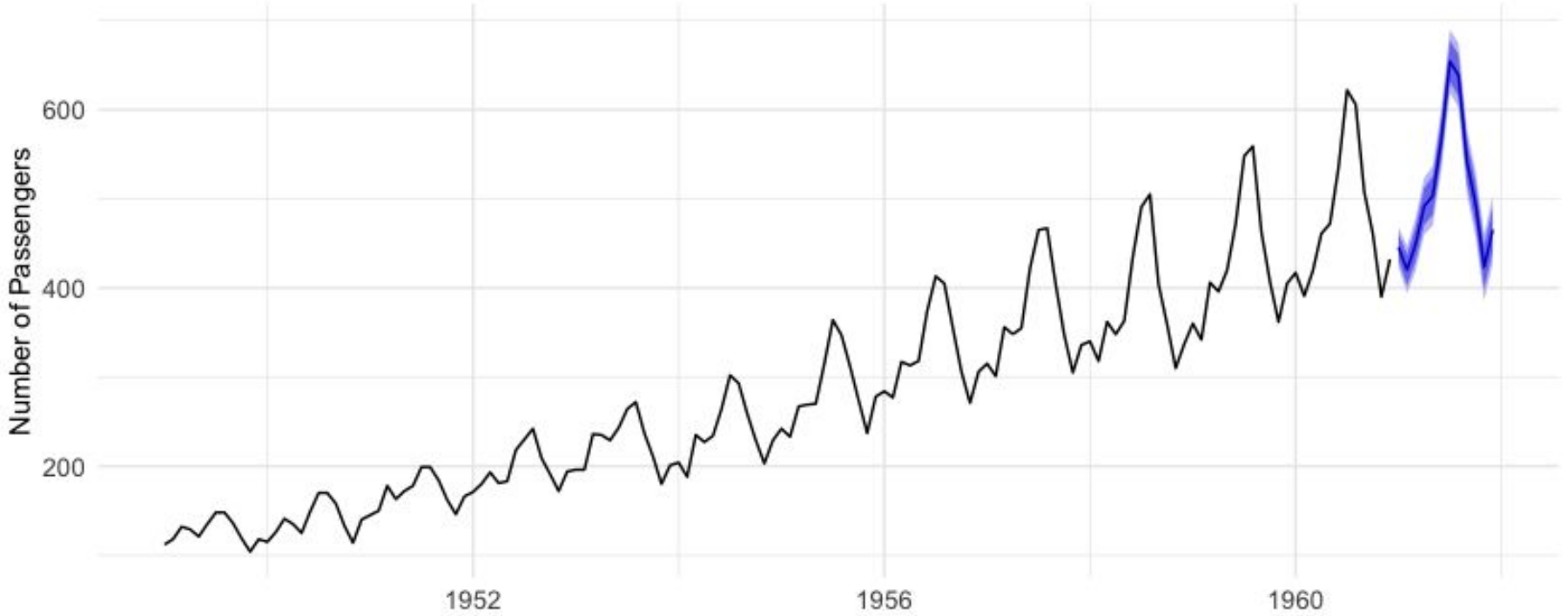
Time Series (TS)


Number of Passengers
600
400
200
1952
1956
1960

Cleaning

- Timezone? UTC?
- Take care with daylight saving times!

Feature engineering

- Datetime
  - Year, month, day, hour, minute, day of the week, day of the year, is weekend, is holiday, season…
  - Pandas already offers a lot of functions:
    - https://pandas.pydata.org/pandas-docs/stable/reference/api/pandas.Timestamp.html

Feature engineering

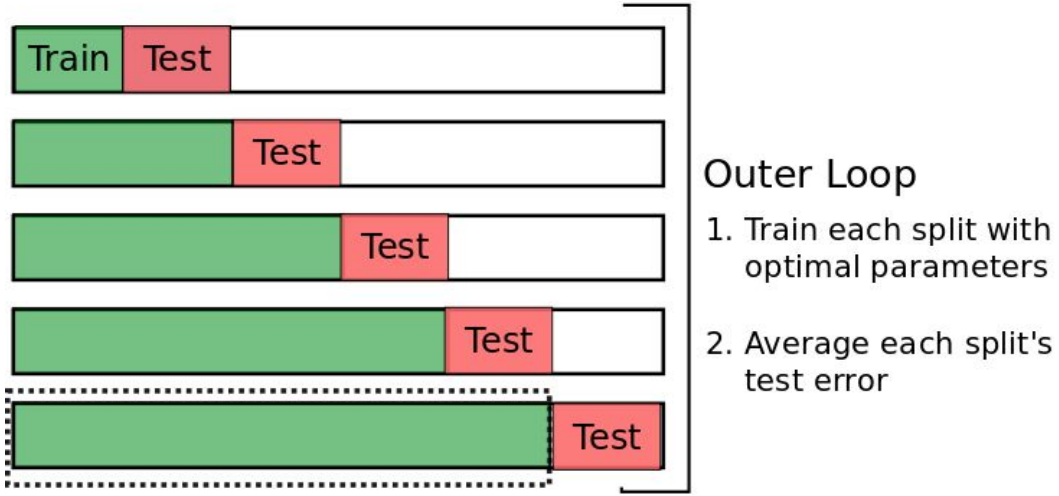
- Datetime
  - Year, month, day, hour, minute, day of the week, day of the year, is weekend, is holiday, season…
  - Pandas already offers a lot of functions:
    - https://pandas.pydata.org/pandas-docs/stable/reference/api/pandas.Timestamp.html
- Lag
  - Use data from previous times (previous hour/day…).

Feature engineering

- Datetime
  - Year, month, day, hour, minute, day of the week, day of the year, is weekend, is holiday, season…
  - Pandas already offers a lot of functions:
    - https://pandas.pydata.org/pandas-docs/stable/reference/api/pandas.Timestamp.html
- Lag
  - Use data from previous times (previous hour/day…).
- Rolling / Expanding windows
  - Mean, weighted mean, median, mode, standard deviation of the last 10/100… elements.

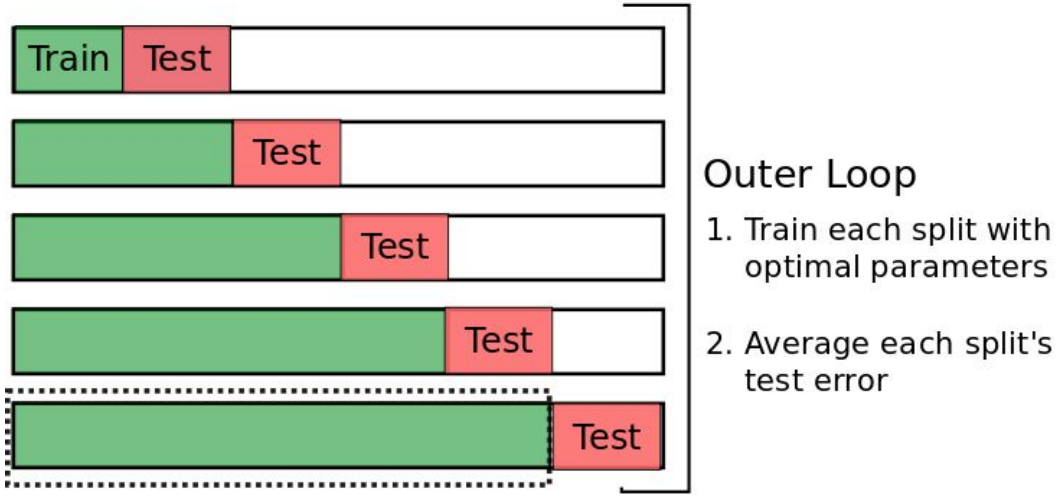
Train / Test

Train / Validation

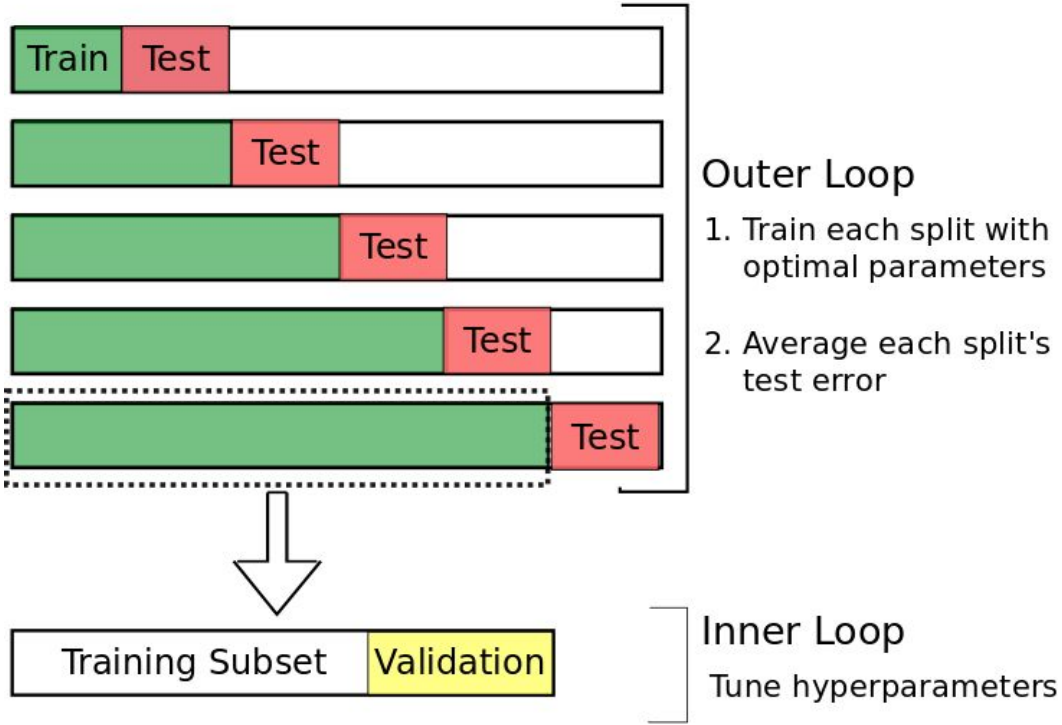


Train / Validation

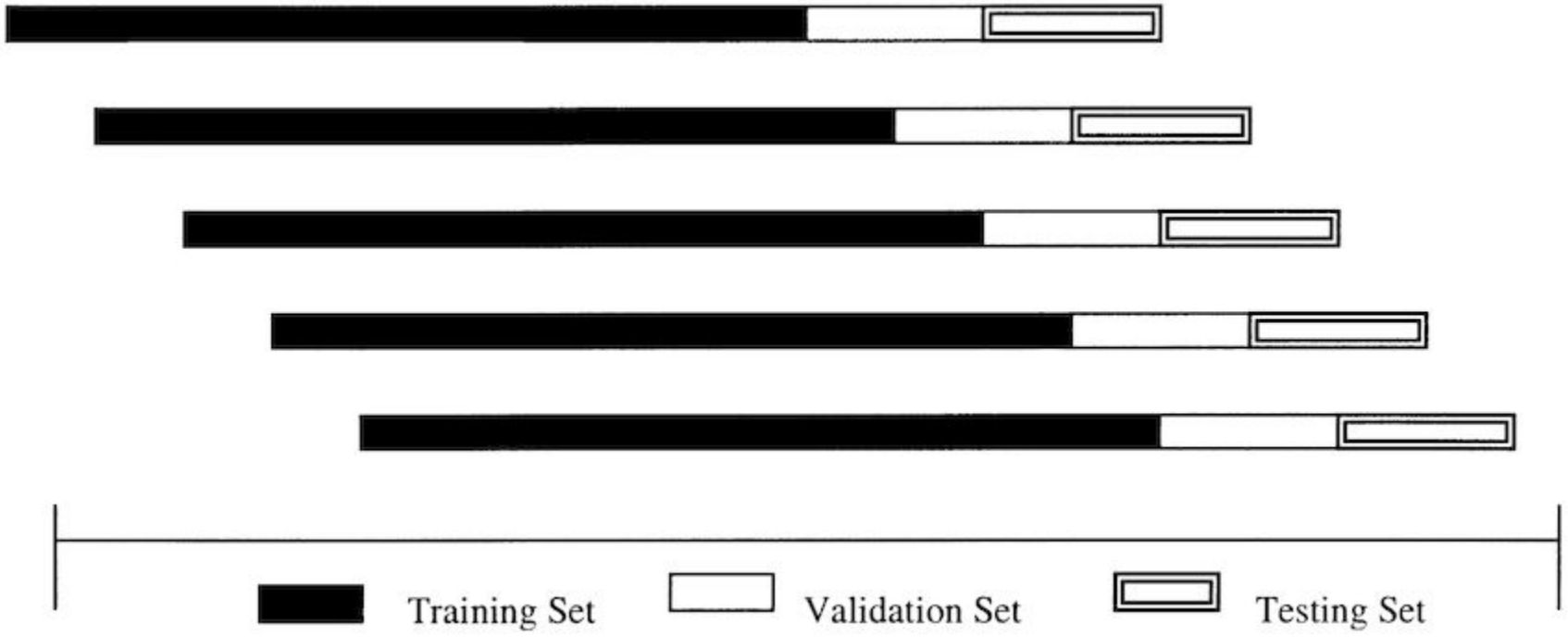
Best hyperparameters?



Walk-forward (nested) cross-validation



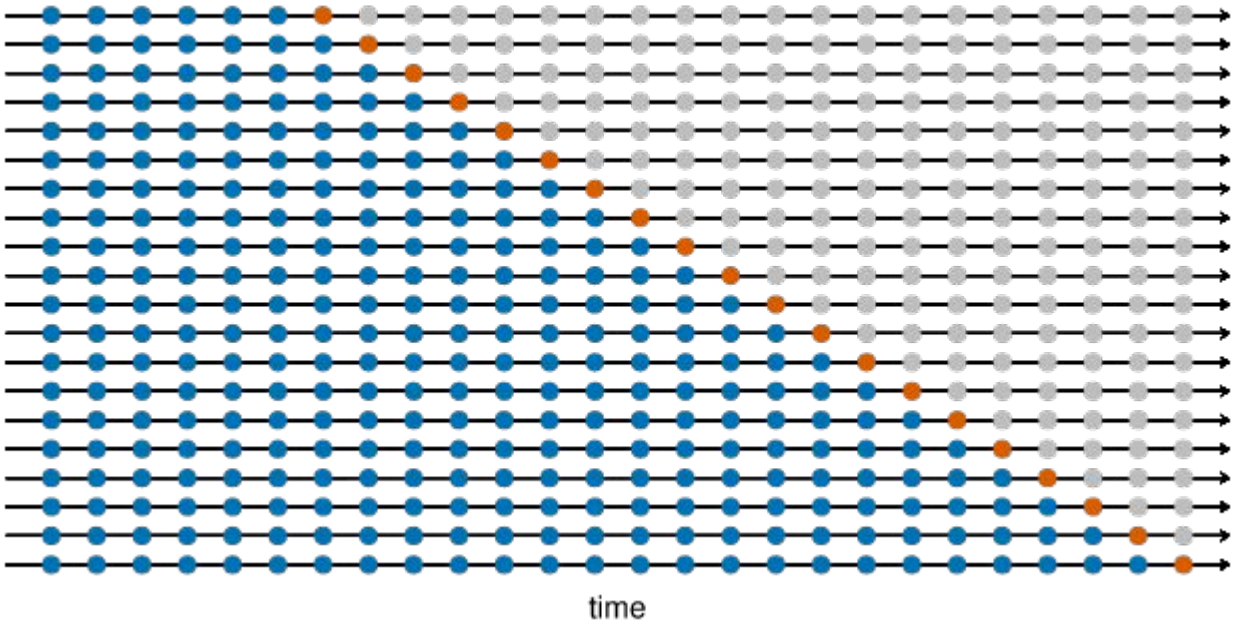
Walk-forward (nested) cross-validation



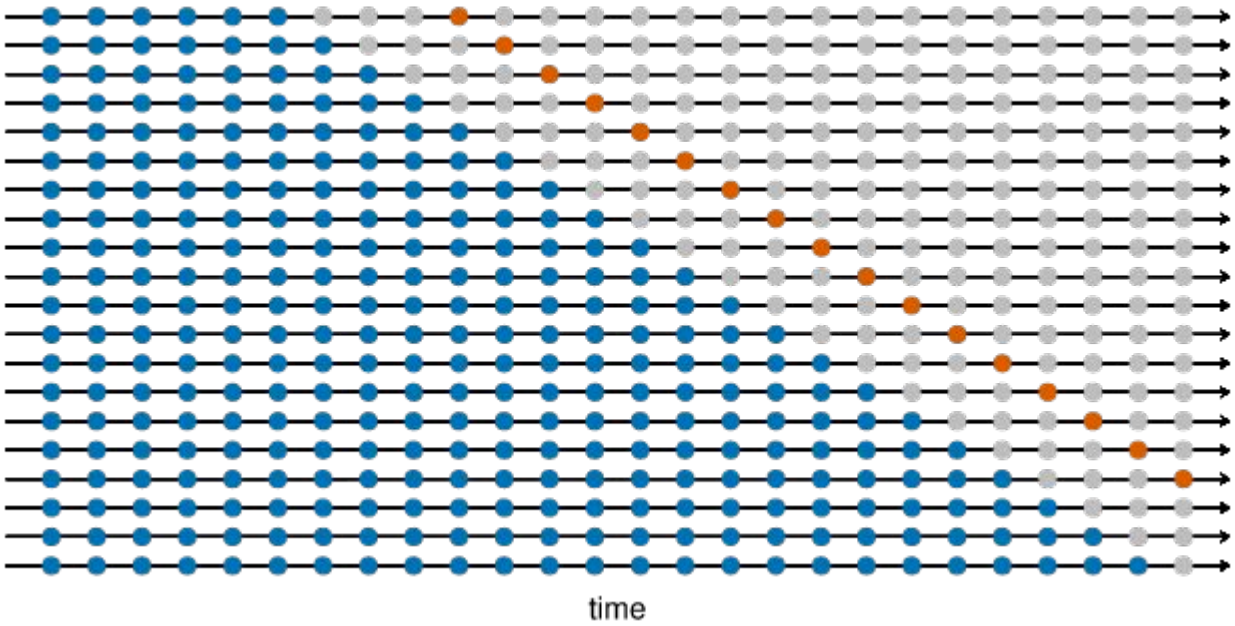
Walk-forward (nested) cross-validation

- Simulation of what would happen if we train our model every T time,
- and use the trained model for prediction in subsequent N time.

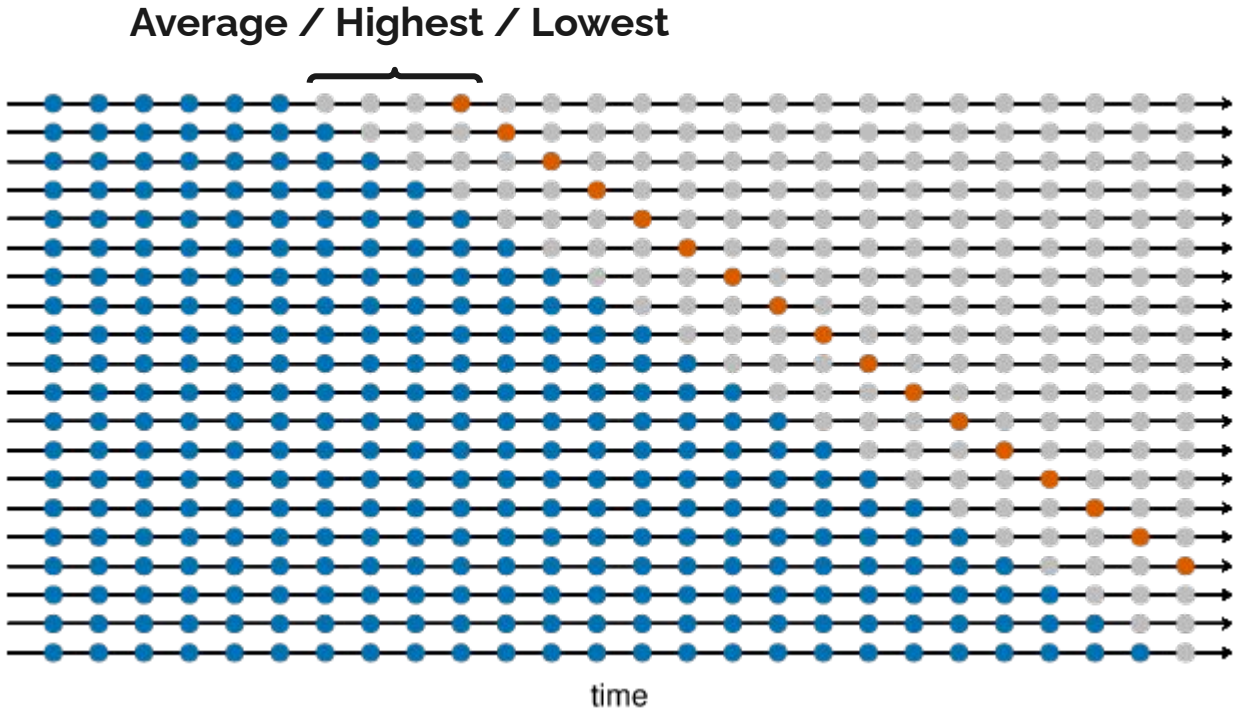
More initial training data



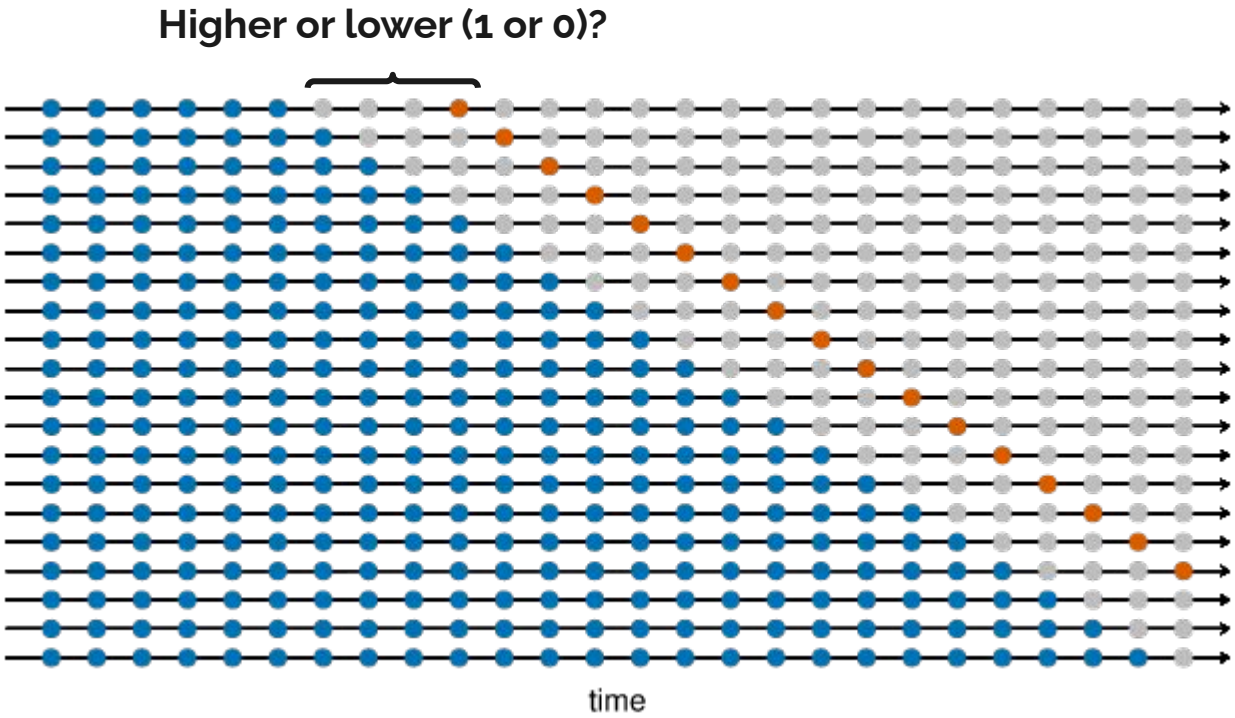
Prediction window / Forecast horizon



Prediction window / Forecast horizon



Prediction window / Forecast horizon



Univariate data

	Today's temperature
25/11/22	10
26/11/22	9
27/11/22	7
28/11/22	9
29/11/22	8

Univariate data – Create features

X

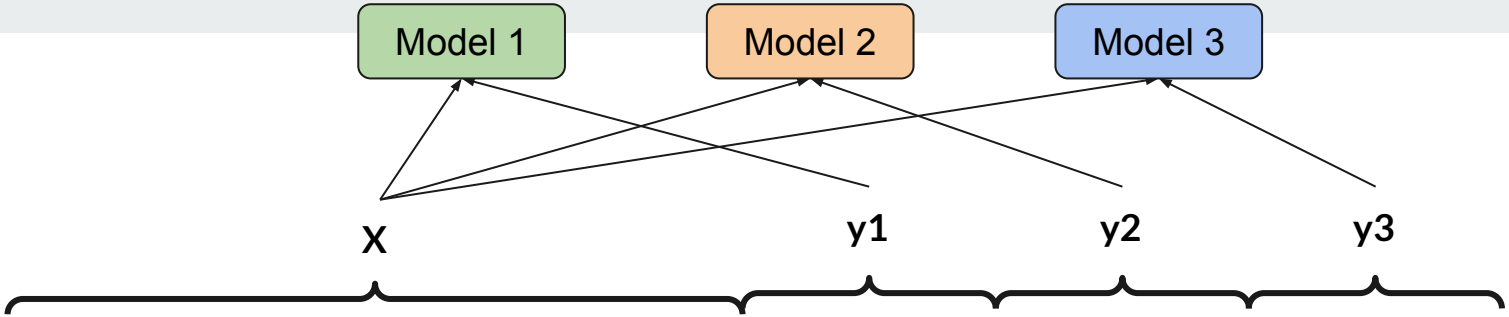
	-2 days temperature	-1 days temperature	Today's temperature
25/11/22	11	12	10
26/11/22	12	10	9
27/11/22	10	9	7
28/11/22	9	7	9
29/11/22	7	9	8

Univariate data – Target feature (1 timestep)

	-2 days temperature	-1 days temperature	Today's temperature	Temperature in 1 day
25/11/22	11	12	10	9
26/11/22	12	10	9	7
27/11/22	10	9	7	9
28/11/22	9	7	9	8
29/11/22	7	9	8	NaN

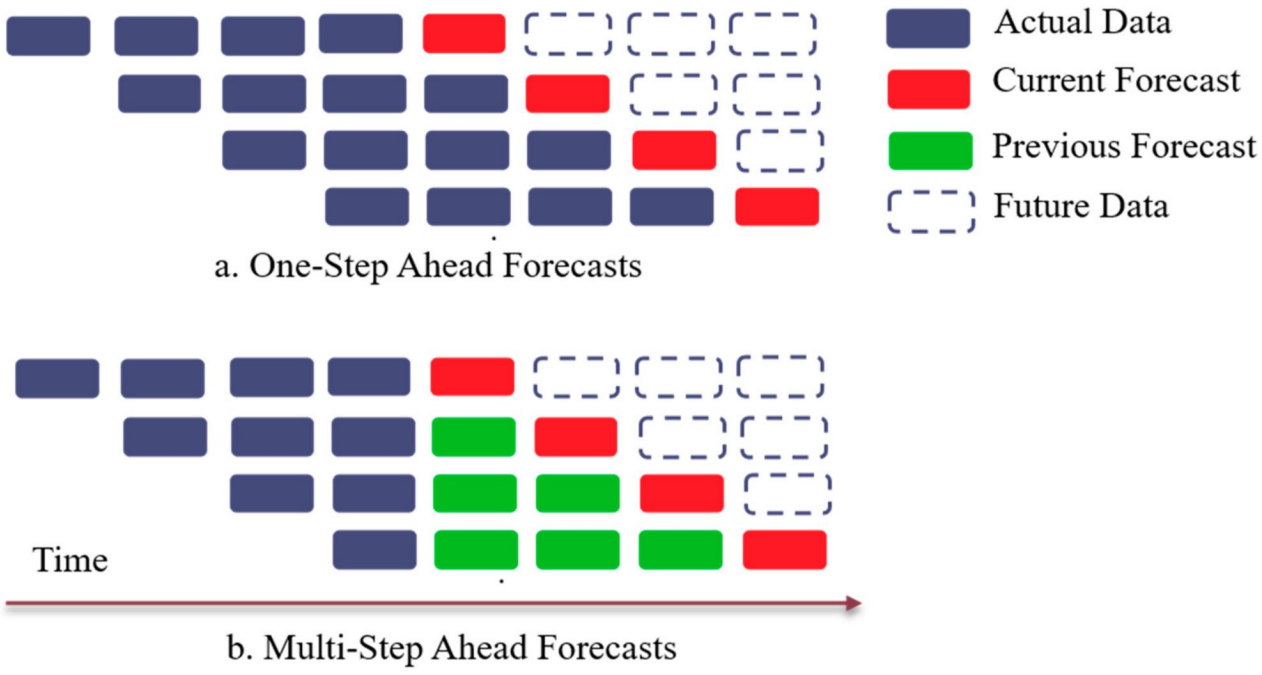
Univariate data – Target feature (1, 2, 3 timesteps)

	-2 days temperature	-1 days temperature	Today's temperature	Temperature in 1 day	Temperature in 2 days	Temperature in 3 days
25/11/22	11	12	10	9	7	9
26/11/22	12	10	9	7	9	8
27/11/22	10	9	7	9	8	NaN
28/11/22	9	7	9	8	NaN	NaN
29/11/22	7	9	8	NaN	NaN	NaN



	-2 days temperature	-1 days temperature	Today's temperature	Temperature in 1 day	Temperature in 2 days	Temperature in 3 days
25/11/22	11	12	10	9	7	9
26/11/22	12	10	9	7	9	8
27/11/22	10	9	7	9	8	NaN
28/11/22	9	7	9	8	NaN	NaN
29/11/22	7	9	8	NaN	NaN	NaN

Prediction: multi-step ahead



X y

	-2 days temperature	-1 days temperature	Today's temperature	Temperature in 1 day
25/11/22	11	12	10	9
26/11/22	12	10	9	7
27/11/22	10	9	7	9
28/11/22	9	7	9	8
29/11/22	7	9	8	NaN

X y

	-2 days temperature	-1 days temperature	Today's temperature	Temperature in 1 day
25/11/22	11	12	10	9
26/11/22	12	10	9	7
27/11/22	10	9	7	9
28/11/22	9	7	9	8
29/11/22	7	9	8	**Predict 1**

Train

X y

	-2 days temperature	-1 days temperature	Today's temperature	Temperature in 1 day
25/11/22	11	12	10	9
26/11/22	12	10	9	7
27/11/22	10	9	7	9
28/11/22	9	7	9	8
29/11/22	7	9	8	**Predict 1**
30/11/22	9	8	**Predict 1**	**Predict 2**

Train

X y

	-2 days temperature	-1 days temperature	Today's temperature	Temperature in 1 day
25/11/22	11	12	10	9
26/11/22	12	10	9	7
27/11/22	10	9	7	9
28/11/22	9	7	9	8
29/11/22	7	9	8	**Predict 1**
30/11/22	9	8	**Predict 1**	**Predict 2**
31/11/22	8	**Predict 1**	**Predict 2**	**Predict 3**

Train

Multivariate

- Similar to univariate, applying the same principles to all features.