

Hyperparameters for various deep learning models

PatchTST model

Parameters:

- **context_length** (*int*, *default* = 96) – Number of time units that condition the predictions
- **patch_len** (*int*, *default* = 16) – Length of the patch.
- **stride** (*int*, *default* = 8) – Stride of the patch.
- **d_model** (*int*, *default* = 32) – Size of hidden layers in the Transformer encoder.
- **nhead** (*int*, *default* = 4) – Number of attention heads in the Transformer encoder which must divide d_model.
- **num_encoder_layers** (*int*, *default* = 2) – Number of layers in the Transformer encoder.
- **distr_output** (*gluonts.torch.distributions.DistributionOutput*, *default* = *StudentTOutput()*) – Distribution to fit.
- **scaling** (*{"mean", "std", None}*, *default* = "mean") –

Scaling applied to each *context window* during training & prediction. One of "mean" (mean absolute scaling), "std" (standardization), None (no scaling).

Note that this is different from the *target_scaler* that is applied to the *entire time series*.

- **max_epochs** (*int*, *default* = 100) – Number of epochs the model will be trained for
- **batch_size** (*int*, *default* = 64) – Size of batches used during training
- **num_batches_per_epoch** (*int*, *default* = 50) – Number of batches processed every epoch
- **lr** (*float*, *default* = 1e-3,) – Learning rate used during training
- **weight_decay** (*float*, *default* = 1e-8) – Weight decay regularization parameter.
- **keep_lightning_logs** (*bool*, *default* = *False*) – If True, lightning_logs directory will NOT be removed after the model finished training.

TiDE Model

Parameters:

- **context_length** (*int*, *default* = $\max(64, 2 * \text{prediction_length})$) – Number of past values used for prediction.
- **disable_static_features** (*bool*, *default* = *False*) – If True, static features won't be used by the model even if they are present in the dataset. If False, static features will be used by the model if they are present in the dataset.
- **disable_known_covariates** (*bool*, *default* = *False*) – If True, known covariates won't be used by the model even if they are present in the dataset. If False, known covariates will be used by the model if they are present in the dataset.
- **disable_past_covariates** (*bool*, *default* = *False*) – If True, past covariates won't be used by the model even if they are present in the dataset. If False, past covariates will be used by the model if they are present in the dataset.
- **feat_proj_hidden_dim** (*int*, *default* = 4) – Size of the feature projection layer.
- **encoder_hidden_dim** (*int*, *default* = 64) – Size of the dense encoder layer.
- **decoder_hidden_dim** (*int*, *default* = 64) – Size of the dense decoder layer.
- **temporal_hidden_dim** (*int*, *default* = 64) – Size of the temporal decoder layer.
- **distr_hidden_dim** (*int*, *default* = 64) – Size of the distribution projection layer.
- **num_layers_encoder** (*int*, *default* = 2) – Number of layers in dense encoder.
- **num_layers_decoder** (*int*, *default* = 2) – Number of layers in dense decoder.
- **decoder_output_dim** (*int*, *default* = 16) – Output size of the dense decoder.
- **dropout_rate** (*float*, *default* = 0.2) – Dropout regularization parameter.
- **num_feat_dynamic_proj** (*int*, *default* = 2) – Output size of feature projection layer.
- **embedding_dimension** (*int*, *default* = $[16] * \text{num_feat_static_cat}$) – Dimension of the embeddings for categorical features
- **layer_norm** (*bool*, *default* = *True*) – Should layer normalization be enabled?
- **scaling** (*{*"mean", "std", None*}*, *default* = *"mean"*) –

Scaling applied to each *context window* during training & prediction. One of "mean" (mean absolute scaling), "std" (standardization), None (no scaling).

Note that this is different from the *target_scaler* that is applied to the *entire time series*.

- **max_epochs** (*int*, *default* = 100) – Number of epochs the model will be trained for
- **batch_size** (*int*, *default* = 256) – Size of batches used during training
- **predict_batch_size** (*int*, *default* = 500) – Size of batches used during prediction.
- **num_batches_per_epoch** (*int*, *default* = 50) – Number of batches processed every epoch

- **lr** (*float, default = 1e-4,*) – Learning rate used during training
- **trainer_kwargs** (*dict, optional*) – Optional keyword arguments passed to `lightning.Trainer`.
- **early_stopping_patience** (*int or None, default = 20*) – Early stop training if the validation loss doesn't improve for this many epochs.
- **keep_lightning_logs** (*bool, default = False*) – If True, `lightning_logs` directory will NOT be removed after the model finished training.

Simple FeedForward Neural Network Model

Parameters:

- **context_length** (*int, default = max(10, 2 * prediction_length)*) – Number of time units that condition the predictions
- **hidden_dimensions** (*List[int], default = [20, 20]*) – Size of hidden layers in the feedforward network
- **distr_output** (*gluonts.torch.distributions.DistributionOutput, default = StudentTOutput()*) – Distribution to fit.
- **batch_normalization** (*bool, default = False*) – Whether to use batch normalization
- **mean_scaling** (*bool, default = True*) – If True, mean absolute scaling will be applied to each *context window* during training & prediction. Note that this is different from the *target_scaler* that is applied to the *entire time series*.
- **max_epochs** (*int, default = 100*) – Number of epochs the model will be trained for
- **batch_size** (*int, default = 64*) – Size of batches used during training
- **predict_batch_size** (*int, default = 500*) – Size of batches used during prediction.
- **num_batches_per_epoch** (*int, default = 50*) – Number of batches processed every epoch
- **lr** (*float, default = 1e-3,*) – Learning rate used during training
- **trainer_kwargs** (*dict, optional*) – Optional keyword arguments passed to `lightning.Trainer`.
- **early_stopping_patience** (*int or None, default = 20*) – Early stop training if the validation loss doesn't improve for this many epochs.
- **keep_lightning_logs** (*bool, default = False*) – If True, `lightning_logs` directory will NOT be removed after the model finished training.

ETS Model

Parameters:

- **model** (*str, default = "AAA"*) – Model string describing the configuration of the E (error), T (trend) and S (seasonal) model components. Each component can be one of "M" (multiplicative), "A" (additive), "N" (omitted). For example when model="ANN" (additive error, no trend, and no seasonality), ETS will explore only a simple exponential smoothing.
- **seasonal_period** (*int or None, default = None*) – Number of time steps in a complete seasonal cycle for seasonal models. For example, 7 for daily data with a weekly cycle or 12 for monthly data with an annual cycle. When set to None, seasonal_period will be inferred from the frequency of the training data. Can also be specified manually by providing an integer > 1. If seasonal_period (inferred or provided) is equal to 1, seasonality will be disabled.
- **damped** (*bool, default = False*) – Whether to dampen the trend.
- **n_jobs** (*int or float, default = 0.5*) – Number of CPU cores used to fit the models in parallel. When set to a float between 0.0 and 1.0, that fraction of available CPU cores is used. When set to a positive integer, that many cores are used. When set to -1, all CPU cores are used.
- **max_ts_length** (*int, default = 2500*) – If not None, only the last max_ts_length time steps of each time series will be used to train the model. This significantly speeds up fitting and usually leads to no change in accuracy.

Seasonal Naive Model

Parameters:

- **seasonal_period** (*int or None, default = None*) – Number of time steps in a complete seasonal cycle for seasonal models. For example, 7 for daily data with a weekly cycle or 12 for monthly data with an annual cycle. When set to None, seasonal_period will be inferred from the frequency of the training data. Can also be specified manually by providing an integer > 1. If seasonal_period (inferred or provided) is equal to 1, will fall back to Naive forecast. Seasonality will also be disabled, if the length of the time series is < seasonal_period.
- **n_jobs** (*int or float, default = 0.5*) – Number of CPU cores used to fit the models in parallel. When set to a float between 0.0 and 1.0, that fraction of available CPU cores is used. When set to a positive integer, that many cores are used. When set to -1, all CPU cores are used.