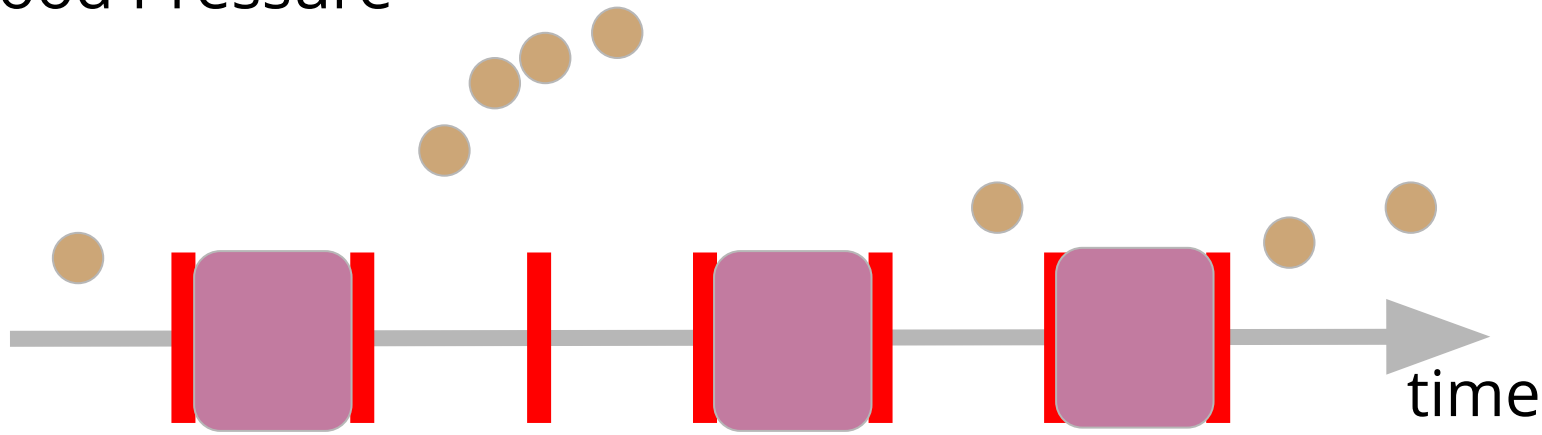
Classification of Irregularly Sampled Clinical Time-Series Data with Convolutional Neural Networks

Group: White Box
Michael Wieck-Sosa, Diego Moreno, Yujuan Fu

Irregularly sampled health data

Blood Pressure



Challenge: data are no missing at random, how to handle irregularities in sampling

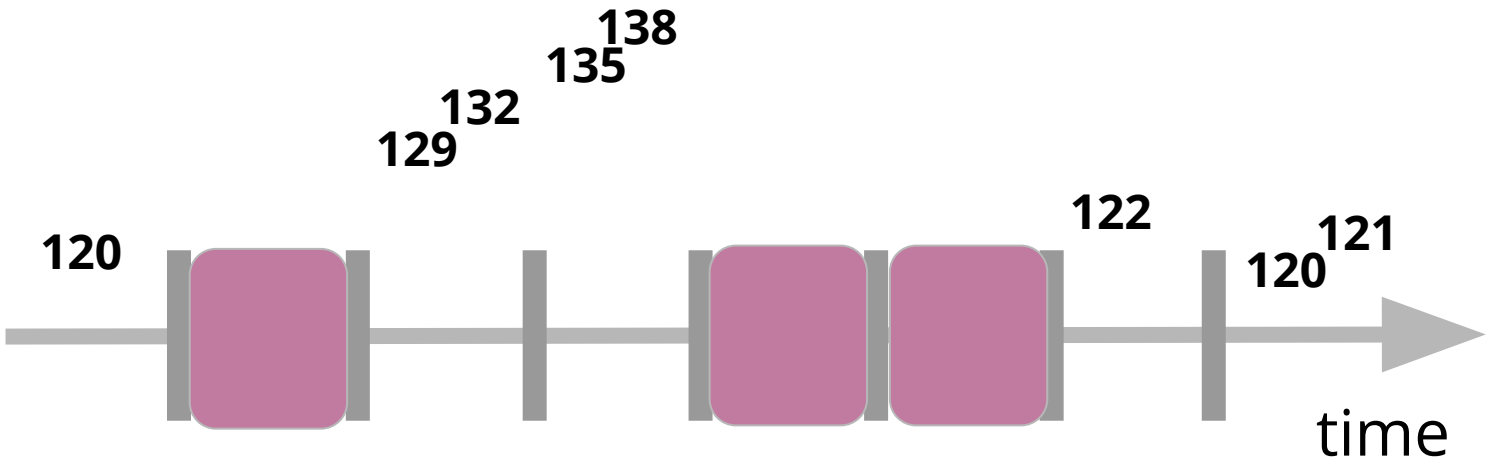
The Proposed Model

- **Hypothesis:** explicitly accounting for missinginess will outperform a model that ignores missingness
- **Idea**: use additional input to CNN that encodes missingness

- **Approach**: compare a baseline CNN to one that explicitly accounts for missingness

Irregularities in the Data

Blood Pressure



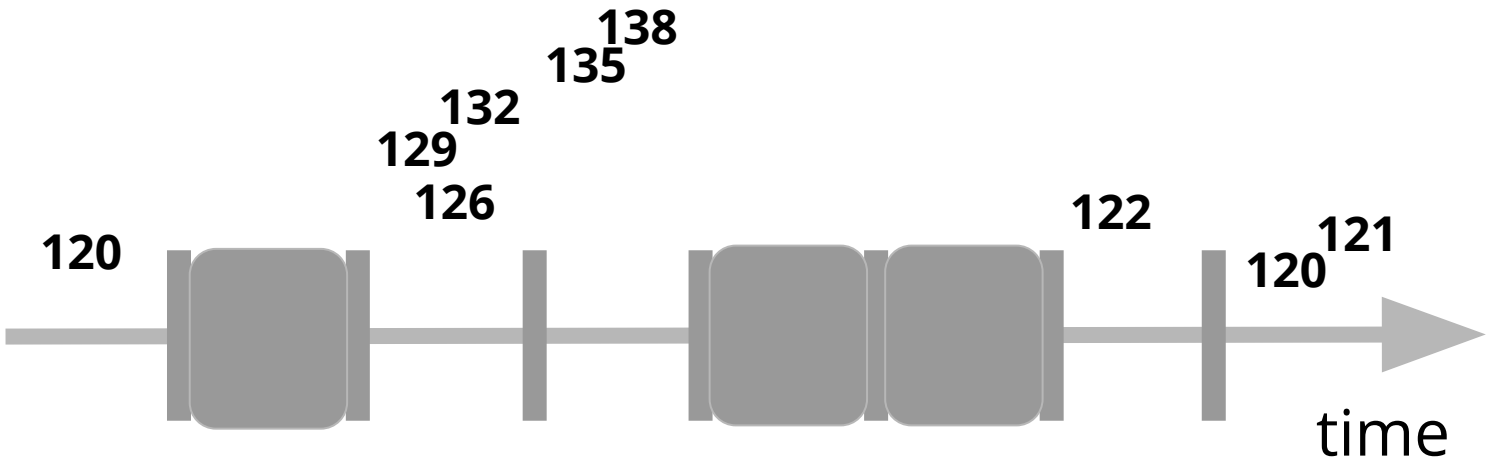
Time interval	1	2	3	4	5	6	7	8
BP	120	nan	130.5	136.5	nan	nan	122	121
Missing indicator	0	1	0	0	1	1	0	0

Relative Intensity of Measurements

Blood Pressure

Total count: 1+3+2+1+2=9

Average count: 9/8=1.125

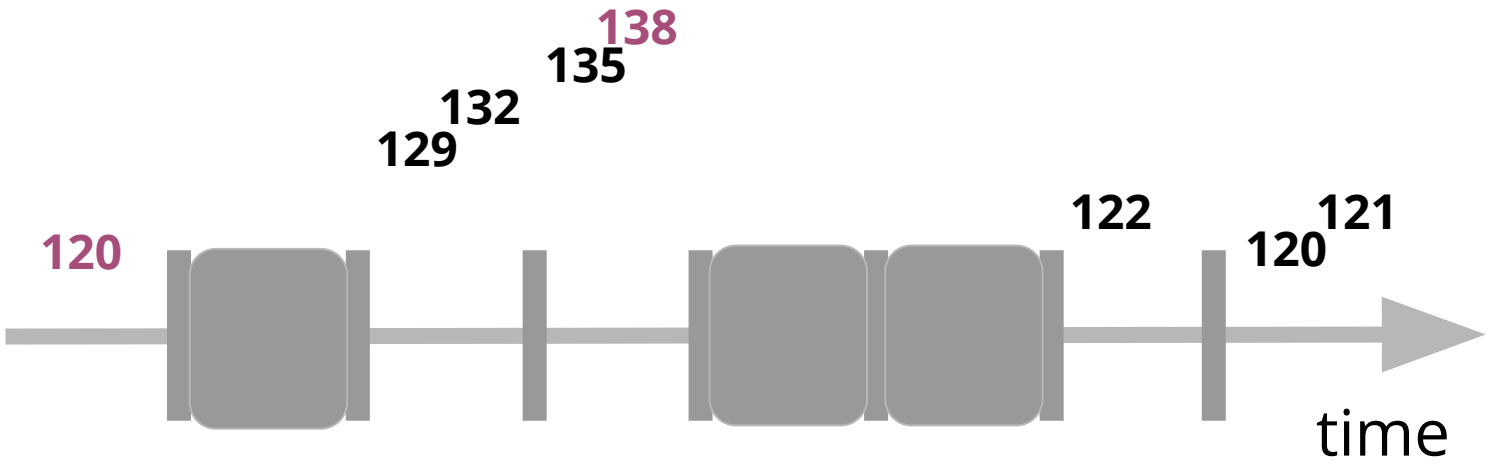


Time interval	1	2	3	4	5	6	7	8
Count	1	0	3	2	0	0	1	2
Ci/C_avg								

***just divide count in time interval by the average number of measurements**

Carry-forward Imputation

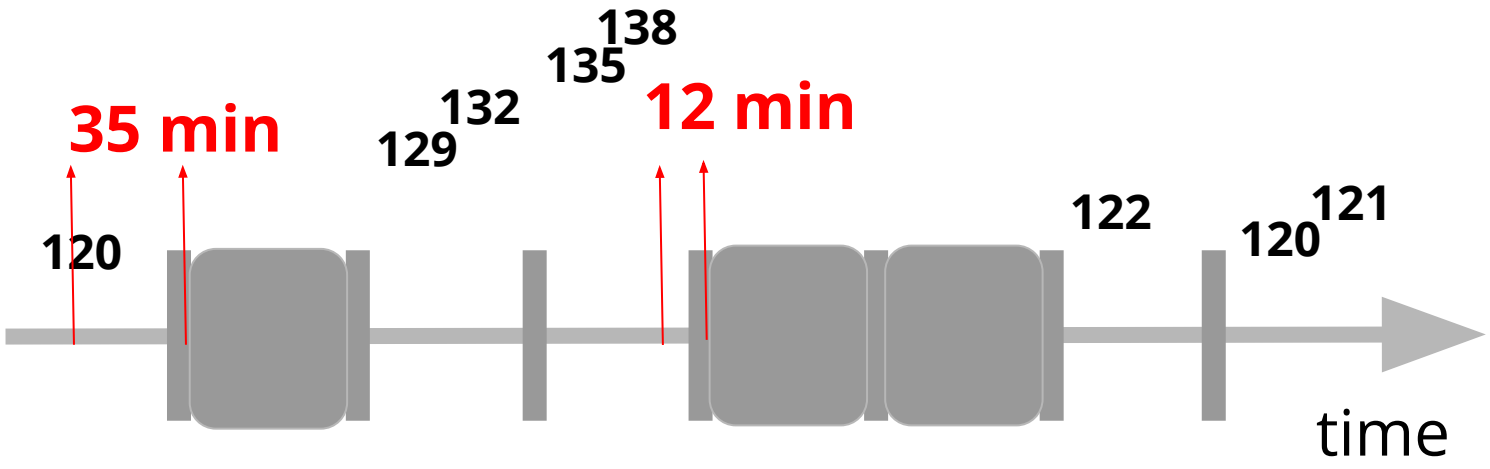
Blood Pressure



Time interval	1	2	3	4	5	6	7	8
Mean BP	120		130.5	136.5			122	120.5

Unreliability Metric

Blood Pressure

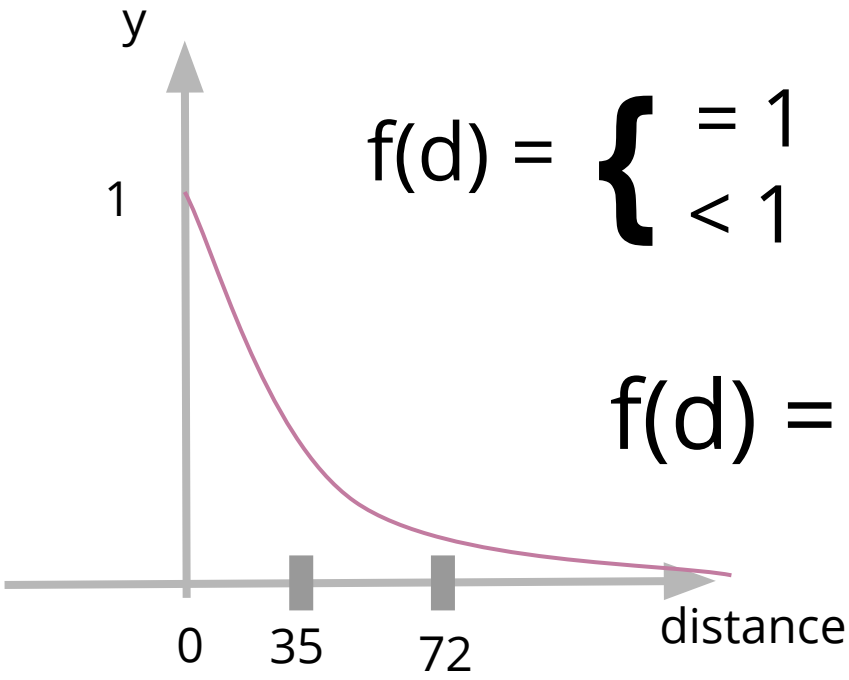


Time interval	1	2	3	4	5	6	7	8
Distance	35		8	12			36	43

*distance from the last measurement

Unreliability Metric

Time interval	1	2	3	4	5	6	7	8
Distance	35	95	8	12	72	132	36	43
Missing indicator	0	1	0	0	1	1	0	0



$$f(d) = \begin{cases} = 1 & \text{if not missing} \\ < 1 & \text{if missing} \end{cases}$$

$$f(d) = e^{-distance * missing}$$

Convolutional Layer

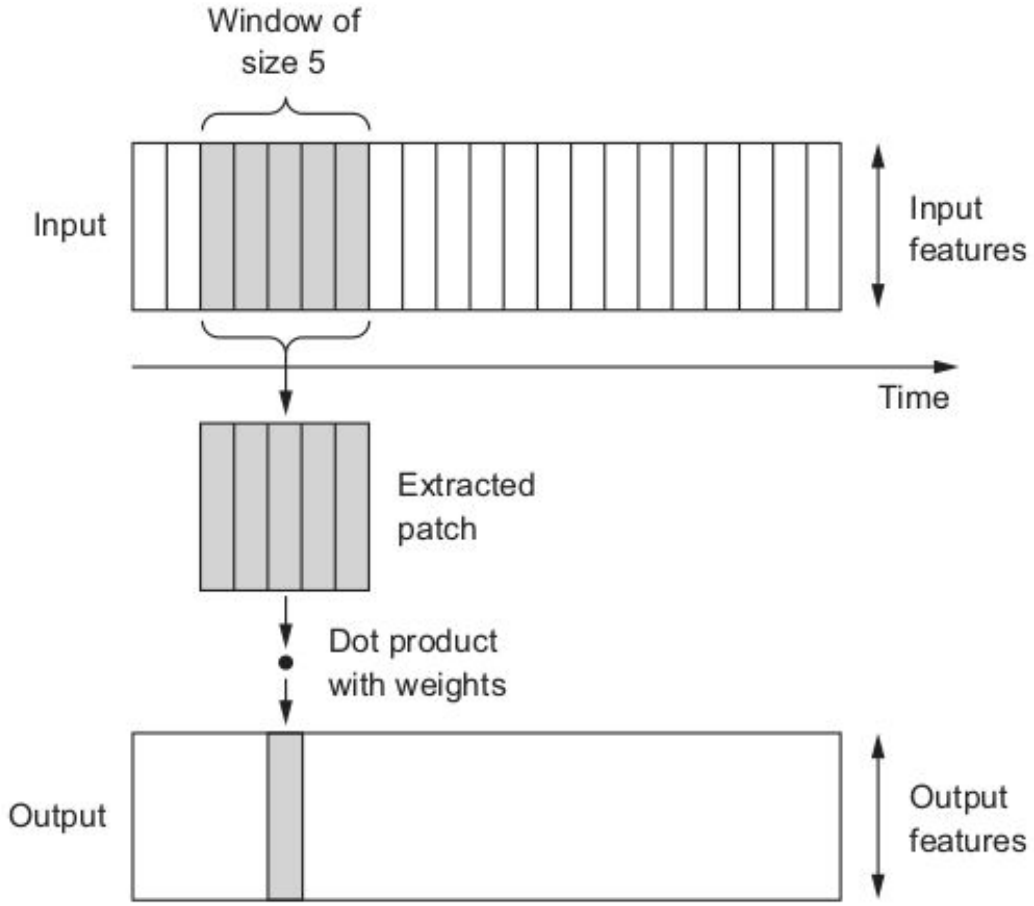
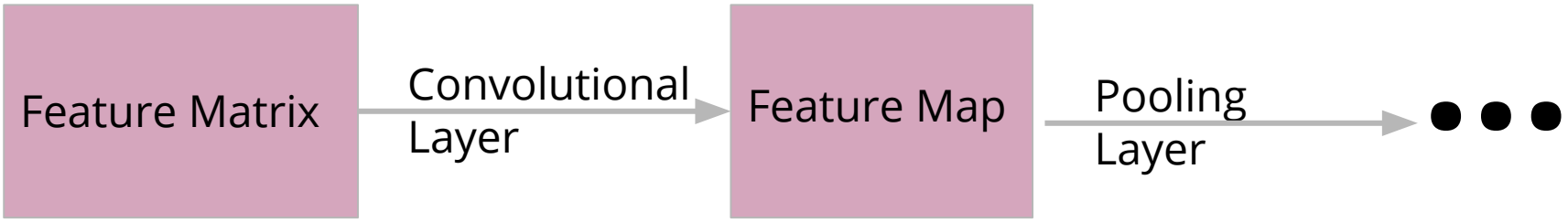
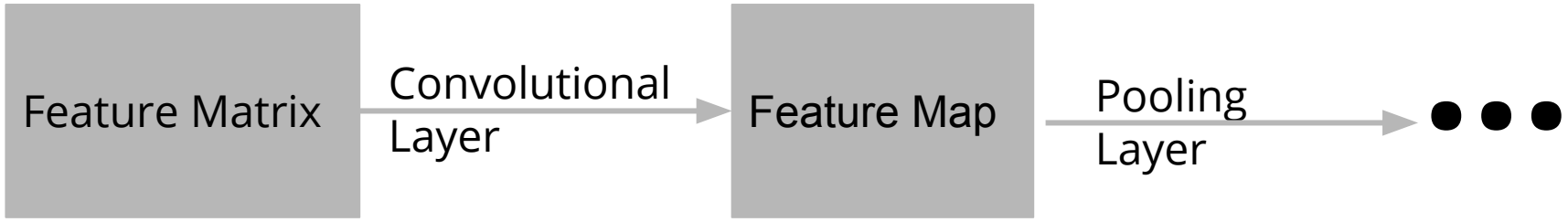


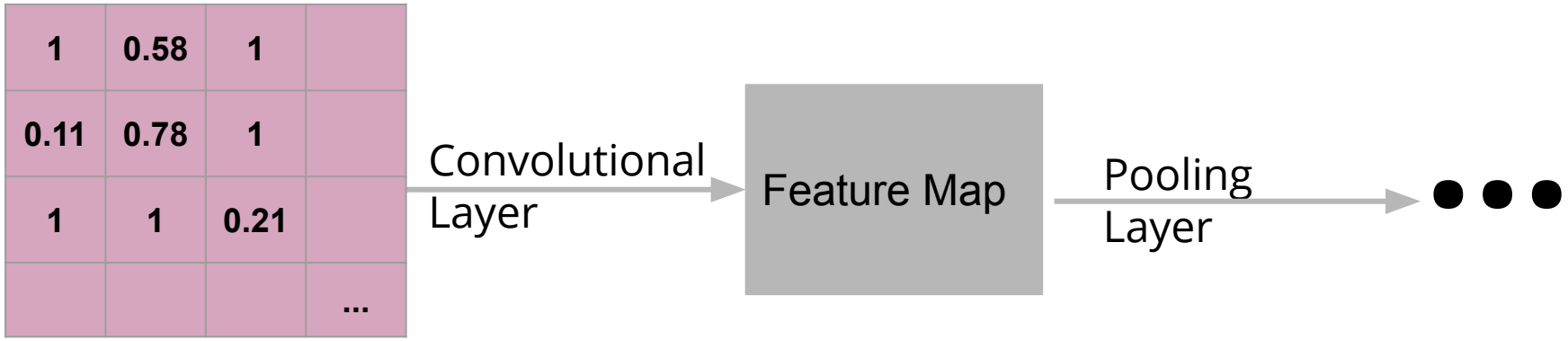
Figure 6.26 How 1D convolution works: each output timestep is obtained from a temporal patch in the input sequence.

Architecture of a Convolutional Neural Network

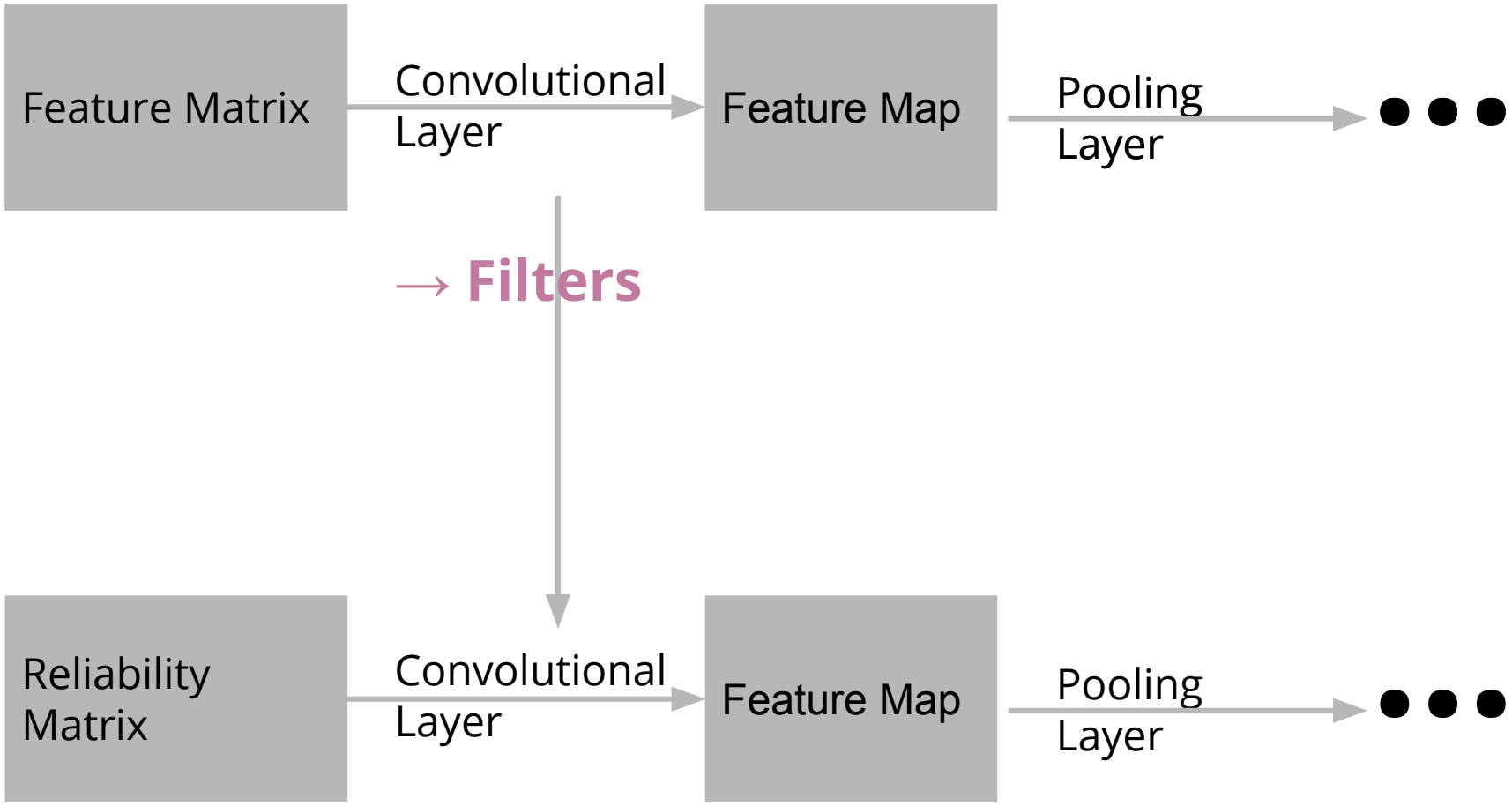


Convolutional Layer

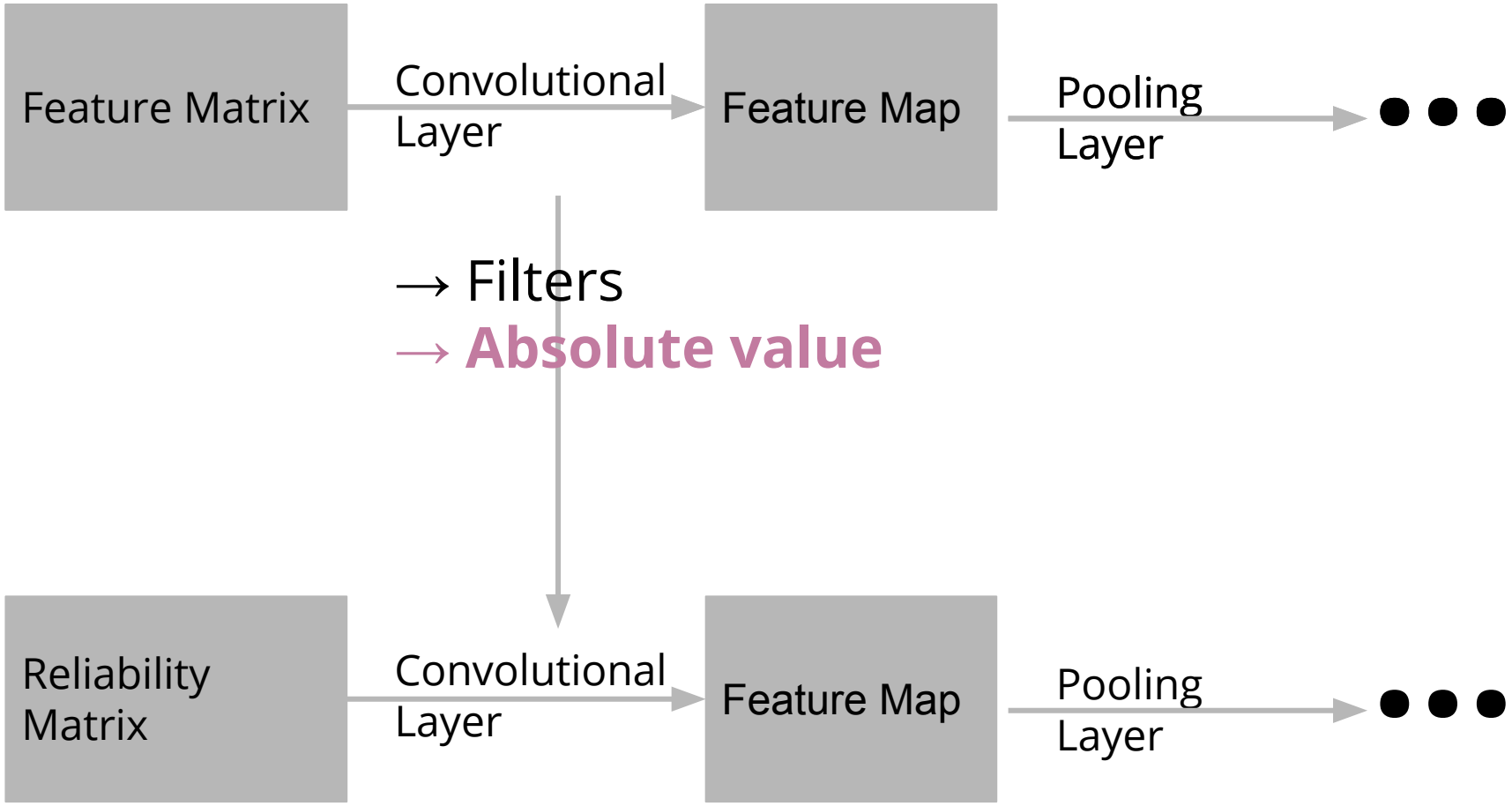




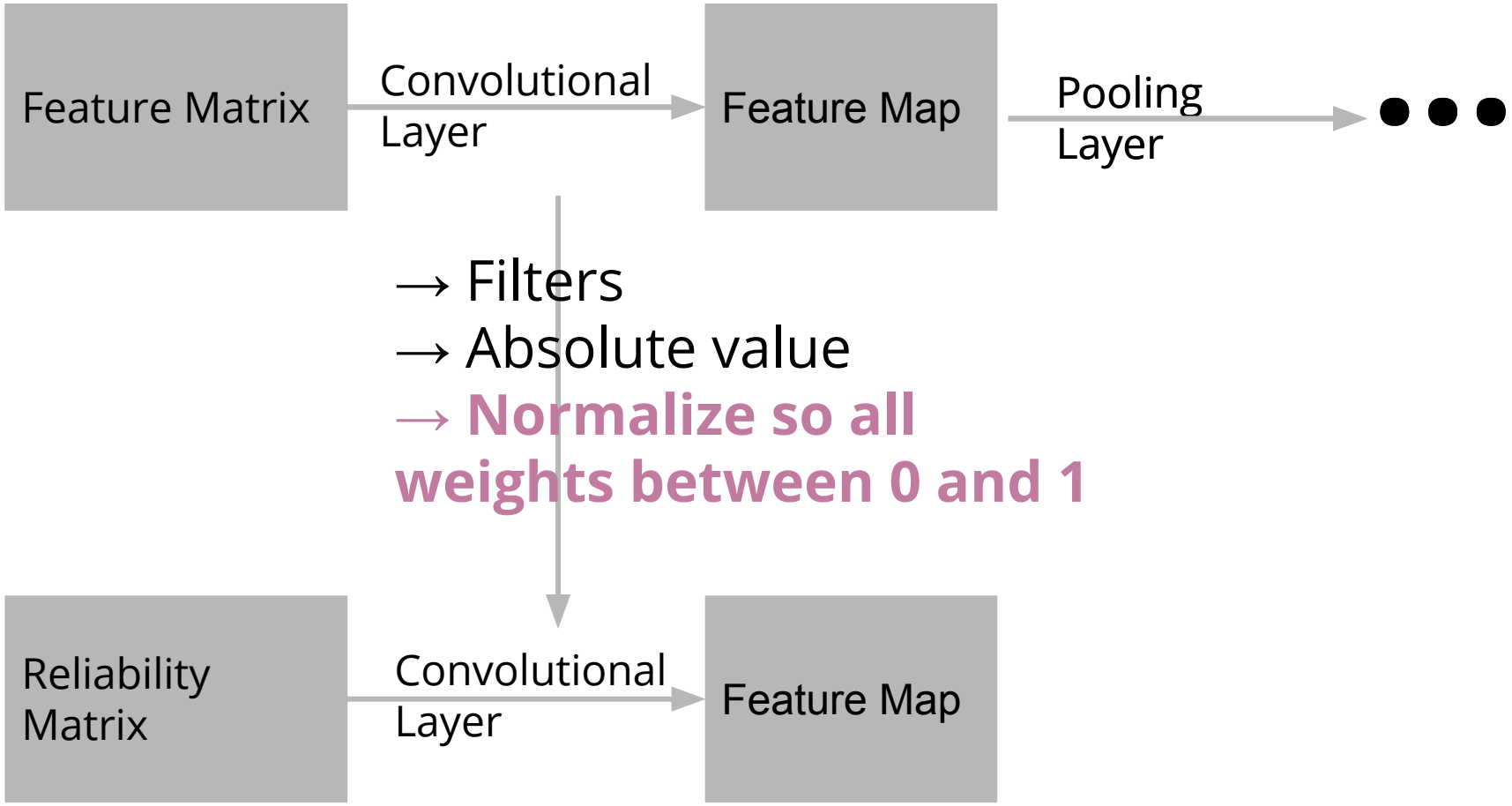
Our Proposed Model



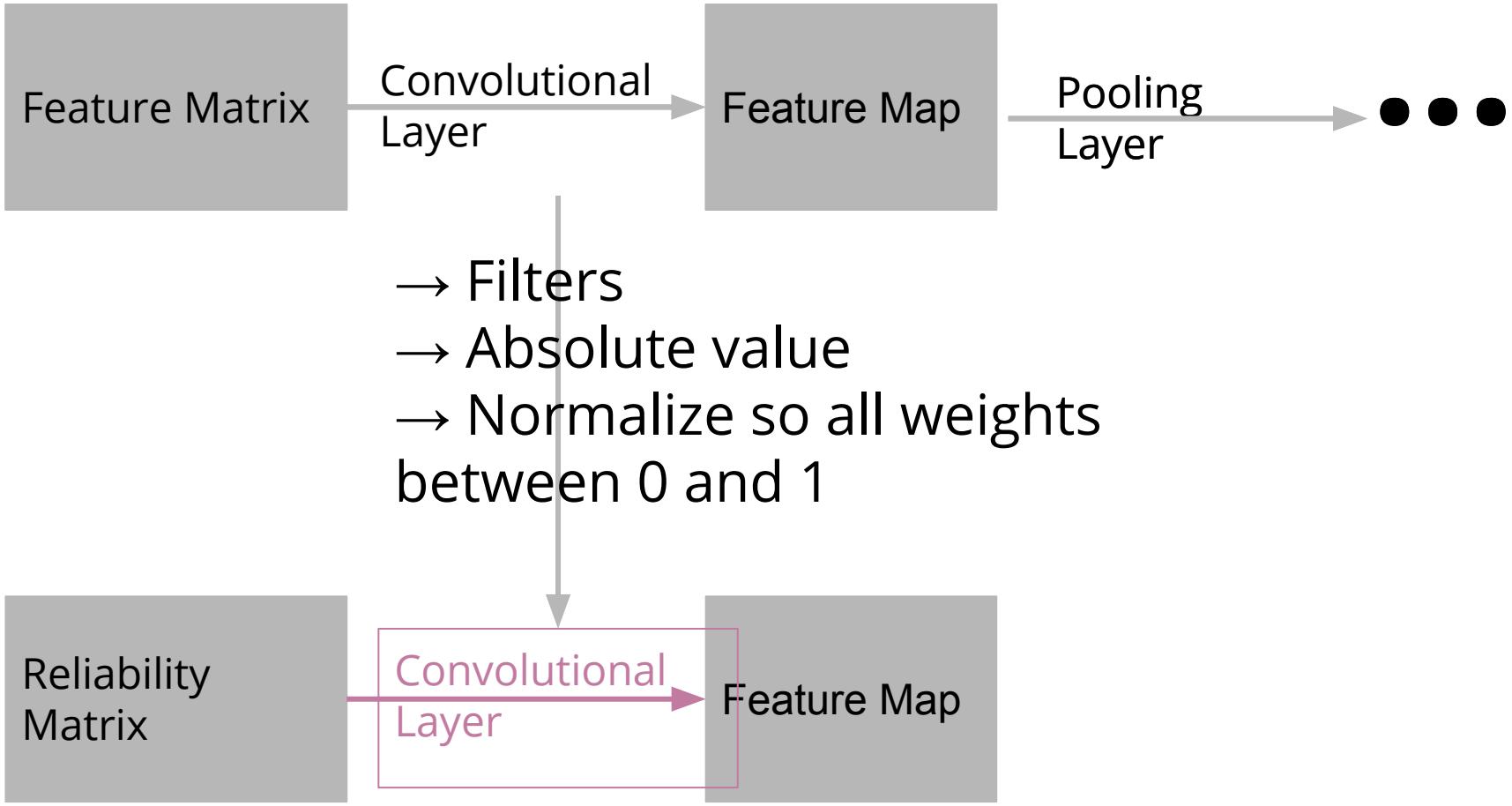
Our Proposed Model



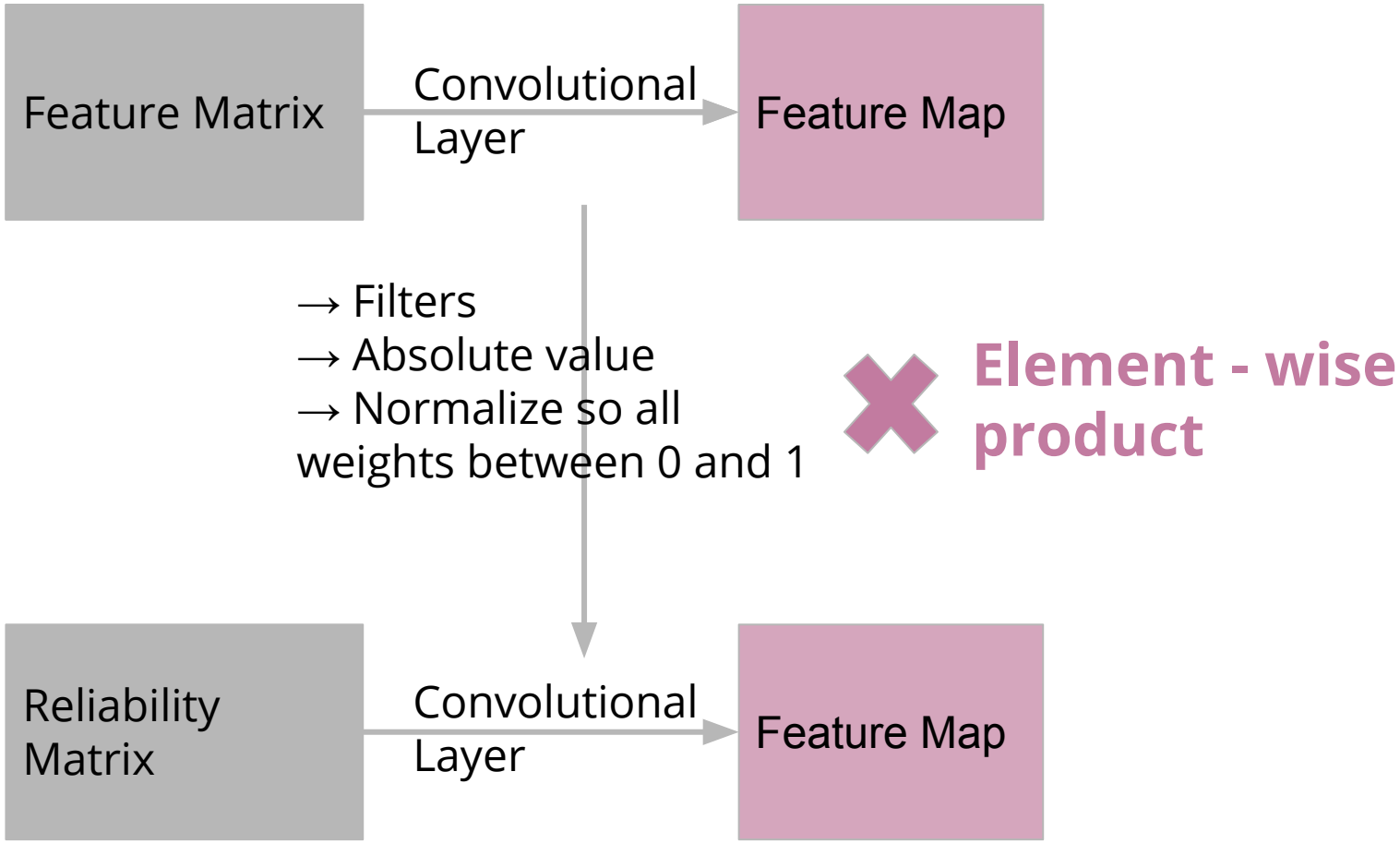
Our Proposed Model



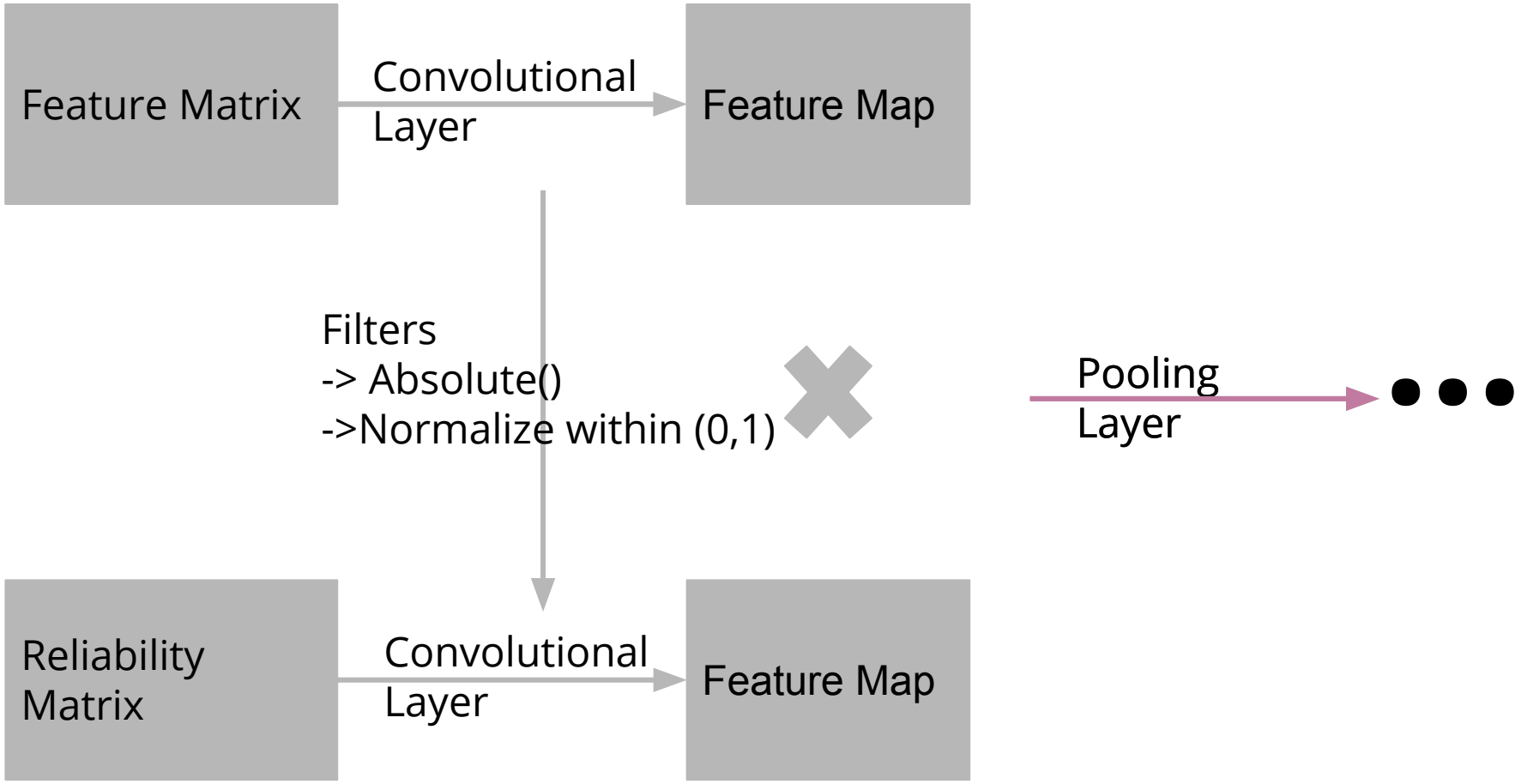
Our Proposed Model



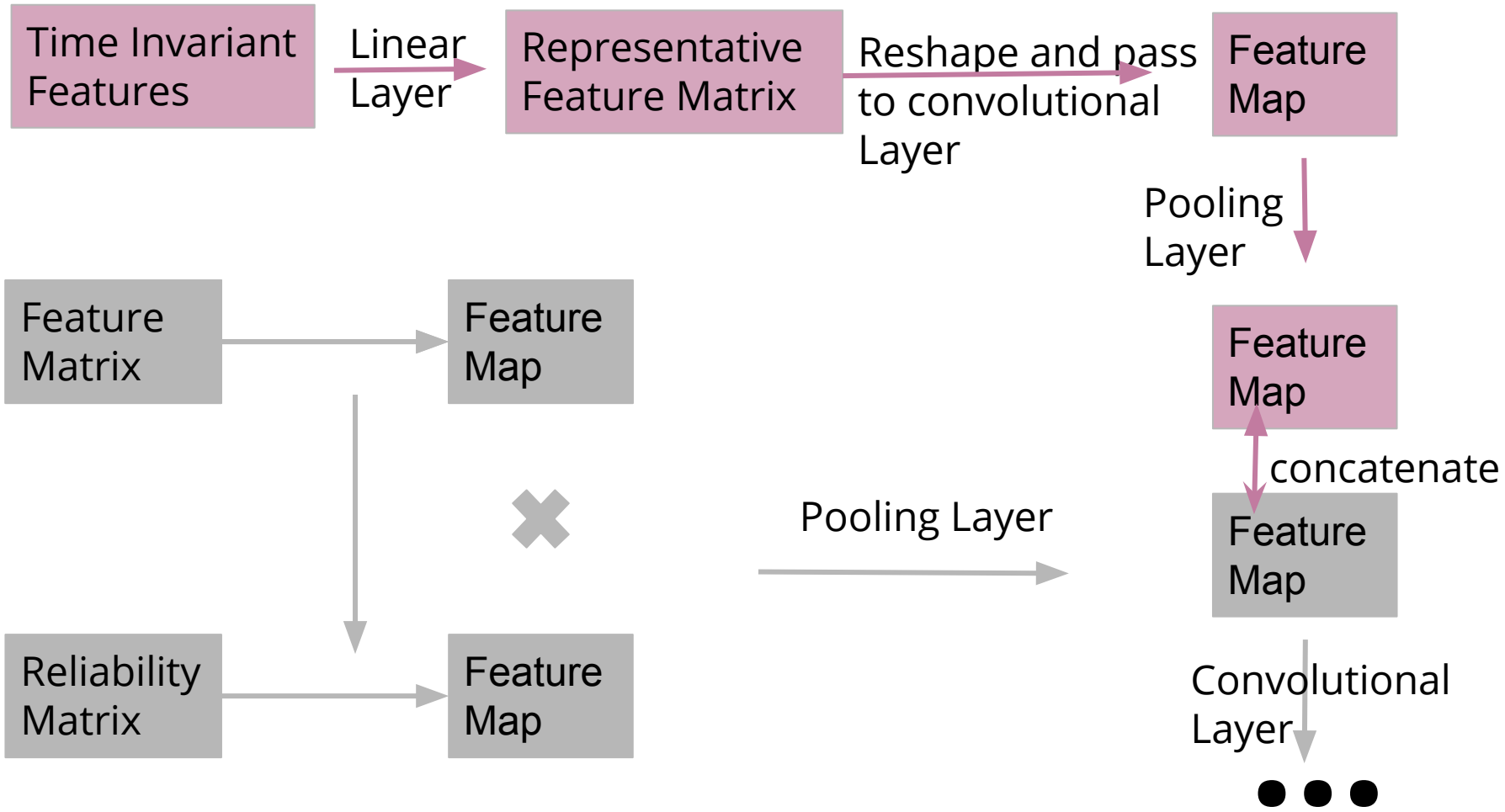
Our Proposed Model



Our Proposed Model



Our Proposed Model



Model Evaluation

Model	AUROC Confidence Interval	Loss Confidence Interval
Baseline CNN	[0.823, 0.847]	[0.312, 0.327]
Proposed CNN	[0.835, 0.858]	[0.293, 0.367]

Conclusion

- The proposed model modestly improves prediction performance

- **Clinical Relevance:** Approach can be applied to irregularly sampled clinical time-series data

- **Technical Relevance:** Approach can be generalized to irregularly sampled time-series data in different domains, provided that the irregularities in the sampling rate are non-random and non-uniform across features

Q&A

Thank You.