch 18:   0%|          | 0/1 [00:00<?, ?it/s, v_num=0, train_loss_step=0.4
03, train_loss_epoch=0.403, valid_loss=0.364]
```

Loading [MathJax]/extensions/Safe.js

```
---------------------------------------------------------------------------
OutOfMemoryError                          Traceback (most recent call last)
File ~/.local/lib/python3.10/site-packages/ray/air/execution/_internal/event
_manager.py:110, in RayEventManager.resolve_future(self, future)
    109 try:
--> 110     result = ray.get(future)
    111 except Exception as e:

File ~/.local/lib/python3.10/site-packages/ray/_private/auto_init_hook.py:2
4, in wrap_auto_init.<locals>.auto_init_wrapper(*args, **kwargs)
     23 auto_init_ray()
---> 24 return fn(*args, **kwargs)

File ~/.local/lib/python3.10/site-packages/ray/_private/client_mode_hook.py:
103, in client_mode_hook.<locals>.wrapper(*args, **kwargs)
    102         return getattr(ray, func.__name__)(*args, **kwargs)
--> 103 return func(*args, **kwargs)

File ~/.local/lib/python3.10/site-packages/ray/_private/worker.py:2549, in g
et(object_refs, timeout)
   2548         else:
-> 2549             raise value
   2551 if is_individual_id:

OutOfMemoryError: Task was killed due to the node running low on memory.
Memory on the node (IP: 172.17.132.167, ID: 12020b96e7d4c1257179a1a072a6d23c
1936d7110758424804cde666) where the task (actor ID: 3f4e94326b59d8d86eafbe98
01000000, name=ImplicitFunc.__init__, pid=14309, memory used=0.90GB) was run
ning was 7.24GB / 7.62GB (0.95008), which exceeds the memory usage threshold
of 0.95. Ray killed this worker (ID: 2b537369b67a0787ac408025f63343000616a0a
5d6290cb781ad8051) because it was the most recently scheduled task; to see m
ore information about memory usage on this node, use `ray logs raylet.out -i
p 172.17.132.167`. To see the logs of the worker, use `ray logs worker-2b537
369b67a0787ac408025f63343000616a0a5d6290cb781ad8051*out -ip 172.17.132.167.
Top 10 memory users:
PID     MEM(GB) COMMAND
13680   1.06    /usr/bin/python3 -m ipykernel_launcher -f /home/blair/.loca
l/share/jupyter/runtime/kernel-cc0fecc7-3...
14309   0.90    ray::ImplicitFunc.train
1447    0.43    /usr/bin/python3 -m ipykernel_launcher -f /home/blair/.loca
l/share/jupyter/runtime/kernel-795da834-6...
8455    0.23    /usr/bin/python3 -m ipykernel_launcher -f /home/blair/.loca
l/share/jupyter/runtime/kernel-19cae6ed-0...
13795   0.23    /home/blair/.local/lib/python3.10/site-packages/ray/core/sr
c/ray/raylet/raylet --raylet_socket_name=...
13739   0.23    /home/blair/.local/lib/python3.10/site-packages/ray/core/sr
c/ray/gcs/gcs_server --log_dir=/tmp/ray/s...
10774   0.10    /usr/bin/python3 -m ipykernel_launcher -f /home/blair/.loca
l/share/jupyter/runtime/kernel-aee0fec4-4...
13780   0.08    /usr/bin/python3 -u /home/blair/.local/lib/python3.10/site-p
ackages/ray/autoscaler/_private/monitor....
13781   0.07    /usr/bin/python3 /home/blair/.local/lib/python3.10/site-pack
ages/ray/dashboard/dashboard.py --host=1...
13796   0.06    /usr/bin/python3 -u /home/blair/.local/lib/python3.10/site-p
ackages/ray/_private/log_monitor.py --lo...
e documentation on how to address the out of memory issue: http
```

Loading [MathJax]/extensions/Safe.js

```
s://docs.ray.io/en/latest/ray-core/scheduling/ray-oom-prevention.html. Consi
der provisioning more memory on this node or reducing task parallelism by re
questing more CPUs per task. Set max_restarts and max_task_retries to enable
retry when the task crashes due to OOM. To adjust the kill threshold, set th
e environment variable `RAY_memory_usage_threshold` when starting Ray. To di
sable worker killing, set the environment variable `RAY_memory_monitor_refre
sh_ms` to zero.

The above exception was the direct cause of the following exception:

RuntimeError                              Traceback (most recent call last)
Cell In[3], line 28
     20 models = [AutoNHITS(h=horizon,
     21                     config=nhits_config,
     22                     num_samples=5)]
     24 nf = NeuralForecast(
     25     models=models,
     26     freq='10min')
---> 28 Y_hat_df = nf.cross_validation(df=Y_df, val_size=val_size,
     29                                 test_size=test_size, n_windows=None)

File ~/.local/lib/python3.10/site-packages/neuralforecast/core.py:520, in Ne
uralForecast.cross_validation(self, df, static_df, n_windows, step_size, val
_size, test_size, sort_df, use_init_models, verbose, **data_kwargs)
    515 fcsts = np.full(
    516     (self.dataset.n_groups * h * n_windows, len(cols)), np.nan, dtyp
e=np.float32
    517 )
    519 for model in self.models:
--> 520     model.fit(dataset=self.dataset, val_size=val_size, test_size=tes
t_size)
    521     model_fcsts = model.predict(
    522         self.dataset, step_size=step_size, **data_kwargs
    523     )
    525     # Append predictions in memory placeholder

File ~/.local/lib/python3.10/site-packages/neuralforecast/common/_base_auto.
py:361, in BaseAuto.fit(self, dataset, val_size, test_size, random_seed)
    359 val_size = val_size if val_size > 0 else self.h
    360 if self.backend == "ray":
--> 361     results = self._tune_model(
    362         cls_model=self.cls_model,
    363         dataset=dataset,
    364         val_size=val_size,
    365         test_size=test_size,
    366         cpus=self.cpus,
    367         gpus=self.gpus,
    368         verbose=self.verbose,
    369         num_samples=self.num_samples,
    370         search_alg=search_alg,
    371         config=self.config,
    372     )
    373     best_config = results.get_best_result().config
    374 else:

File ~/.local/lib/python3.10/site-packages/neuralforecast/common/_base_auto.
```

Loading [MathJax]/extensions/Safe.js

```
py:259, in BaseAuto._tune_model(self, cls_model, dataset, val_size, test_siz
e, cpus, gpus, verbose, num_samples, search_alg, config)
    240     device_dict = {"cpu": cpus}
    242 tuner = tune.Tuner(
    243     tune.with_resources(train_fn_with_parameters, device_dict),
    244     run_config=air.RunConfig(
   (...)
    257     param_space=config,
    258 )
--> 259 results = tuner.fit()
    260 return results

File ~/.local/lib/python3.10/site-packages/ray/tune/tuner.py:372, in Tuner.f
it(self)
    370 if not self._is_ray_client:
    371     try:
--> 372         return self._local_tuner.fit()
    373     except TuneError as e:
    374         raise TuneError(
    375             _TUNER_FAILED_MSG.format(
    376                 path=self._local_tuner.get_experiment_checkpoint_dir
()
    377             )
    378         ) from e

File ~/.local/lib/python3.10/site-packages/ray/tune/impl/tuner_internal.py:5
79, in TunerInternal.fit(self)
    577 param_space = copy.deepcopy(self.param_space)
    578 if not self._is_restored:
--> 579     analysis = self._fit_internal(trainable, param_space)
    580 else:
    581     analysis = self._fit_resume(trainable, param_space)

File ~/.local/lib/python3.10/site-packages/ray/tune/impl/tuner_internal.py:6
99, in TunerInternal._fit_internal(self, trainable, param_space)
    686 """Fitting for a fresh Tuner."""
    687 args = {
    688     **self._get_tune_run_arguments(trainable),
    689     **dict(
   (...)
    697     **self._tuner_kwargs,
    698 }
--> 699 analysis = run(
    700     **args,
    701 )
    702 self.clear_remote_string_queue()
    703 return analysis

File ~/.local/lib/python3.10/site-packages/ray/tune/tune.py:1103, in run(run
_or_experiment, name, metric, mode, stop, time_budget_s, config, resources_p
er_trial, num_samples, storage_path, storage_filesystem, search_alg, schedul
er, checkpoint_config, verbose, progress_reporter, log_to_file, trial_name_c
reator, trial_dirname_creator, sync_config, export_formats, max_failures, fa
il_fast, restore, server_port, resume, reuse_actors, raise_on_failed_trial,
callbacks, max_concurrent_trials, keep_checkpoints_num, checkpoint_score_att
r, checkpoint_freq, checkpoint_at_end, chdir_to_trial_dir, local_dir, _exper
```

Loading [MathJax]/extensions/Safe.js

```
iment_checkpoint_dir, _remote, _remote_string_queue, _entrypoint)
   1099 try:
   1100     while (
   1101         not runner.is_finished() and not experiment_interrupted_even
t.is_set()
   1102     ):
-> 1103         runner.step()
   1104         if has_verbosity(Verbosity.V1_EXPERIMENT):
   1105             _report_progress(runner, progress_reporter)

File ~/.local/lib/python3.10/site-packages/ray/tune/execution/tune_controlle
r.py:850, in TuneController.step(self)
    847 self._maybe_add_actors()
    849 # Handle one event
--> 850 if not self._actor_manager.next(timeout=0.1):
    851     # If there are no actors running, warn about potentially
    852     # insufficient resources
    853     if not self._actor_manager.num_live_actors:
    854         self._insufficient_resources_manager.on_no_available_trials(
    855             self.get_trials()
    856         )

File ~/.local/lib/python3.10/site-packages/ray/air/execution/_internal/actor
_manager.py:224, in RayActorManager.next(self, timeout)
    222     self._actor_state_events.resolve_future(future)
    223 elif future in actor_task_futures:
--> 224     self._actor_task_events.resolve_future(future)
    225 else:
    226     self._handle_ready_resource_future()

File ~/.local/lib/python3.10/site-packages/ray/air/execution/_internal/event
_manager.py:113, in RayEventManager.resolve_future(self, future)
    111 except Exception as e:
    112     if on_error:
--> 113         on_error(e)
    114     else:
    115         raise e

File ~/.local/lib/python3.10/site-packages/ray/air/execution/_internal/actor
_manager.py:770, in RayActorManager._schedule_tracked_actor_task.<locals>.on
_error(exception)
    769 def on_error(exception: Exception):
--> 770     self._actor_task_failed(
    771         tracked_actor_task=tracked_actor_task, exception=exception
    772     )

File ~/.local/lib/python3.10/site-packages/ray/air/execution/_internal/actor
_manager.py:291, in RayActorManager._actor_task_failed(self, tracked_actor_t
ask, exception)
    289         tracked_actor_task._on_error(tracked_actor, exception)
    290 else:
--> 291     raise RuntimeError(
    292         f"Caught unexpected exception: {exception}"
    293     ) from exception

RuntimeError: Caught unexpected exception: Task was killed due to the node r
```

Loading [MathJax]/extensions/Safe.js

```
unning low on memory.
Memory on the node (IP: 172.17.132.167, ID: 12020b96e7d4c1257179a1a072a6d23c
1936d7110758424804cde666) where the task (actor ID: 3f4e94326b59d8d86eafbe98
01000000, name=ImplicitFunc.__init__, pid=14309, memory used=0.90GB) was run
ning was 7.24GB / 7.62GB (0.95008), which exceeds the memory usage threshold
of 0.95. Ray killed this worker (ID: 2b537369b67a0787ac408025f63343000616a0a
5d6290cb781ad8051) because it was the most recently scheduled task; to see m
ore information about memory usage on this node, use `ray logs raylet.out -i
p 172.17.132.167`. To see the logs of the worker, use `ray logs worker-2b537
369b67a0787ac408025f63343000616a0a5d6290cb781ad8051*out -ip 172.17.132.167.
Top 10 memory users:
PID     MEM(GB) COMMAND
13680   1.06    /usr/bin/python3 -m ipykernel_launcher -f /home/blair/.loca
l/share/jupyter/runtime/kernel-cc0fecc7-3...
14309   0.90    ray::ImplicitFunc.train
1447    0.43    /usr/bin/python3 -m ipykernel_launcher -f /home/blair/.loca
l/share/jupyter/runtime/kernel-795da834-6...
8455    0.23    /usr/bin/python3 -m ipykernel_launcher -f /home/blair/.loca
l/share/jupyter/runtime/kernel-19cae6ed-0...
13795   0.23    /home/blair/.local/lib/python3.10/site-packages/ray/core/sr
c/ray/raylet/raylet --raylet_socket_name=...
13739   0.23    /home/blair/.local/lib/python3.10/site-packages/ray/core/sr
c/ray/gcs/gcs_server --log_dir=/tmp/ray/s...
10774   0.10    /usr/bin/python3 -m ipykernel_launcher -f /home/blair/.loca
l/share/jupyter/runtime/kernel-aee0fec4-4...
13780   0.08    /usr/bin/python3 -u /home/blair/.local/lib/python3.10/site-p
ackages/ray/autoscaler/_private/monitor....
13781   0.07    /usr/bin/python3 /home/blair/.local/lib/python3.10/site-pack
ages/ray/dashboard/dashboard.py --host=1...
13796   0.06    /usr/bin/python3 -u /home/blair/.local/lib/python3.10/site-p
ackages/ray/_private/log_monitor.py --lo...
Refer to the documentation on how to address the out of memory issue: http
s://docs.ray.io/en/latest/ray-core/scheduling/ray-oom-prevention.html. Consi
der provisioning more memory on this node or reducing task parallelism by re
questing more CPUs per task. Set max_restarts and max_task_retries to enable
retry when the task crashes due to OOM. To adjust the kill threshold, set th
e environment variable `RAY_memory_usage_threshold` when starting Ray. To di
sable worker killing, set the environment variable `RAY_memory_monitor_refre
sh_ms` to zero.
```

In [2]:
```python
from neuralforecast.losses.numpy import mae, mse # INPUT SIZE = 96

print('MAE: ', mae(Y_hat_df['y'], Y_hat_df['AutoNHITS']))
print('MSE: ', mse(Y_hat_df['y'], Y_hat_df['AutoNHITS']))
```

```
MAE:  0.23199070626996326
MSE:  0.20198460555905476
```

NHITS - NOT AUTONHITS

In [10]:
```python
horizon = 96

models = [NHITS(h=horizon,
                input_size = 5*horizon,
                batch_size = 21,
```

Loading [MathJax]/extensions/Safe.js

```
                windows_batch_size = 256,
                max_steps = 200,
                val_check_steps=20
               )]

nf = NeuralForecast(
    models=models,
    freq='10min')

Y_hat_df = nf.cross_validation(df=Y_df, val_size=val_size,
                               test_size=test_size, n_windows=None)
```

```
Seed set to 1
2023-11-02 18:22:24.437749: I tensorflow/core/util/port.cc:111] oneDNN custo
m operations are on. You may see slightly different numerical results due to
floating-point round-off errors from different computation orders. To turn t
hem off, set the environment variable `TF_ENABLE_ONEDNN_OPTS=0`.
2023-11-02 18:22:24.469122: I tensorflow/tsl/cuda/cudart_stub.cc:28] Could n
ot find cuda drivers on your machine, GPU will not be used.
2023-11-02 18:22:24.617762: E tensorflow/compiler/xla/stream_executor/cuda/c
uda_dnn.cc:9342] Unable to register cuDNN factory: Attempting to register fa
ctory for plugin cuDNN when one has already been registered
2023-11-02 18:22:24.617786: E tensorflow/compiler/xla/stream_executor/cuda/c
uda_fft.cc:609] Unable to register cuFFT factory: Attempting to register fac
tory for plugin cuFFT when one has already been registered
2023-11-02 18:22:24.618739: E tensorflow/compiler/xla/stream_executor/cuda/c
uda_blas.cc:1518] Unable to register cuBLAS factory: Attempting to register
factory for plugin cuBLAS when one has already been registered
2023-11-02 18:22:24.698294: I tensorflow/core/platform/cpu_feature_guard.cc:
182] This TensorFlow binary is optimized to use available CPU instructions i
n performance-critical operations.
To enable the following instructions: AVX2 AVX_VNNI FMA, in other operation
s, rebuild TensorFlow with the appropriate compiler flags.
2023-11-02 18:22:25.704940: W tensorflow/compiler/tf2tensorrt/utils/py_util
s.cc:38] TF-TRT Warning: Could not find TensorRT
Sanity Checking: |
| 0/? [00:00…
Training: |
| 0/? [00:00…
Validation: |
| 0/? [00:00…
Validation: |
| 0/? [00:00…
Validation: |
| 0/? [00:00…
Validation: |
| 0/? [00:00…
Validation: |
| 0/? [00:00…
Validation: |
| 0/? [00:00…
Validation: |
| 0/? [00:00…
Validation: |
| 0/? [00:00…
Validation: |
| 0/? [00:00…
```

Loading [MathJax]/extensions/Safe.js

```
Validation: |
| 0/? [00:00…
Validation: |
| 0/? [00:03…
Predicting: |
| 0/? [00:00…
```

In [11]:
```
from neuralforecast.losses.numpy import mae, mse

print('MAE: ', mae(Y_hat_df['y'], Y_hat_df['NHITS']))
print('MSE: ', mse(Y_hat_df['y'], Y_hat_df['NHITS']))
```

```
MAE:  0.19757447049855434
MSE:  0.16556661850686696
```

In [12]:
```
Y_hat_df.to_csv('results/Weather/NHITS.csv')
```

In [14]:
```
data = {'Informer_MSE': mse(Y_hat_df['y'], Y_hat_df['NHITS']),
        'Informer_MAE': mae(Y_hat_df['y'], Y_hat_df['NHITS'])}

df = pd.DataFrame(data, index=['Weather'])
df.to_csv('results/Weather/df_NHITS.csv')
```

Loading [MathJax]/extensions/Safe.js