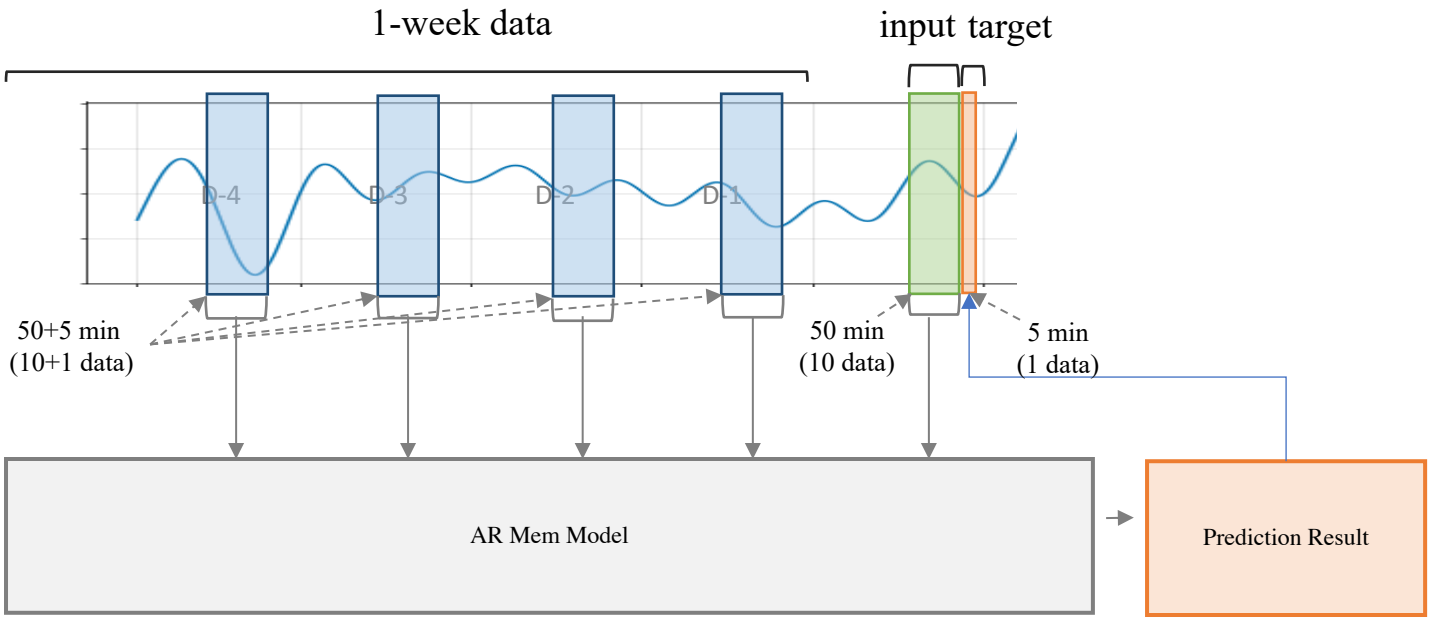
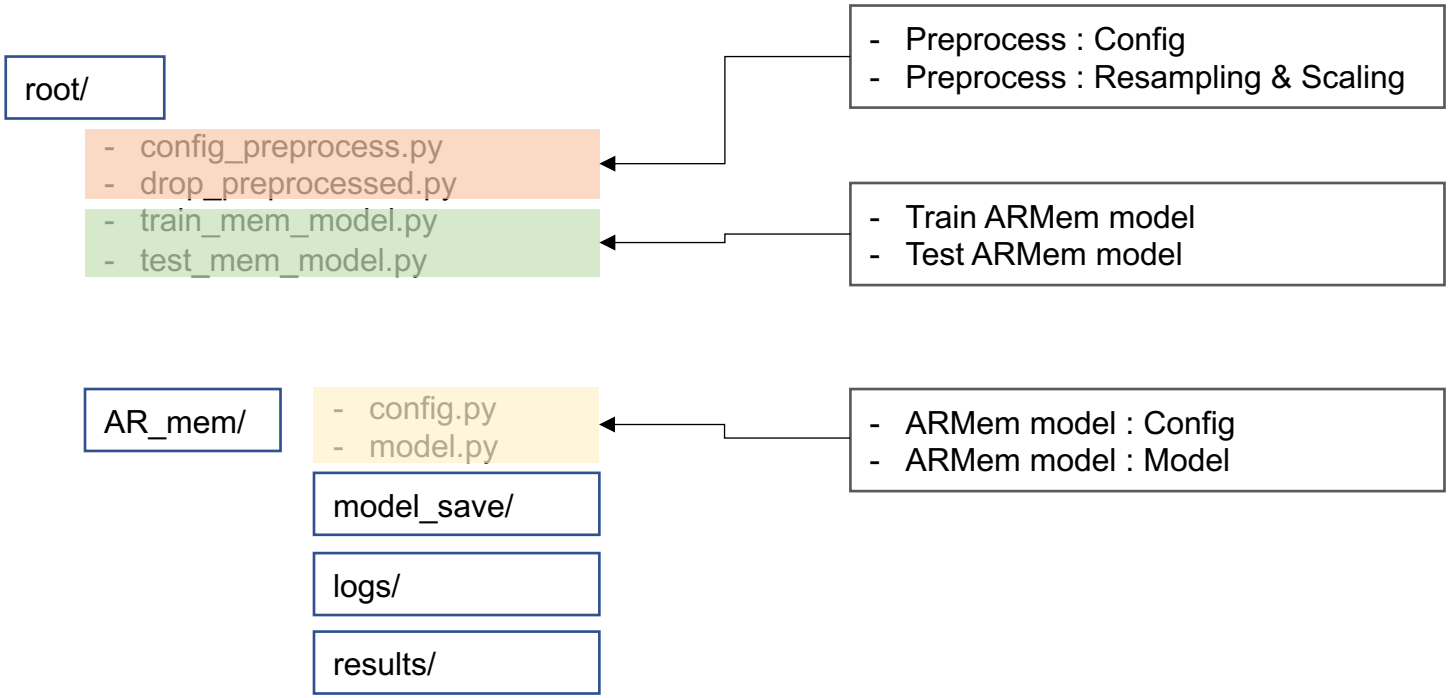
ARMem Model : Architecture

- Model architecture with **Memory Components**
  - 5-minute average value data is used.
  - ‘target’ is the value after ‘input’ to be predicted.
  - ‘input’ is the last 50 minutes data.
  - ‘1-week data’ is 50 + 5 minutes worth of data in the same time zone as ‘input’ and ‘target’.



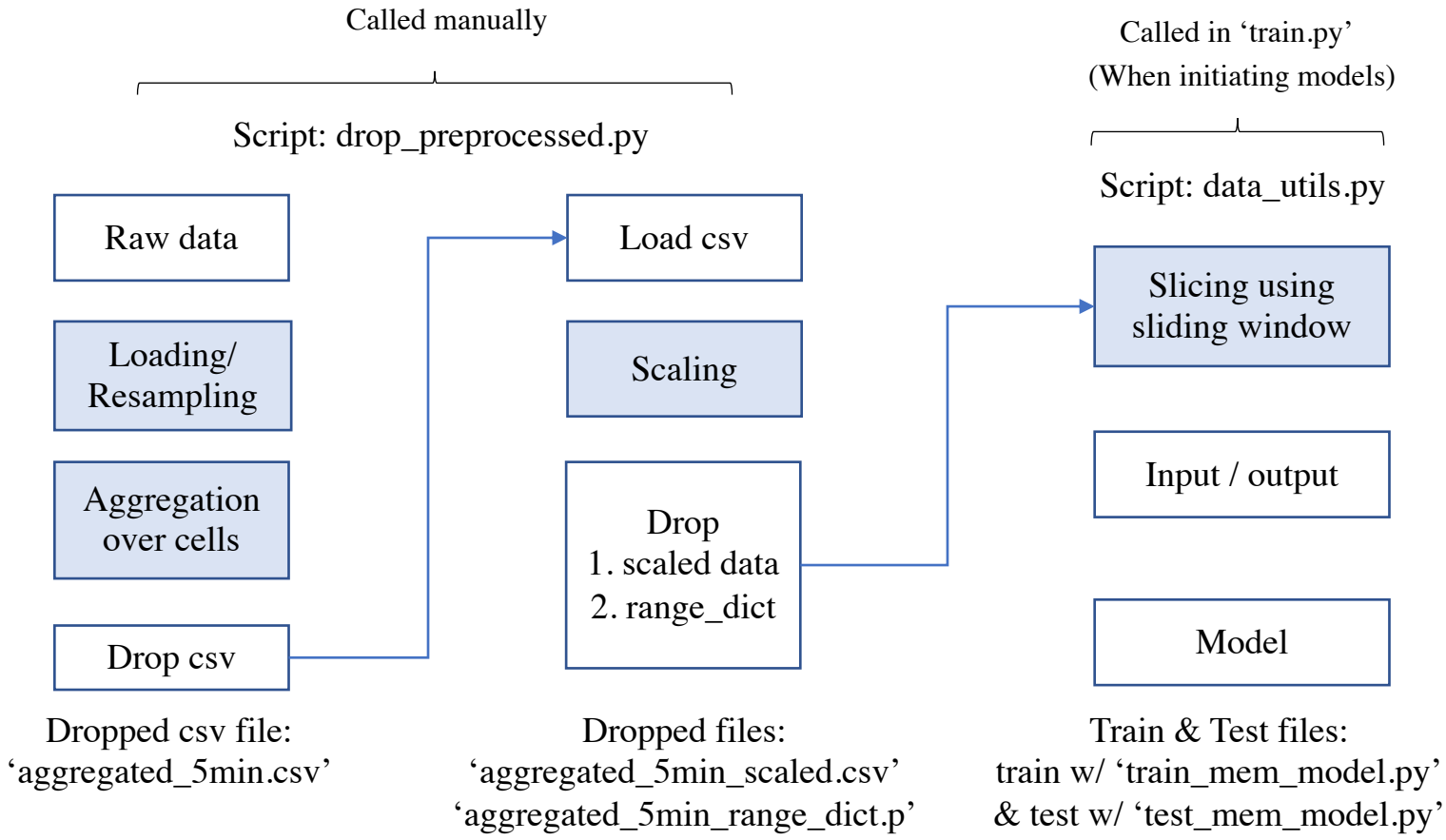
Preprocess : Scripts

- Source Code Overview



Preprocess : Overview

- Overview
  - config_preprocess.py



Preprocess : Step 1

- Resampling and Aggregation
  - Resample to 5 min & Aggregate over every cell

Raw data (one cell)

Timestamp	Col1
2019-05-01 00:00	5
2019-05-01 00:01	4
2019-05-01 00:02	NA
2019-05-01 00:03	6
2019-05-01 00:04	8
2019-05-01 00:05	1
2019-05-01 00:06	2
…	…

Average

Average

Resampling

Timestamp	Col1
2019-05-01 00:00	5.4
2019-05-01 00:05	2.5
2019-05-01 00:10	…
2019-05-01 00:15	…
…	…
…	…
…	…
…	…

Aggregate all cells

Timestamp	Col1	Cell
2019-05-01 00:00	5.4	0
2019-05-01 00:05	2.5	0
…	…	…
2019-05-01 00:00	6.6	1
2019-05-01 00:05	3.4	1
…	…	…
…	…	…
2019-05-31 23:55	10.8	19

Preprocess : Step 2

- Slicing using sliding window
  - Generate train/test data using sliding window

Scaled data (one cell)

Timestamp	Col1
2019-05-01 00:00	0.11
2019-05-01 00:05	0.12
2019-05-01 00:10	0.15
2019-05-01 00:15	0.14
…	…
2019-05-02 00:05	0.36
2019-05-02 00:10	0.40
…	…

Group by day

Timestamp	Col1
2019-05-01 00:00	0.11
2019-05-01 00:05	0.12
2019-05-01 00:10	0.15
2019-05-01 00:15	0.14
…	…

Sliding window size=2

X

Y

Preprocess : Step 2

- Slicing using sliding window
  - Generate train/test data using sliding window

Scaled data (one cell)

Timestamp	Col1
2019-05-01 00:00	0.11
2019-05-01 00:05	0.12
2019-05-01 00:10	0.15
2019-05-01 00:15	0.14
…	…
2019-05-02 00:05	0.36
2019-05-02 00:10	0.40
…	…

Group by day

Timestamp	Col1
2019-05-01 00:00	0.11
2019-05-01 00:05	0.12
2019-05-01 00:10	0.15
2019-05-01 00:15	0.14
…	…

X

Y

Sliding window size=2

index	X1	X2	Y
0	0.11	0.12	0.15
1			
2			
3			
4			

Preprocess : Step 2

- Slicing using sliding window
  - Generate train/test data using sliding window

Scaled data (one cell)

Timestamp	Col1
2019-05-01 00:00	0.11
2019-05-01 00:05	0.12
2019-05-01 00:10	0.15
2019-05-01 00:15	0.14
…	…
2019-05-02 00:05	0.36
2019-05-02 00:10	0.40
…	…

Group by day

Timestamp	Col1
2019-05-01 00:00	0.11
2019-05-01 00:05	0.12
2019-05-01 00:10	0.15
2019-05-01 00:15	0.14
…	…

X

Y

Sliding window size=2

index	X1	X2	Y
0	0.11	0.12	0.15
1			
2			
3			
4			

Preprocess : Step 2

- Slicing using sliding window
  - Generate train/test data using sliding window

Scaled data (one cell)

Timestamp	Col1
2019-05-01 00:00	0.11
2019-05-01 00:05	0.12
2019-05-01 00:10	0.15
2019-05-01 00:15	0.14
…	…
2019-05-02 00:05	0.36
2019-05-02 00:10	0.40
…	…

Group by day

Timestamp	Col1
2019-05-01 00:00	0.11
2019-05-01 00:05	0.12
2019-05-01 00:10	0.15
2019-05-01 00:15	0.14
…	…

X

Y

Sliding window size=2

index	X1	X2	Y
0	0.11	0.12	0.15
1	0.12	0.15	0.14
2	…	…	…
3	…	…	…
4	…	…	…

- The above is an example, and the actual Window Size is 10. (5 min x 10 = 50 min)
- We do the same for memories. (We will do the above for all dates.)
- The number of memory used (msteps) is 7 because it trains the latest seven days.