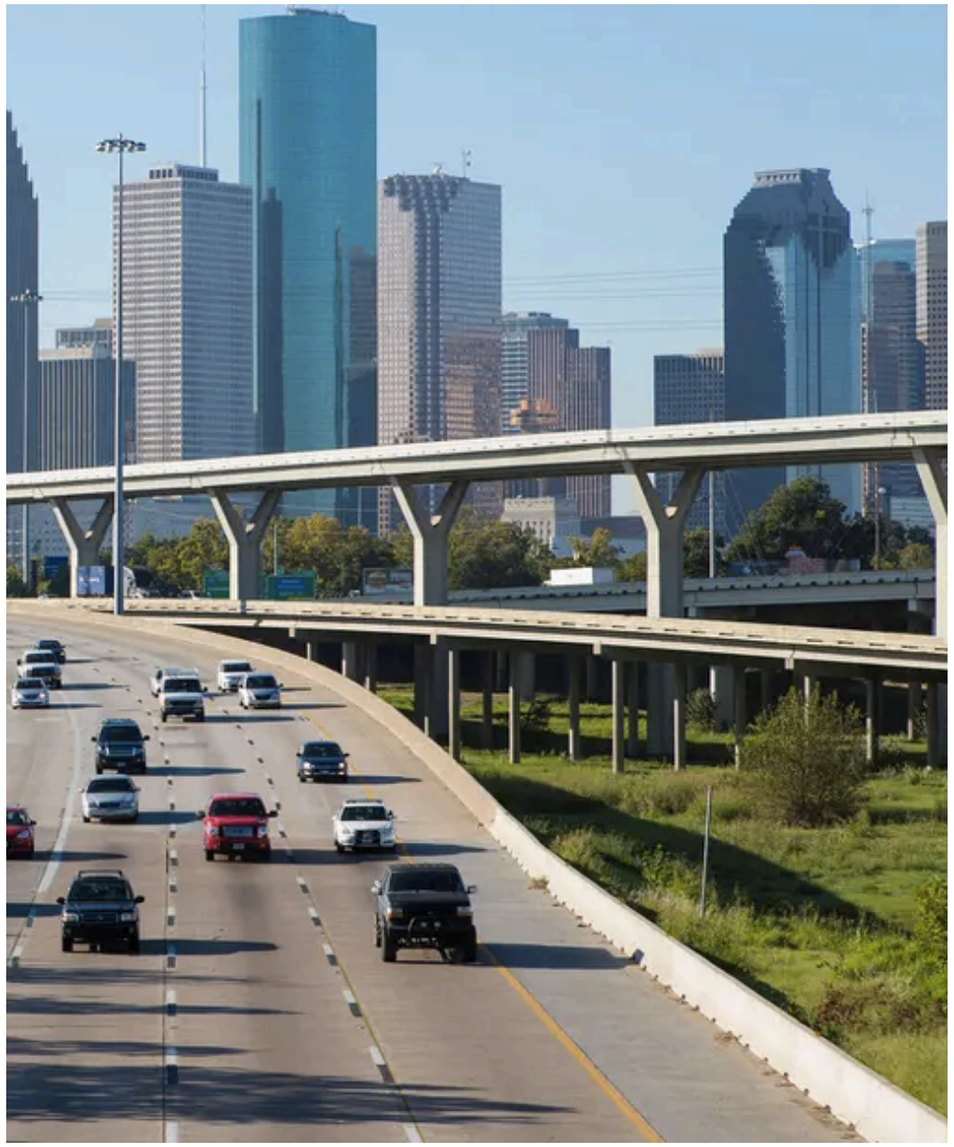
Multi-output neural processes

Zinzan Gurney

Supervised by Prof. Richard Turner

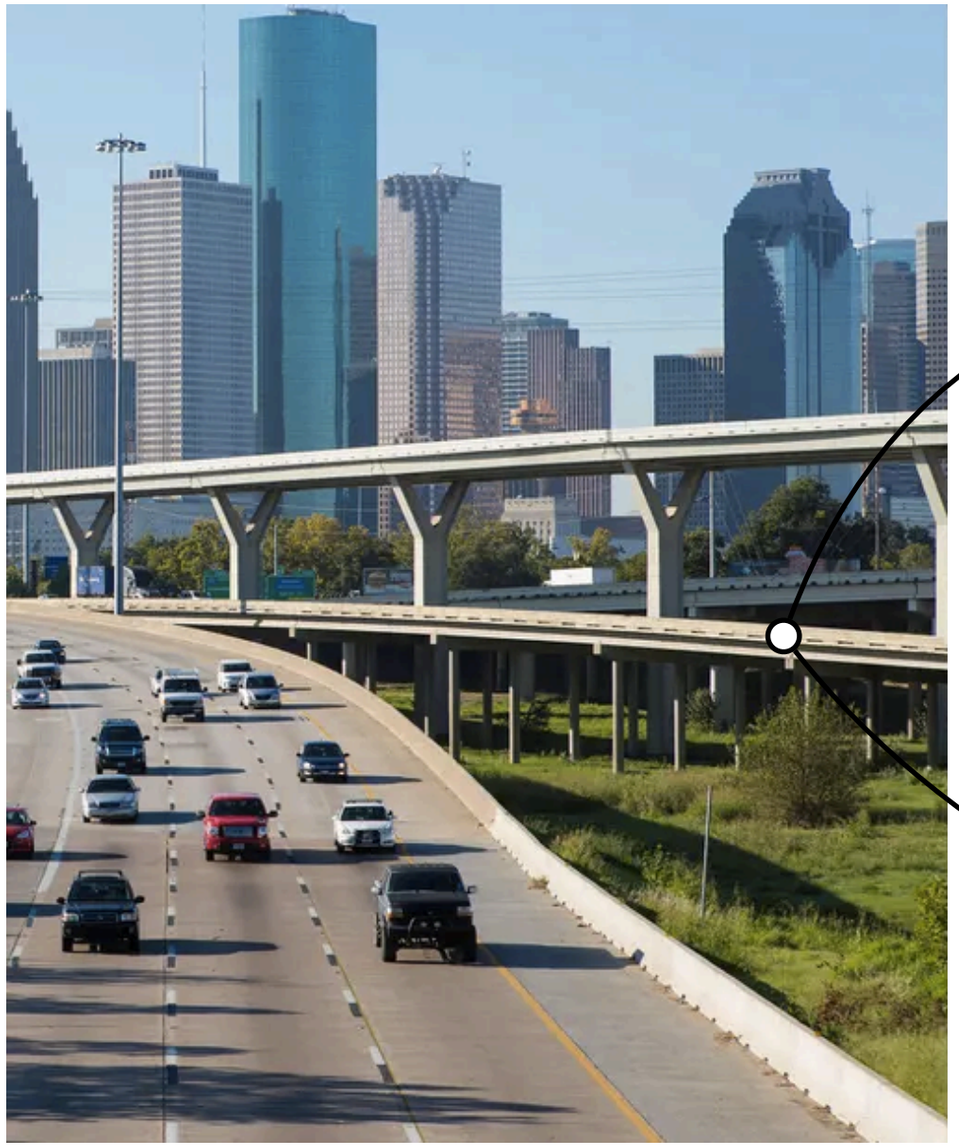
24 November 2021

What problem are we addressing?

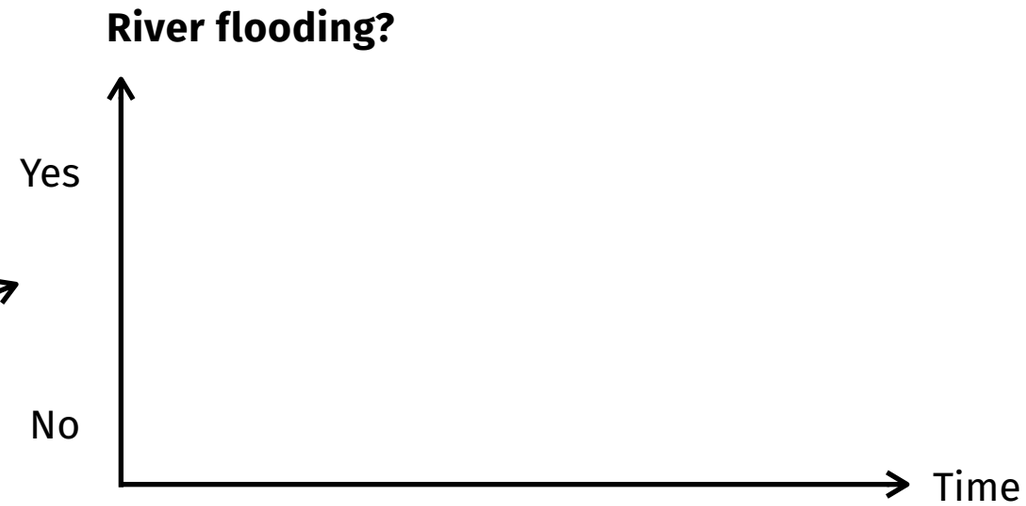


[1]

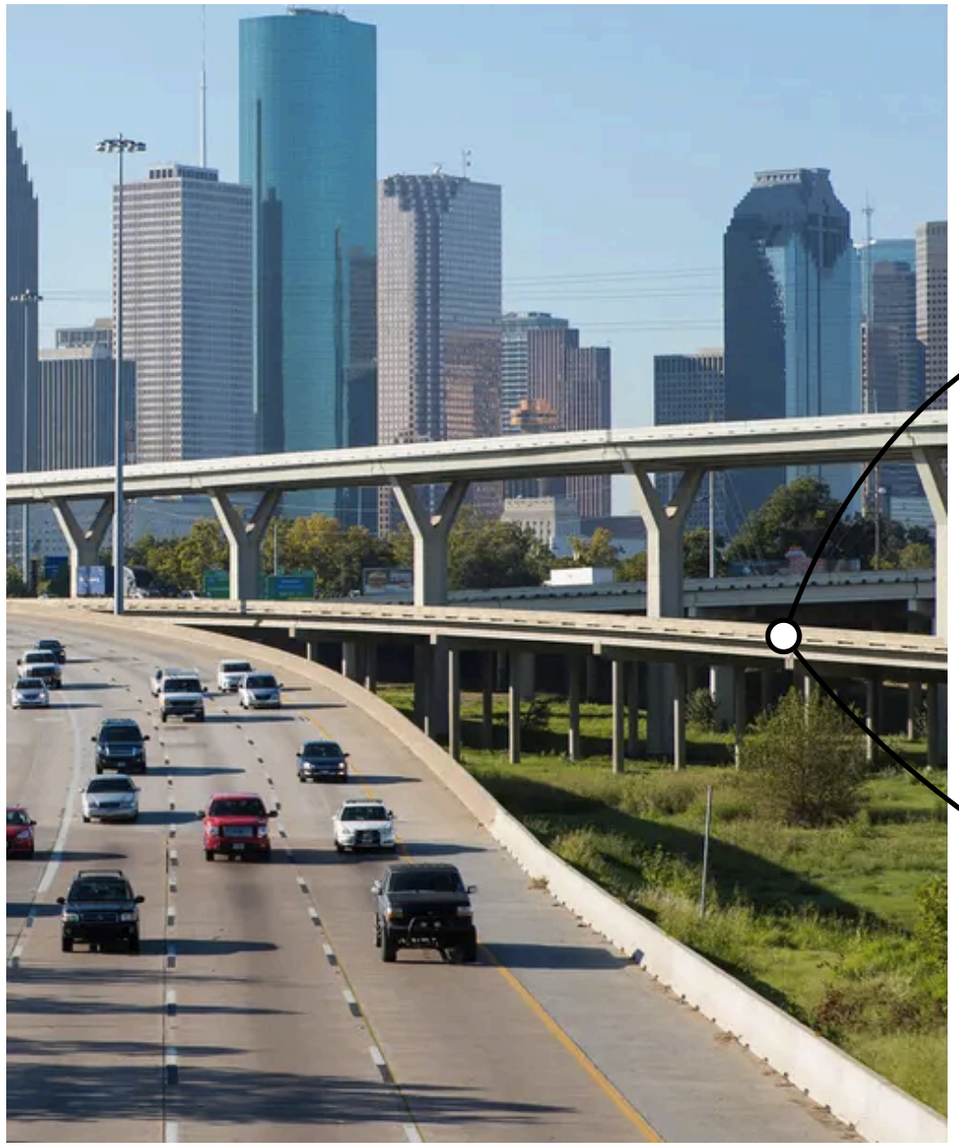
What problem are we addressing?



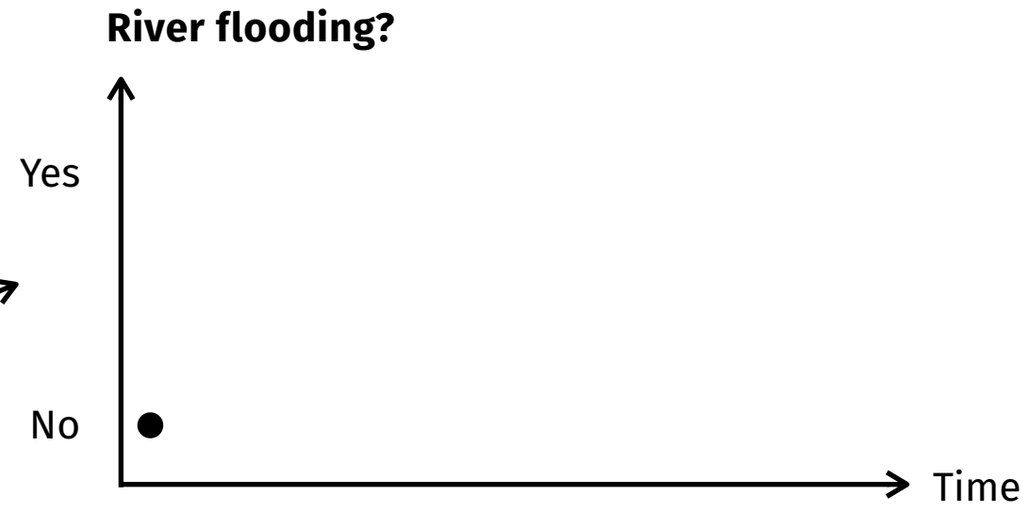
[1]



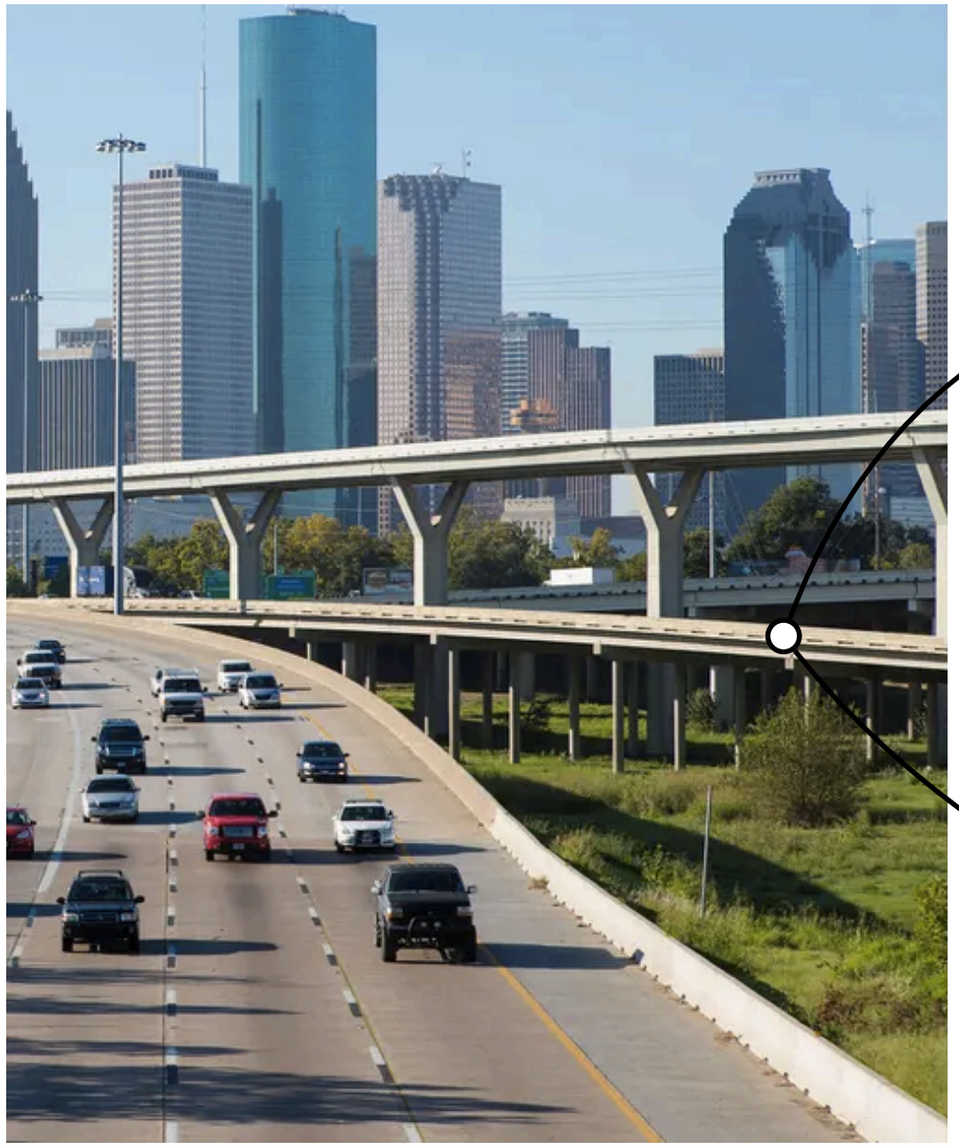
What problem are we addressing?



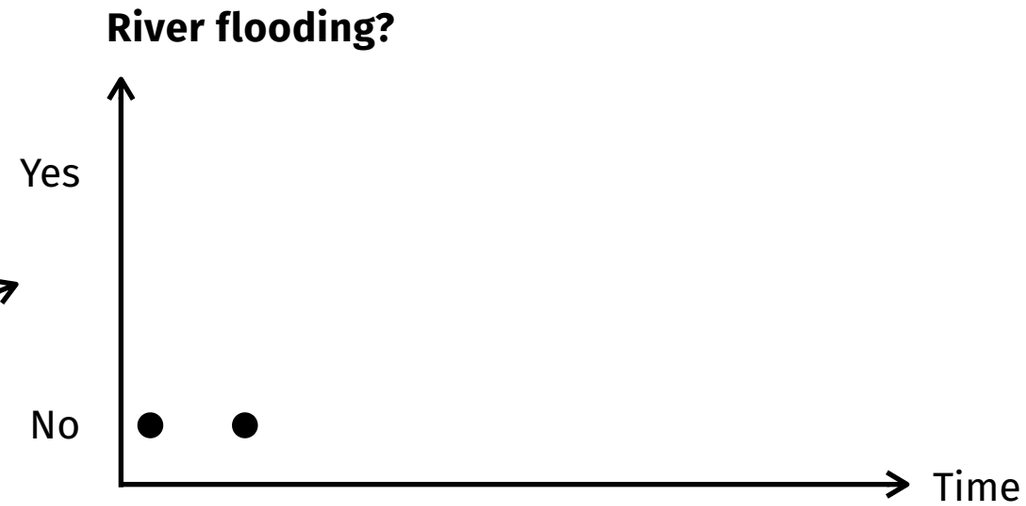
[1]



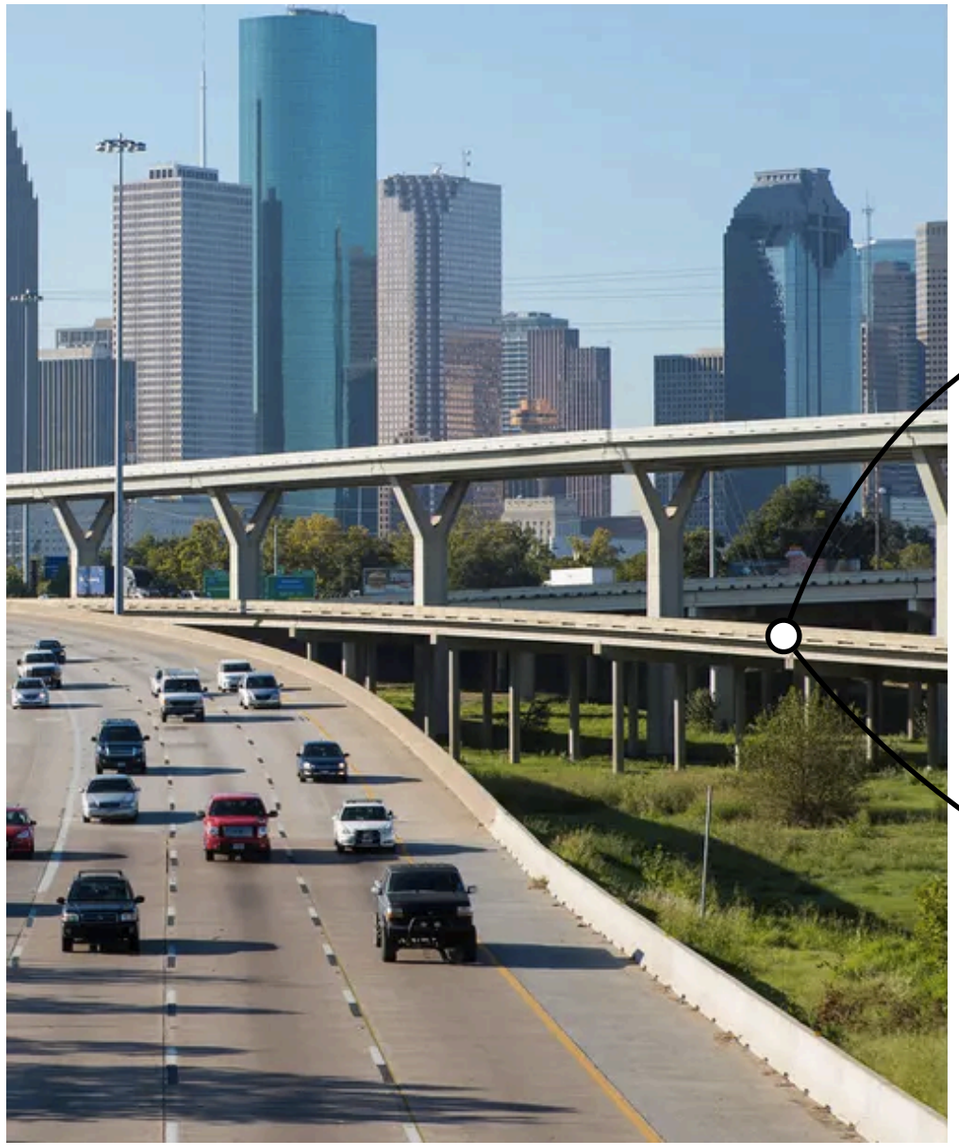
What problem are we addressing?



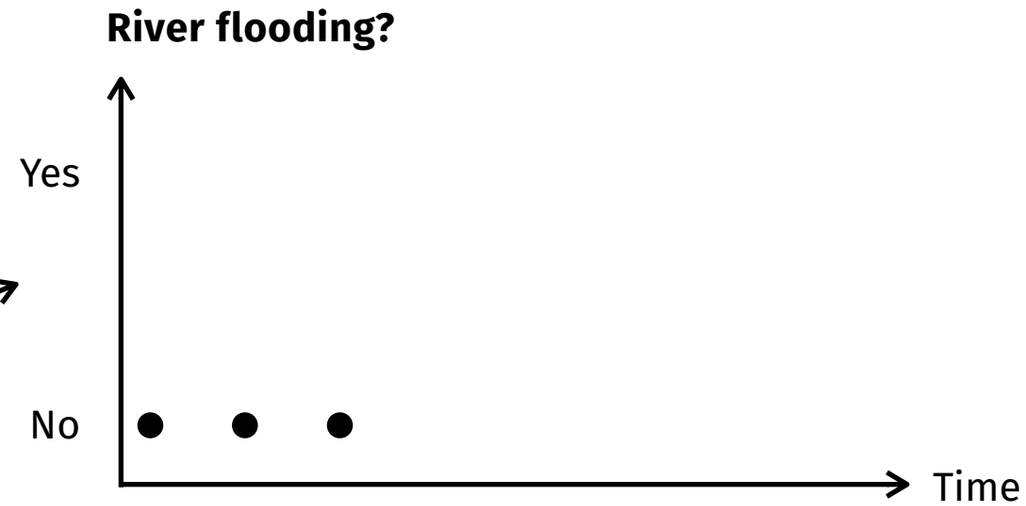
[1]

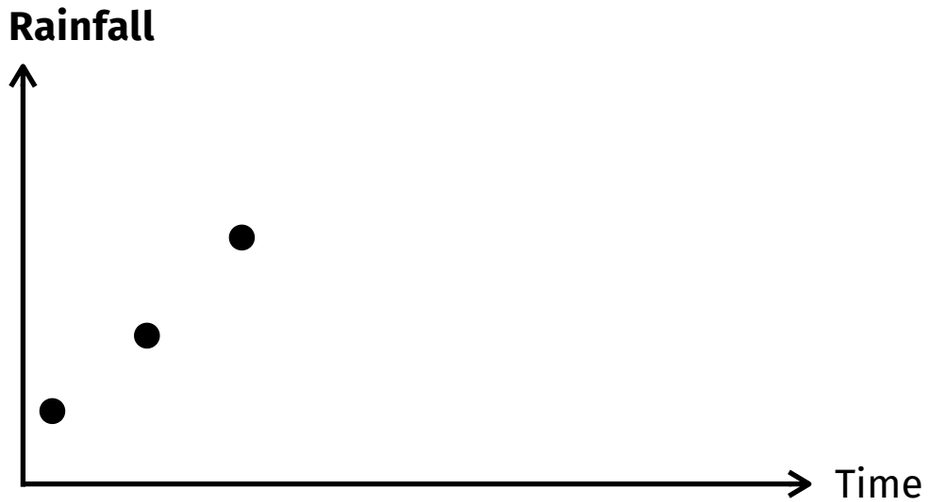


What problem are we addressing?

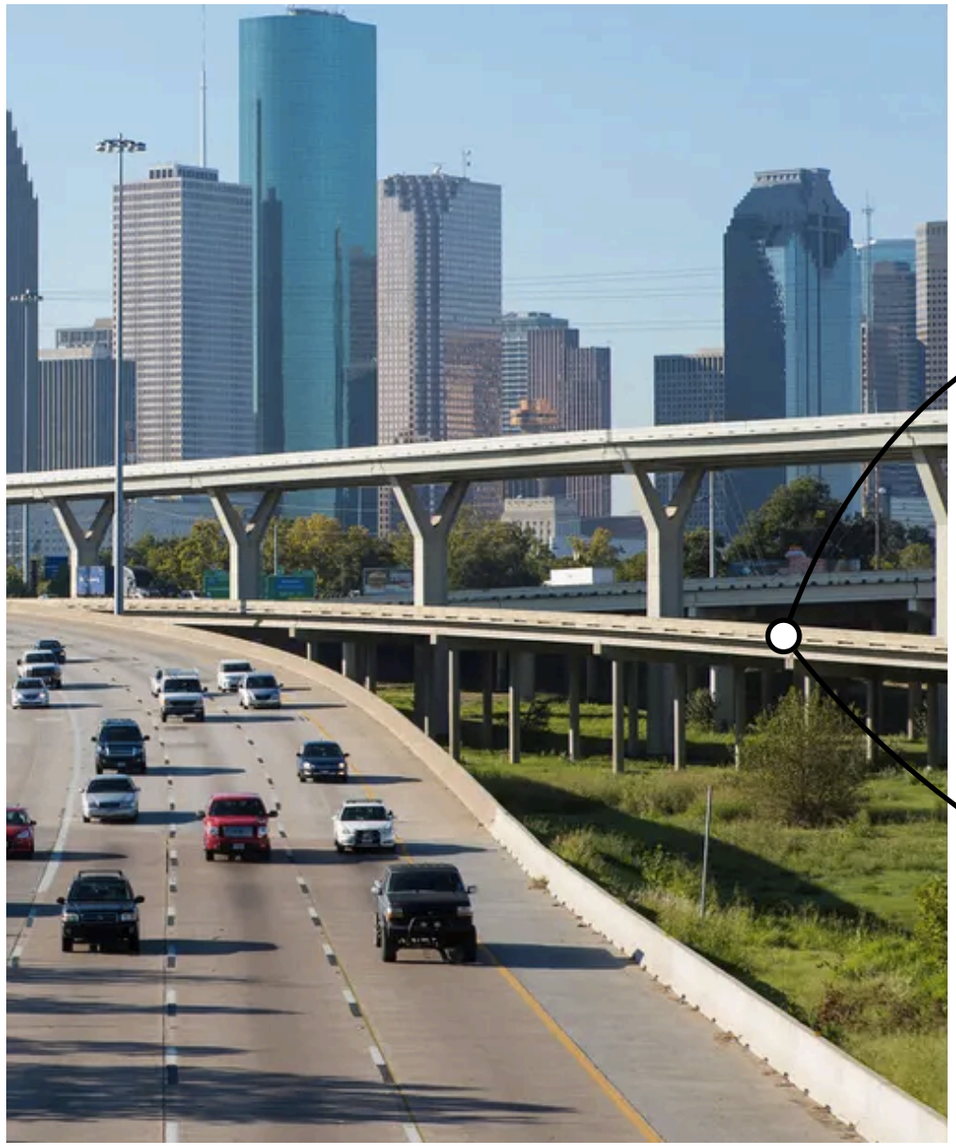


[1]

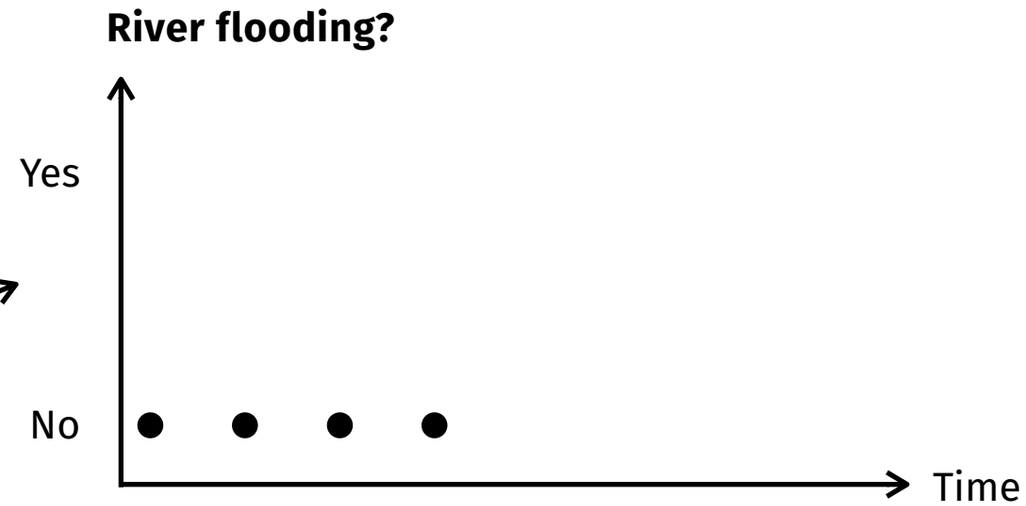


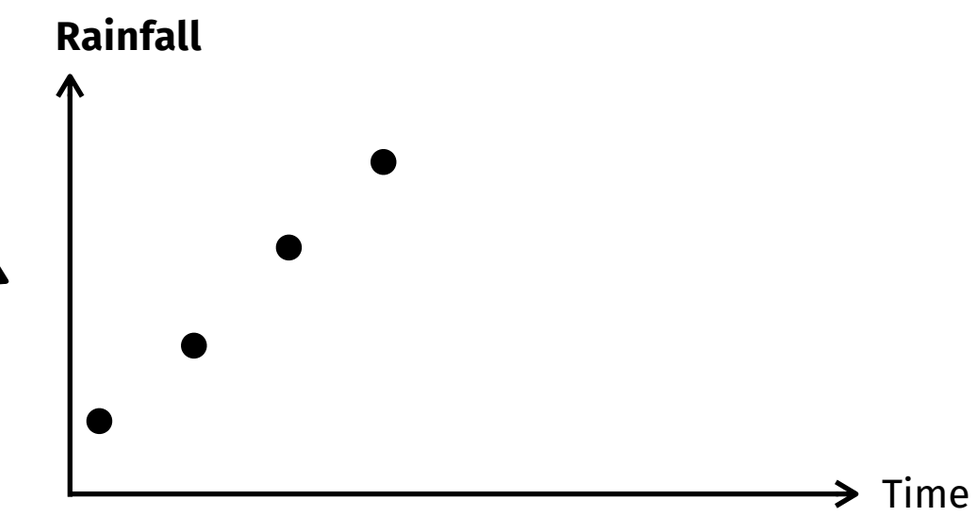


What problem are we addressing?

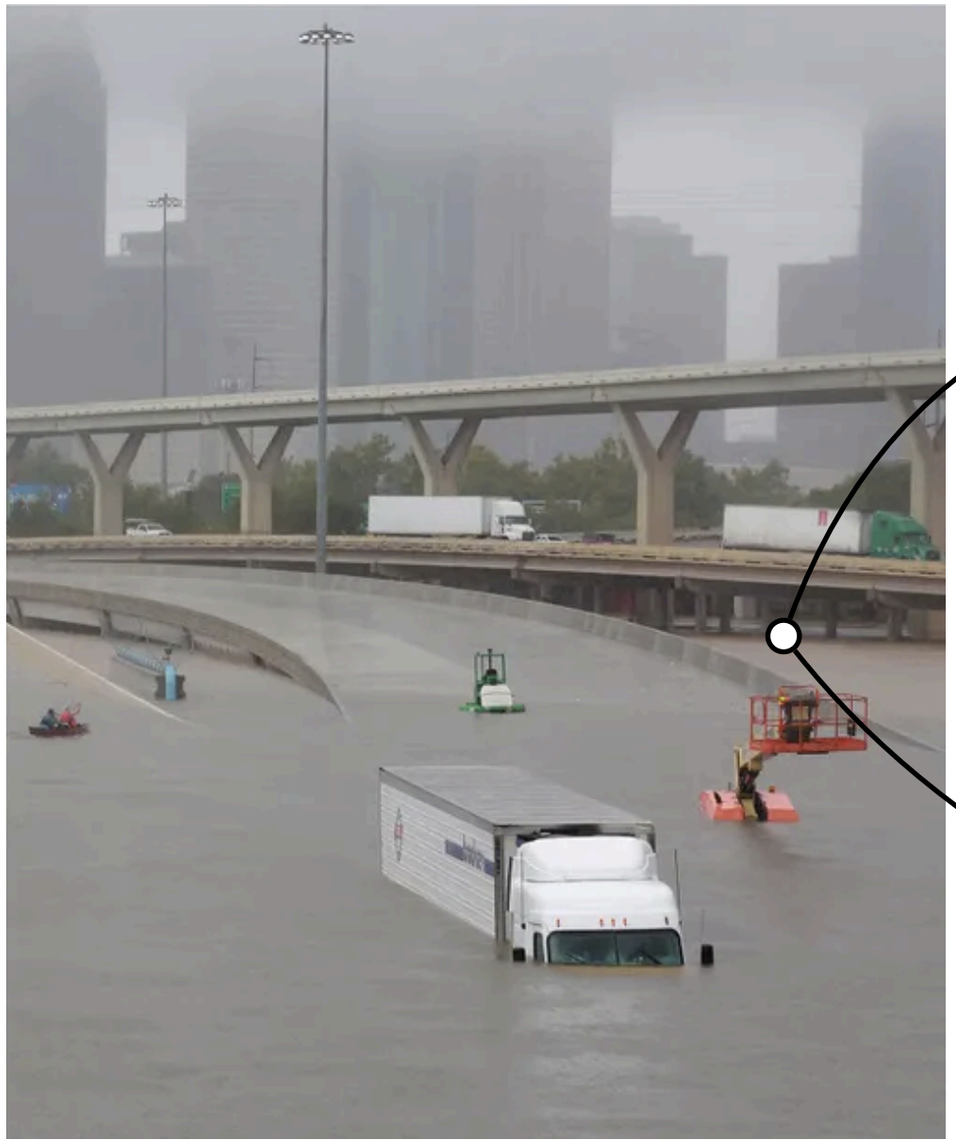


[1]

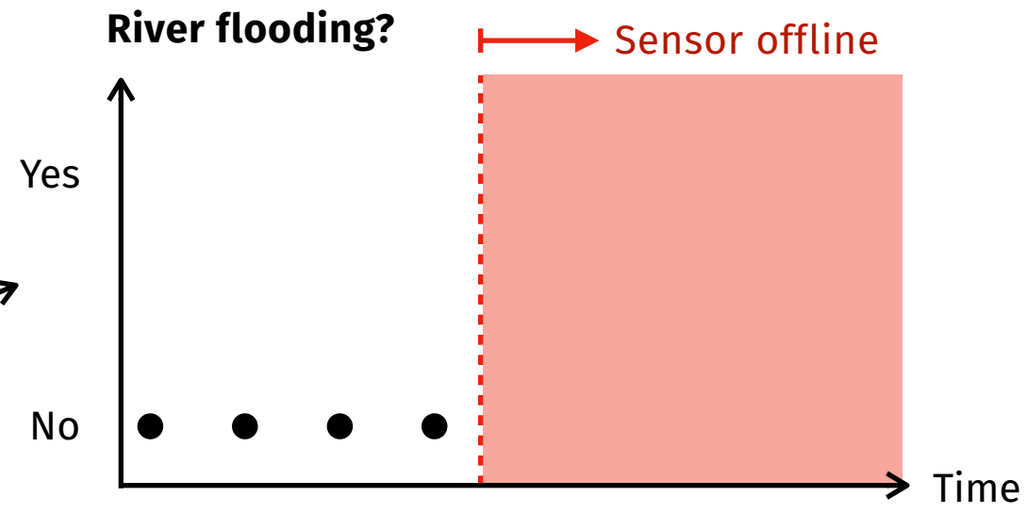


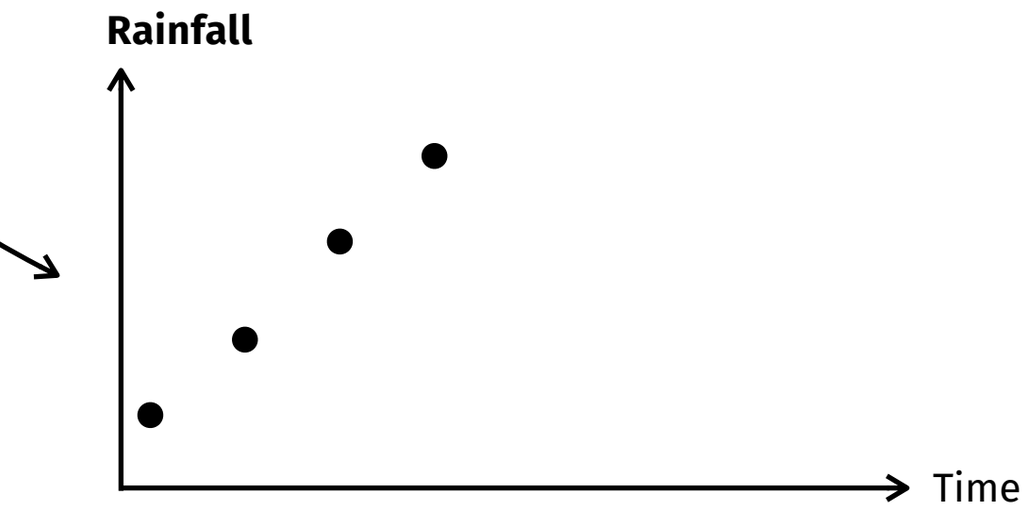


What problem are we addressing?

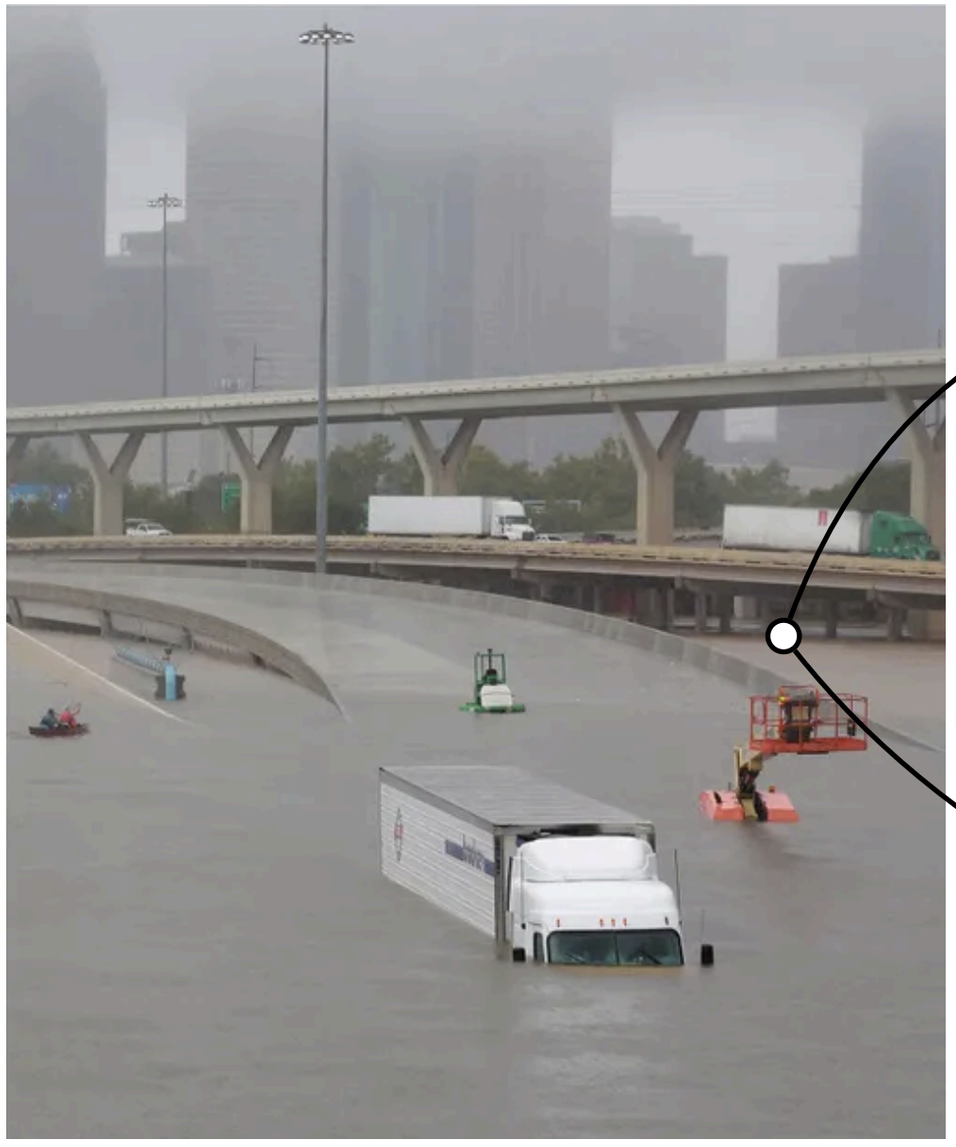


[1]

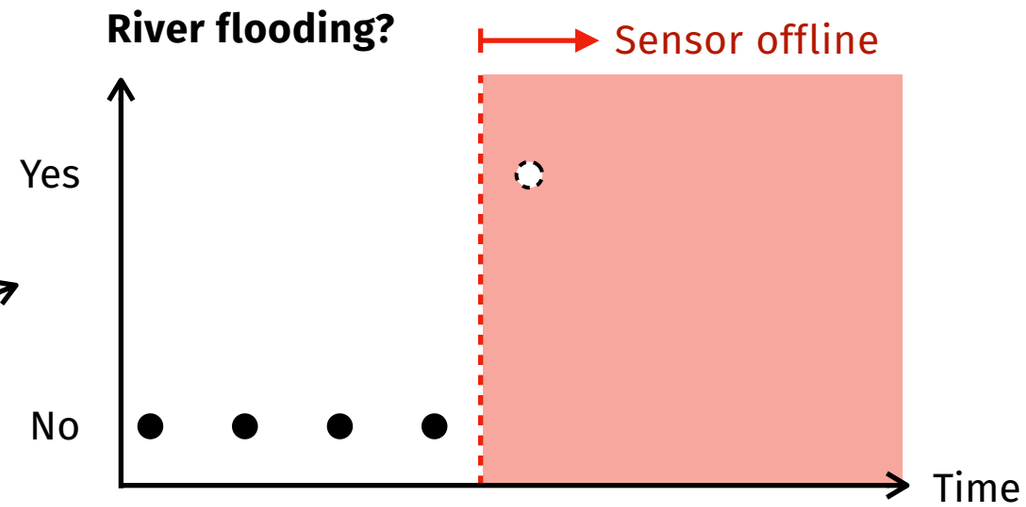


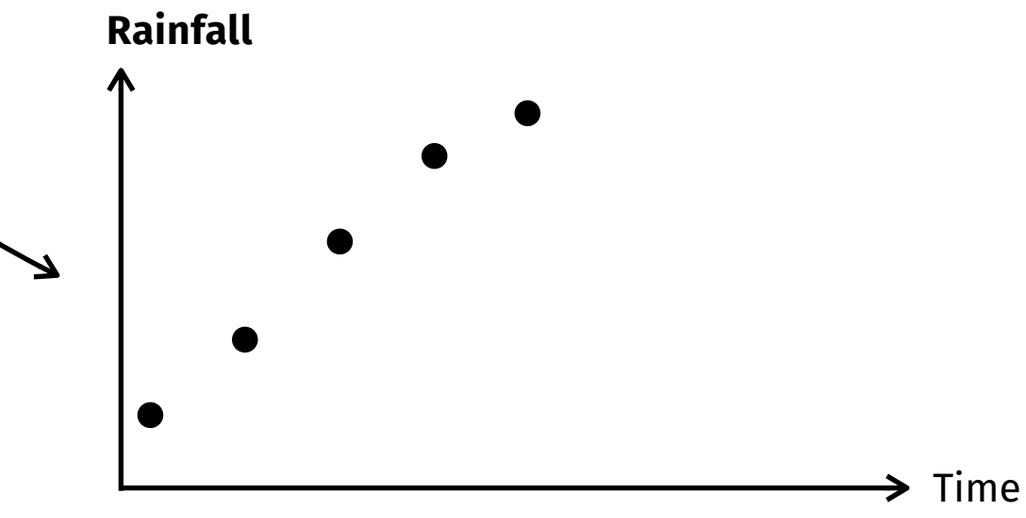


What problem are we addressing?

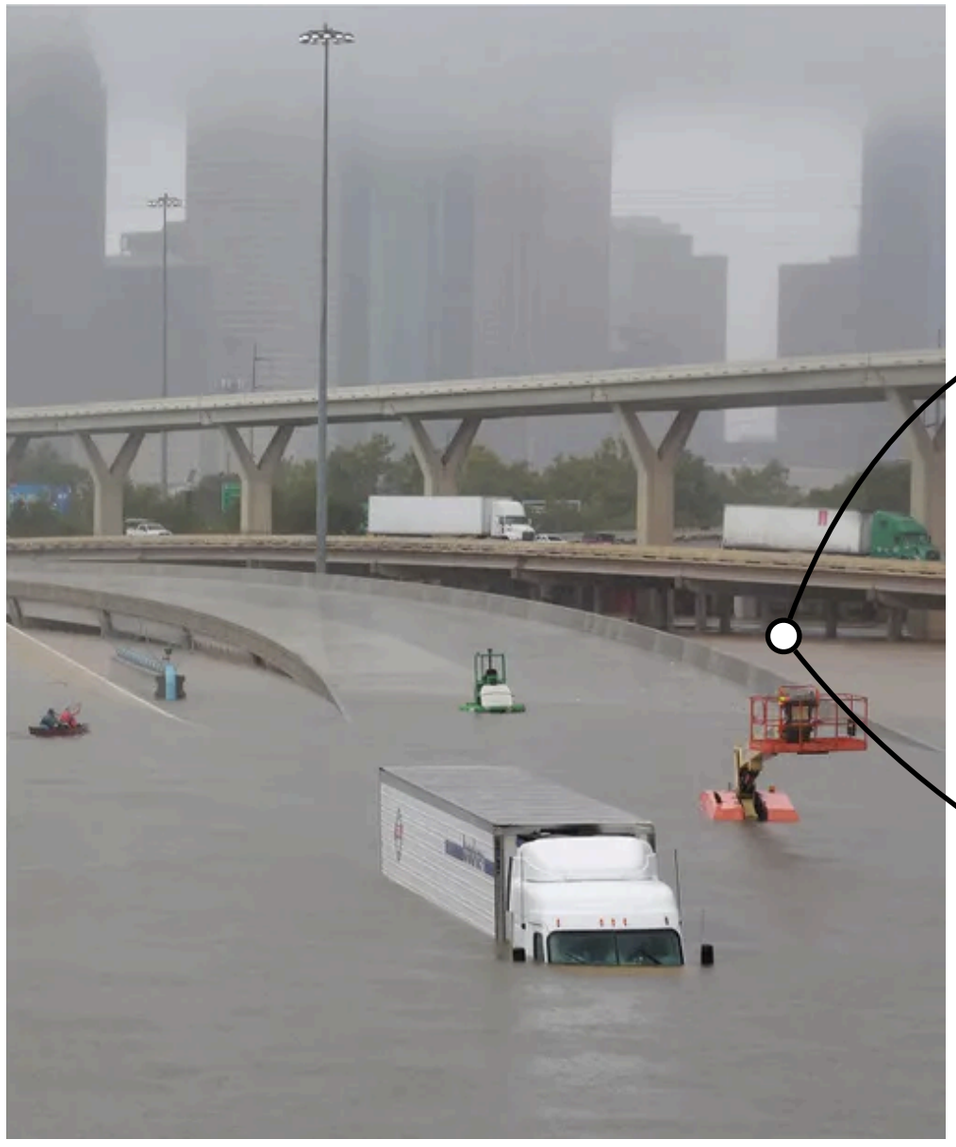


[1]

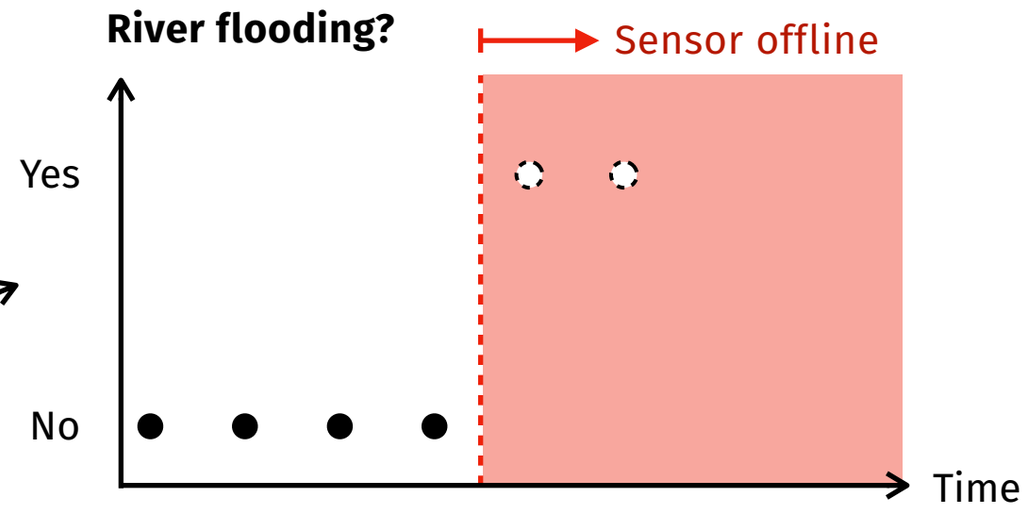


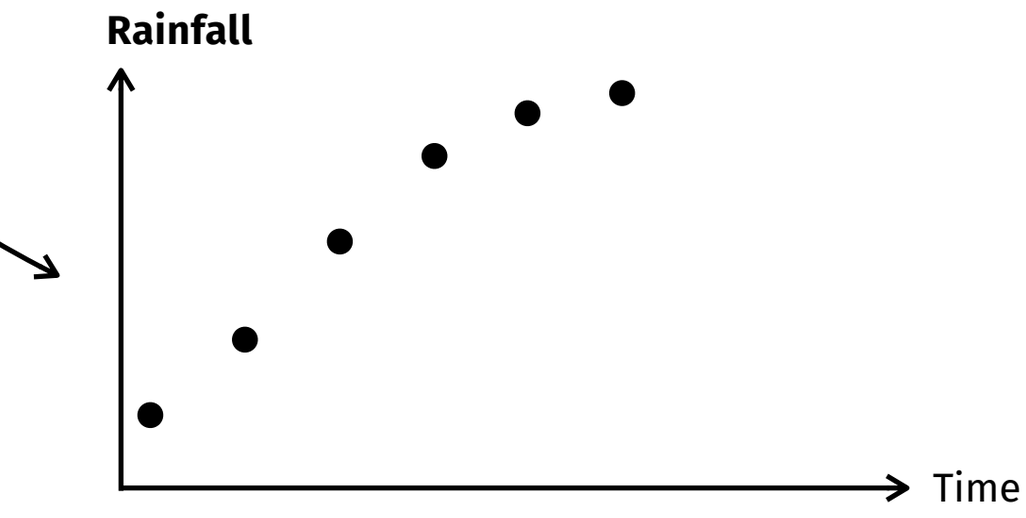


What problem are we addressing?

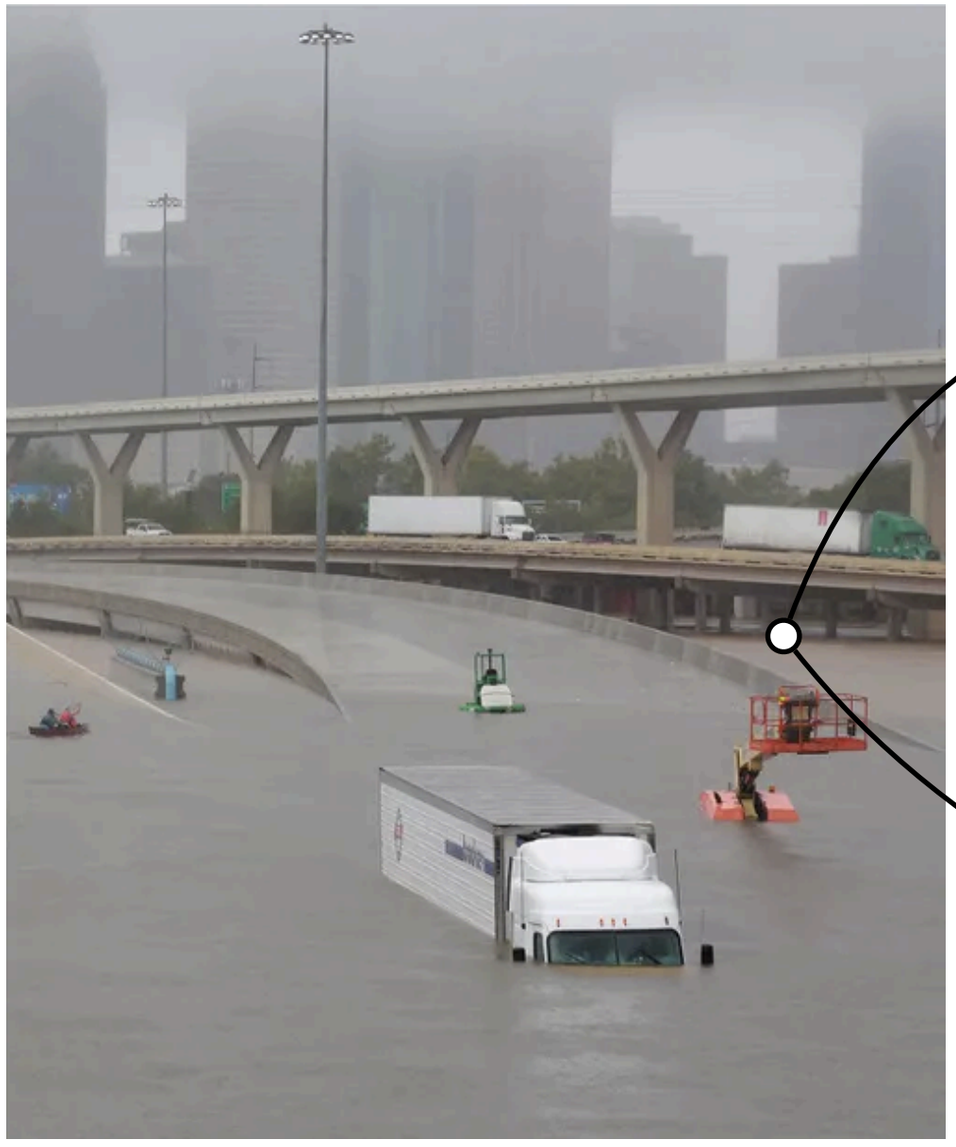


[1]

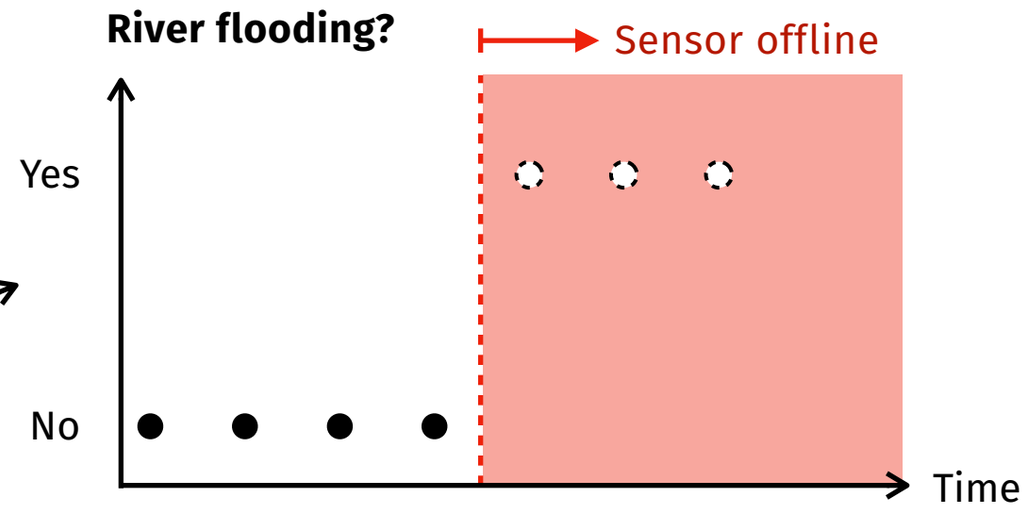


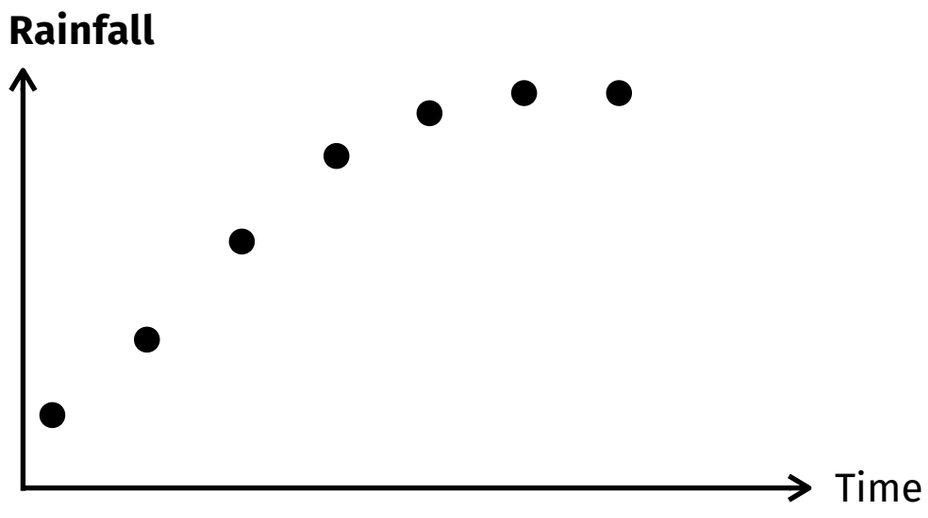


What problem are we addressing?

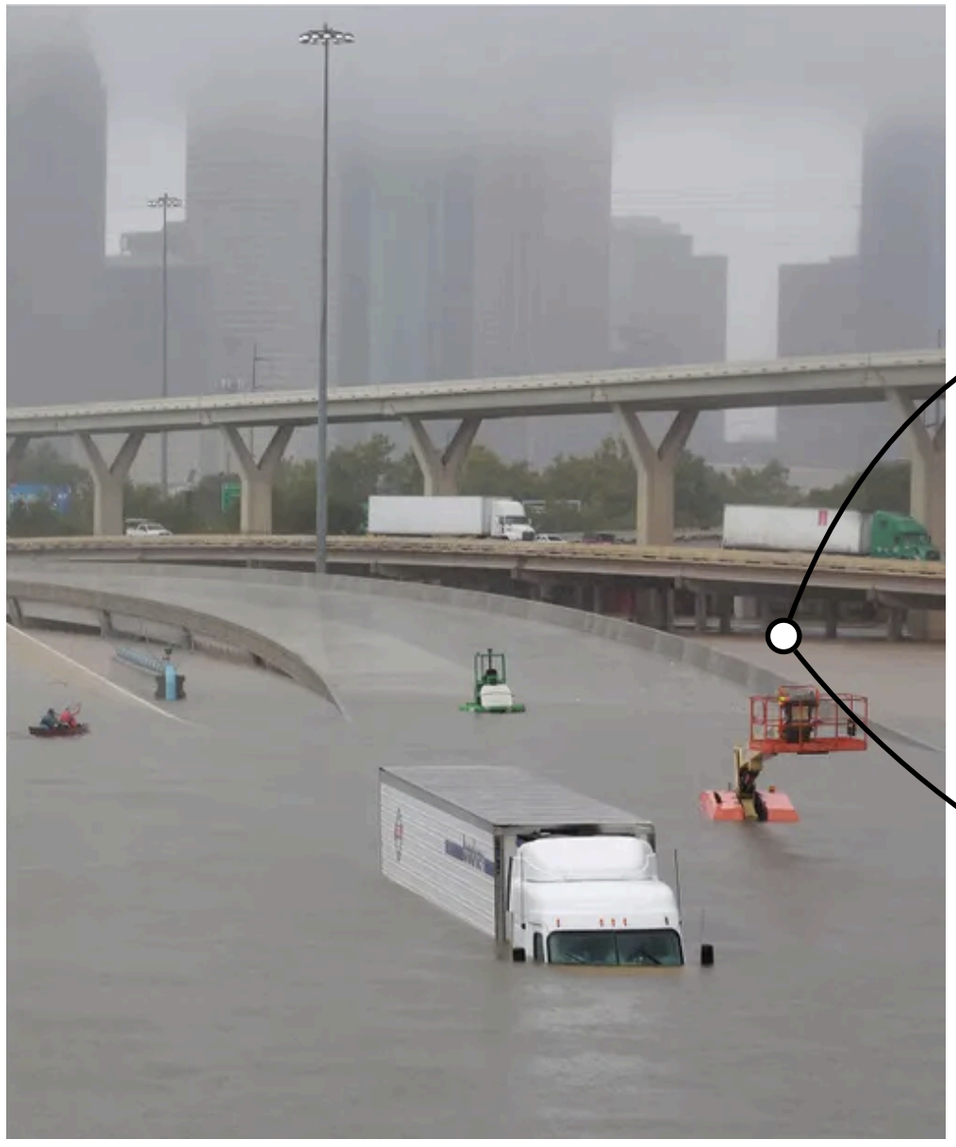


[1]

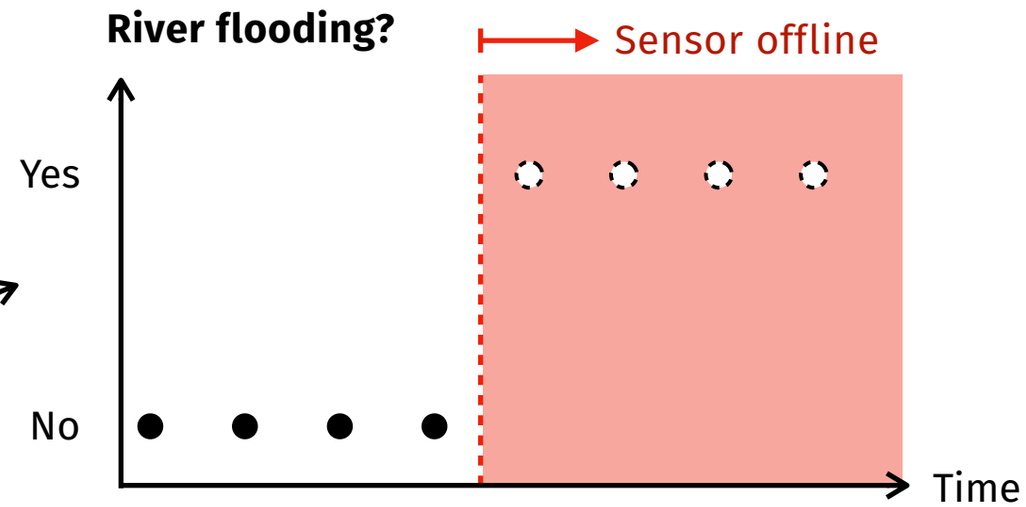


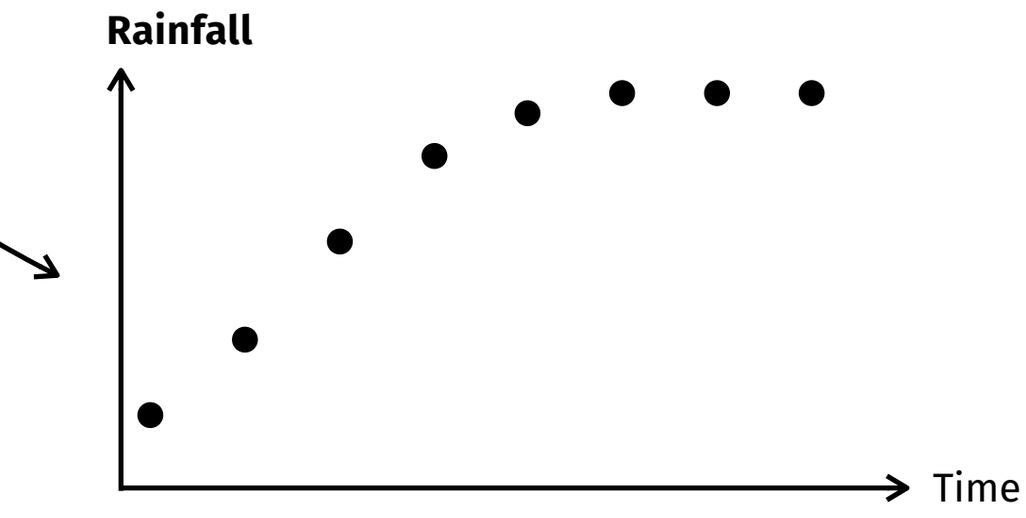


What problem are we addressing?

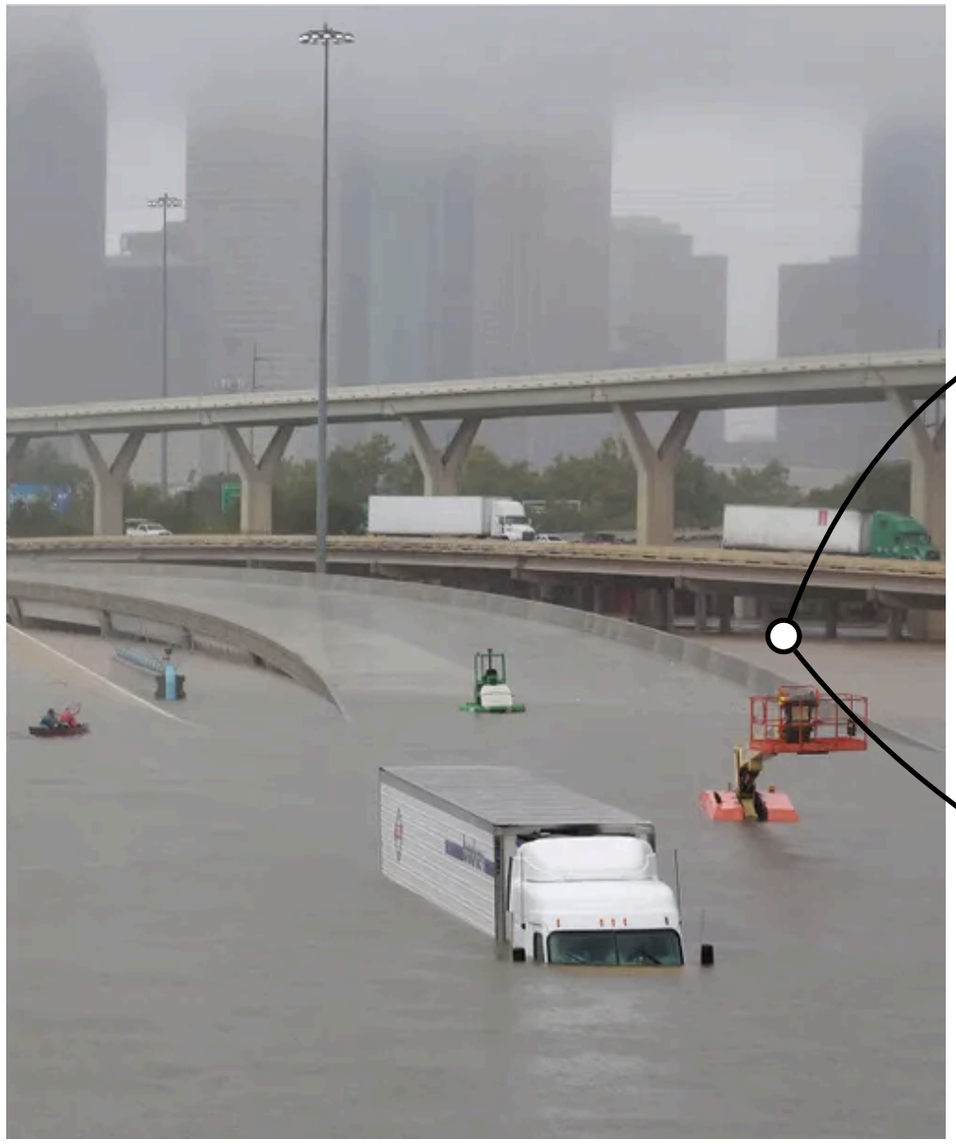


[1]

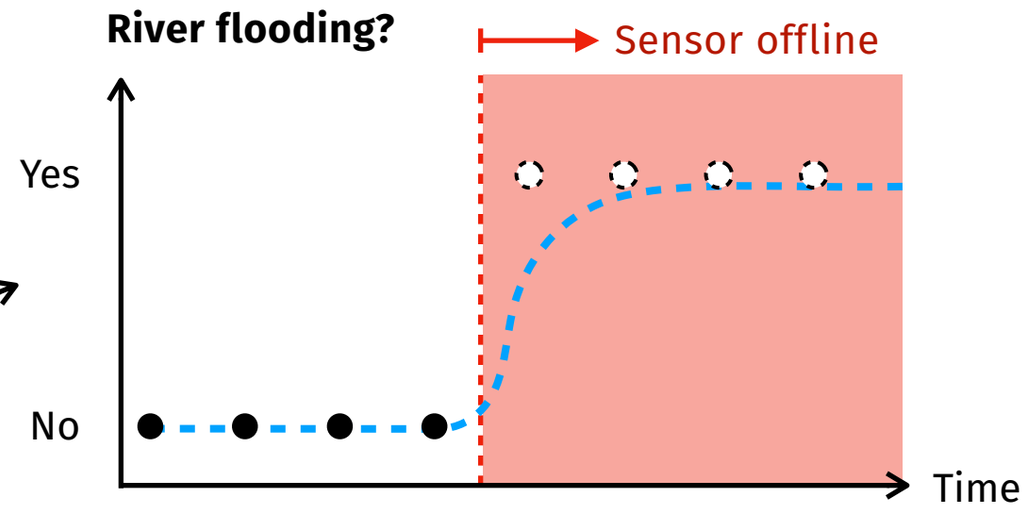


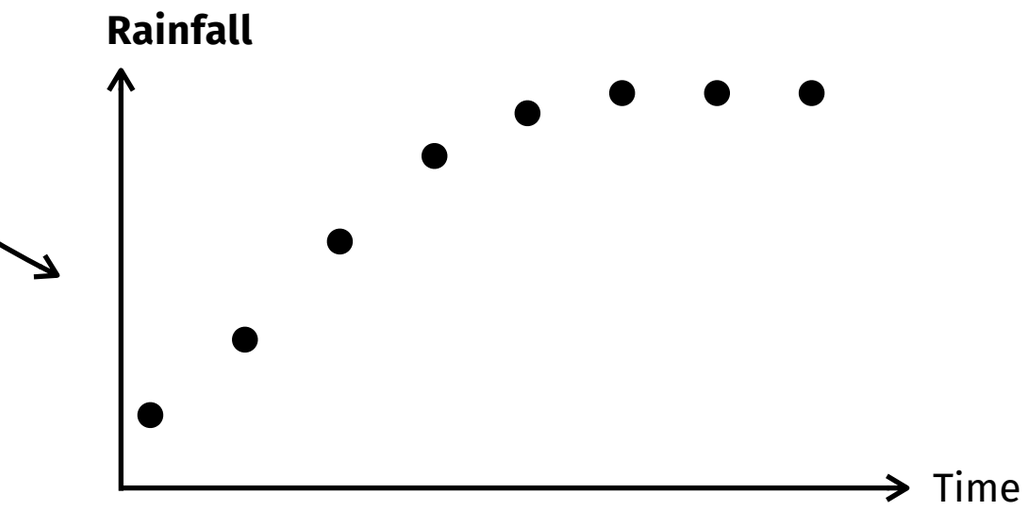


What problem are we addressing?

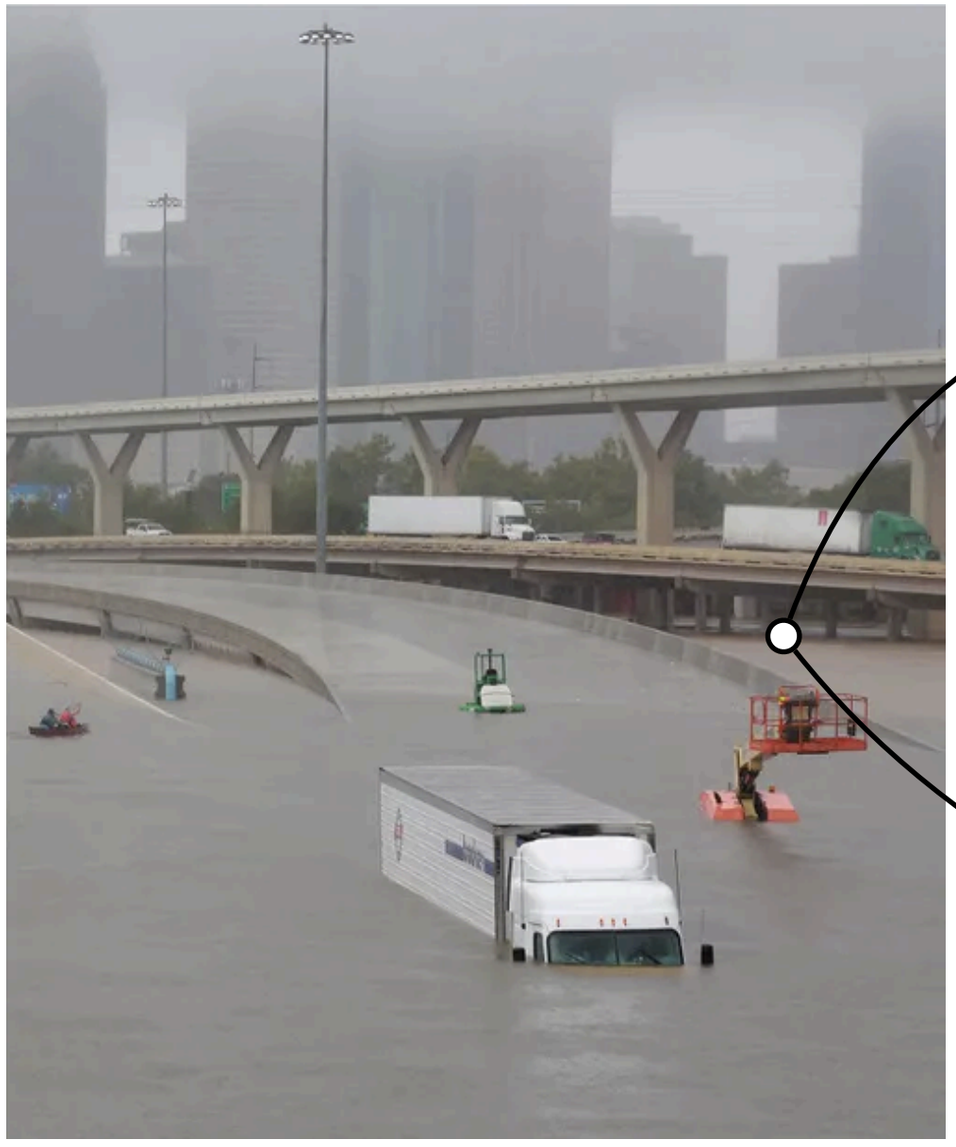


[1]

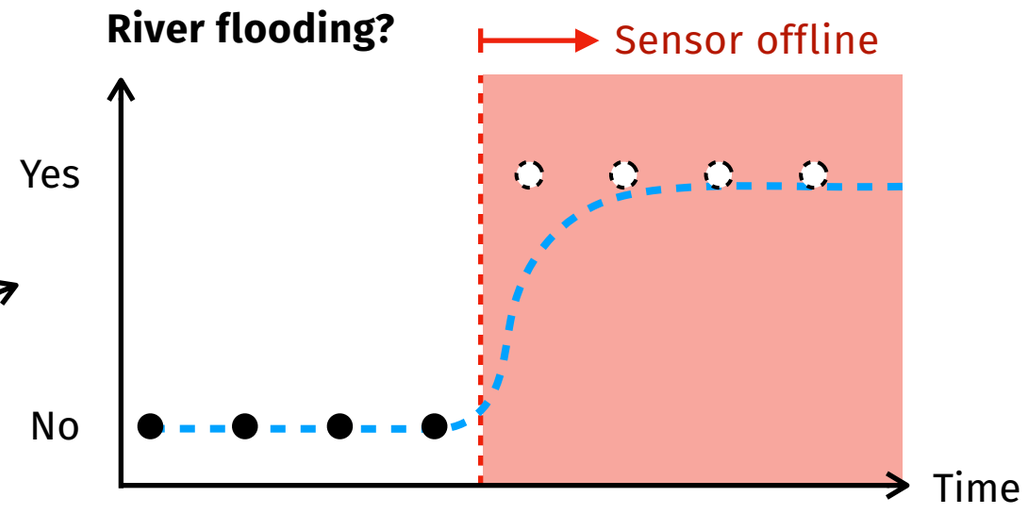


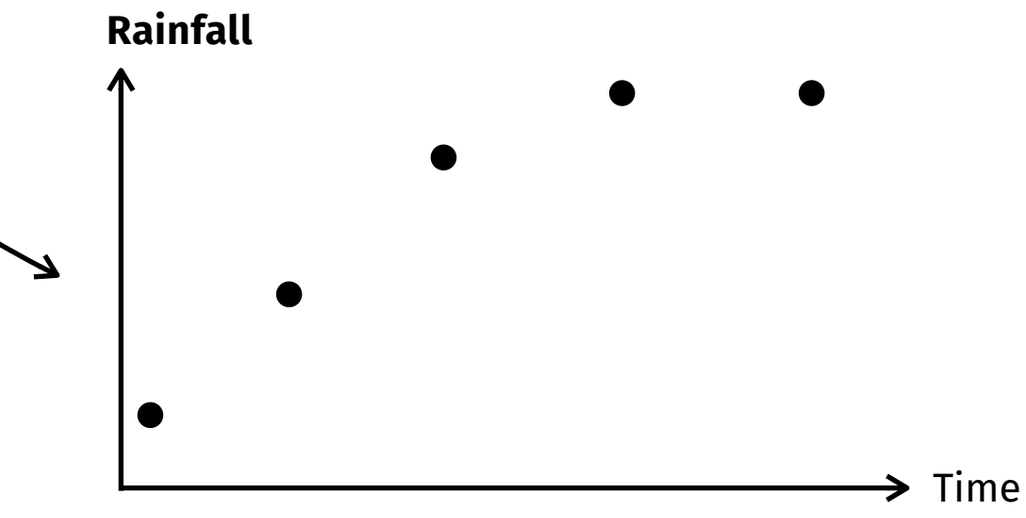


What problem are we addressing?

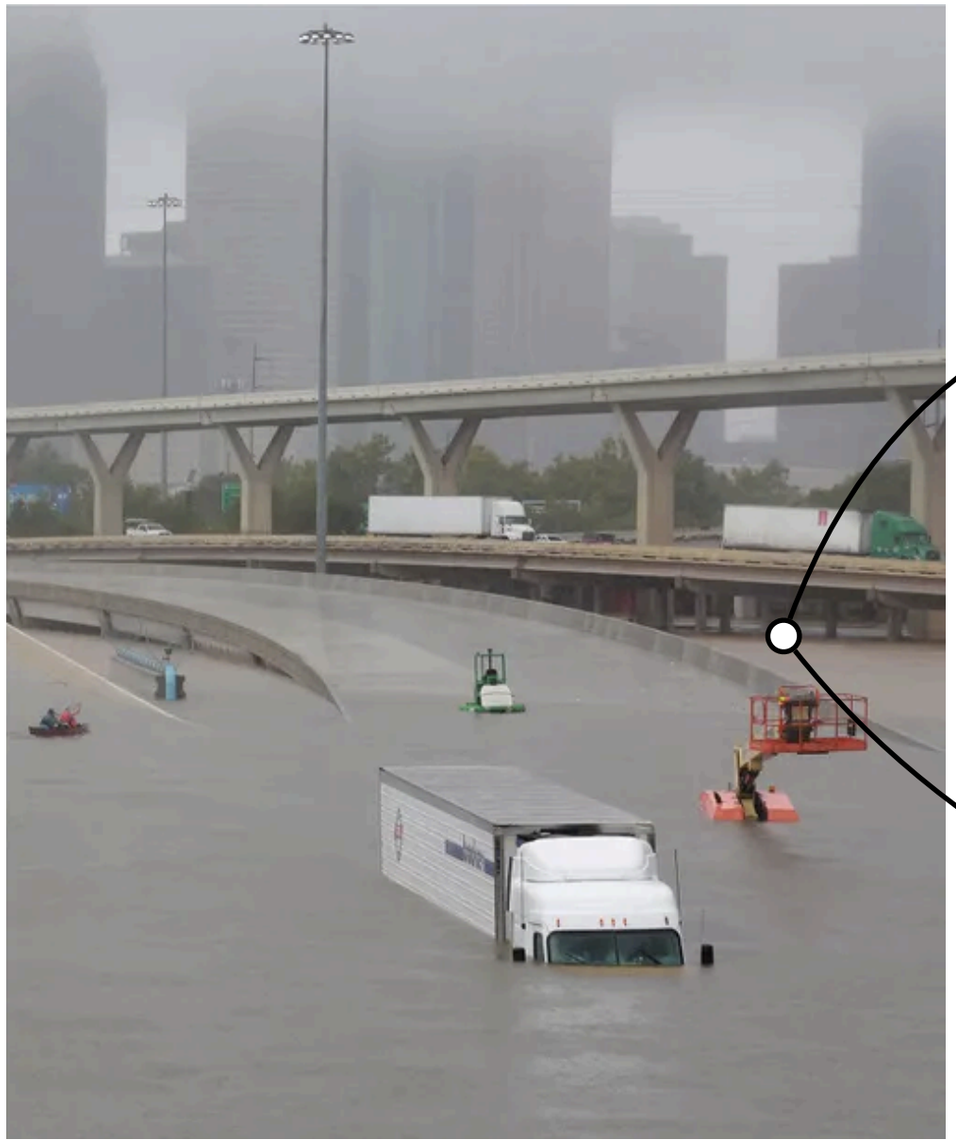


[1]

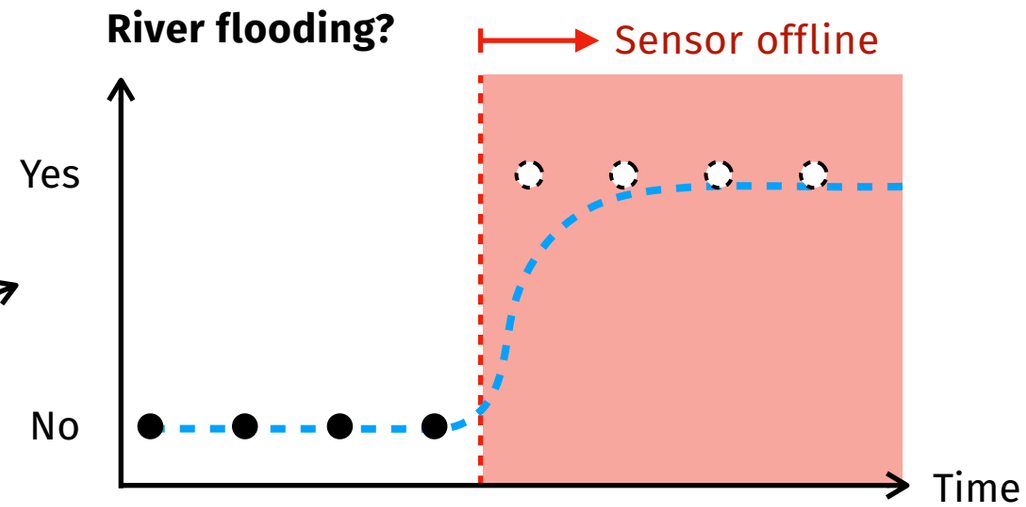


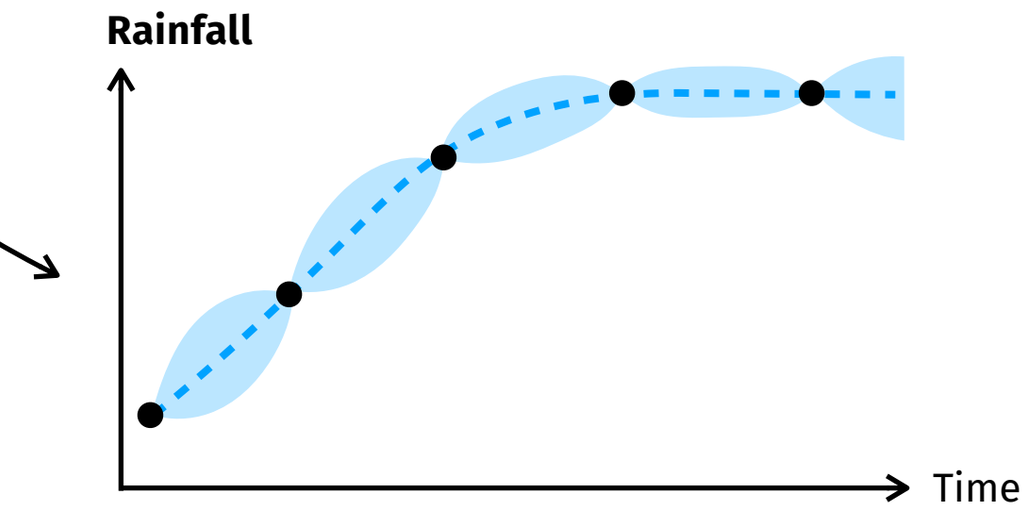


What problem are we addressing?

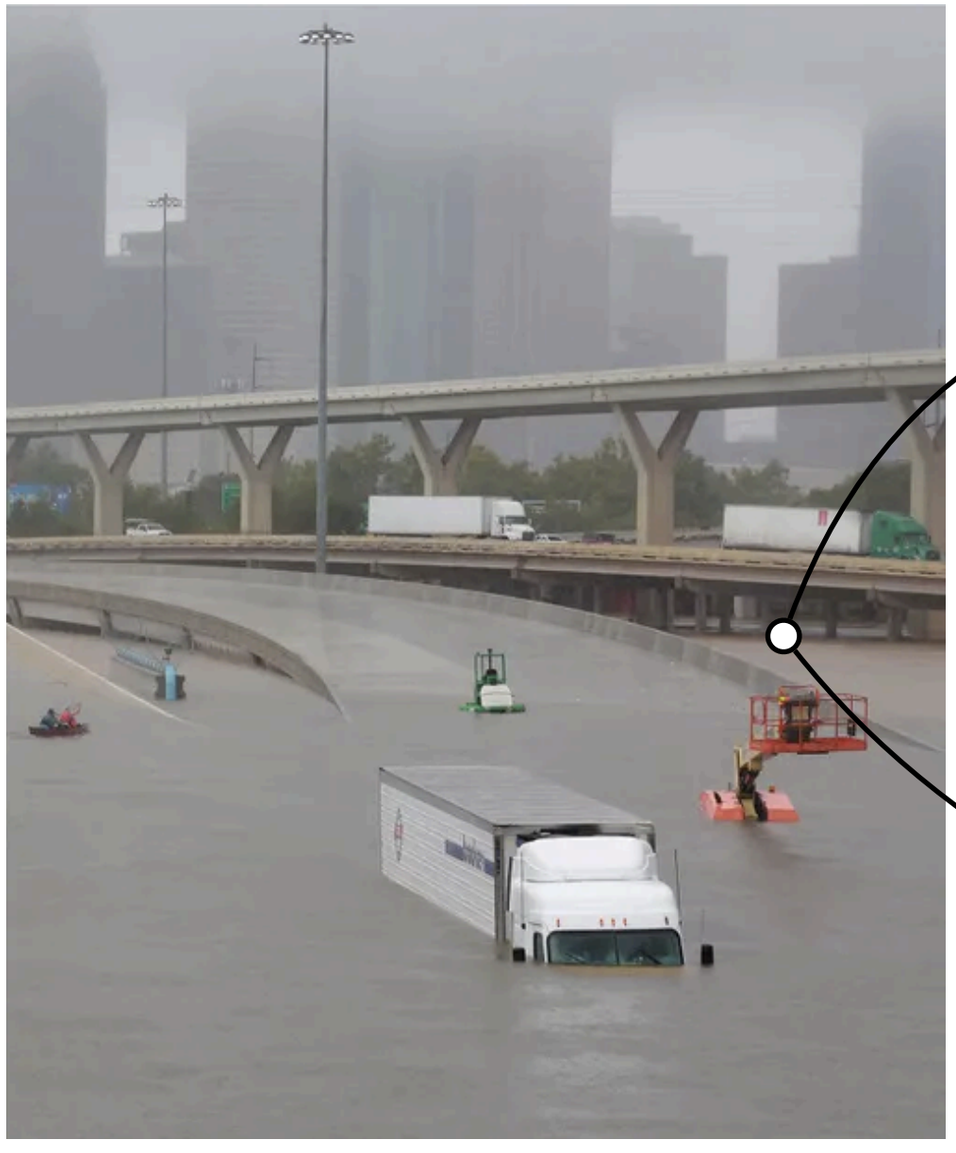


[1]

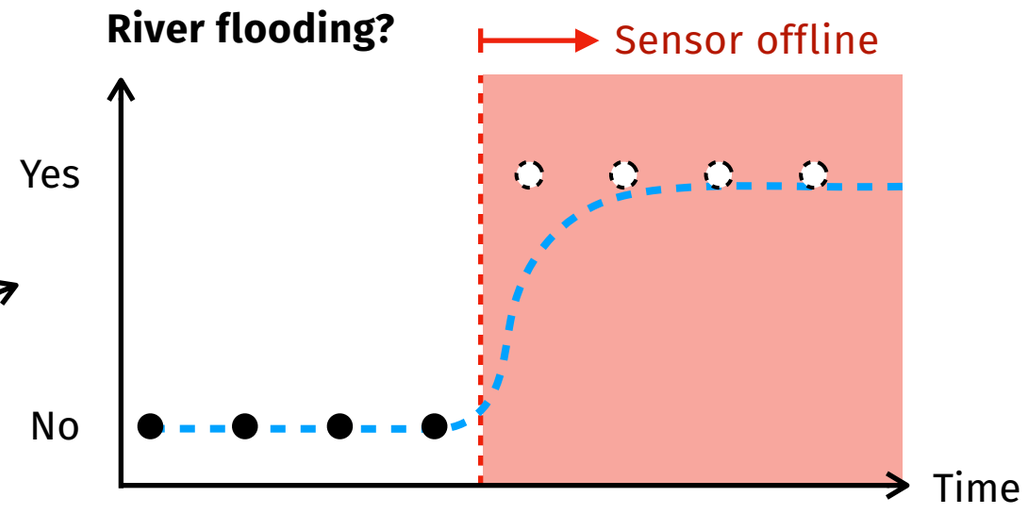


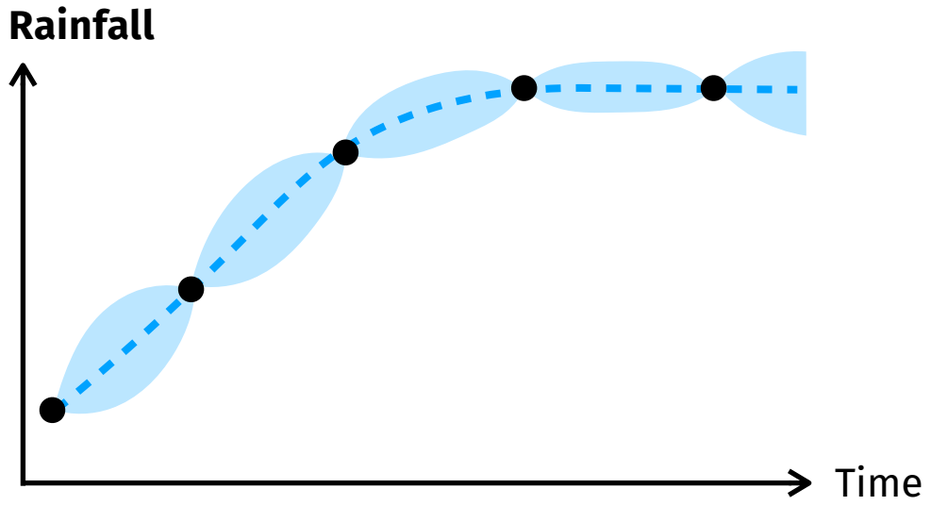


What problem are we addressing?



[1]





Can we make joint predictions over multiple outputs where data is missing with uncertainty?

Multi-output neural processes

What are neural processes?

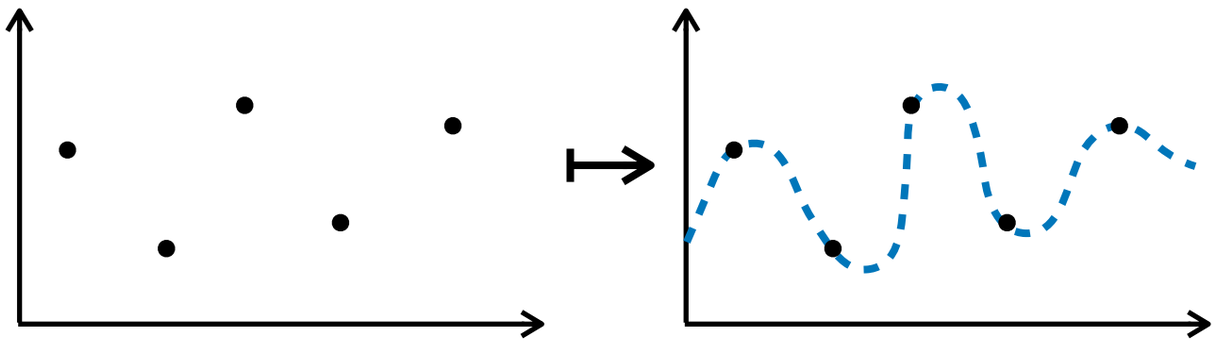
Traditional supervised learning

Neural processes

What are neural processes?

Traditional supervised learning

Dataset $\mathscr{D}$ $\longmapsto$ Prediction $f(x)$



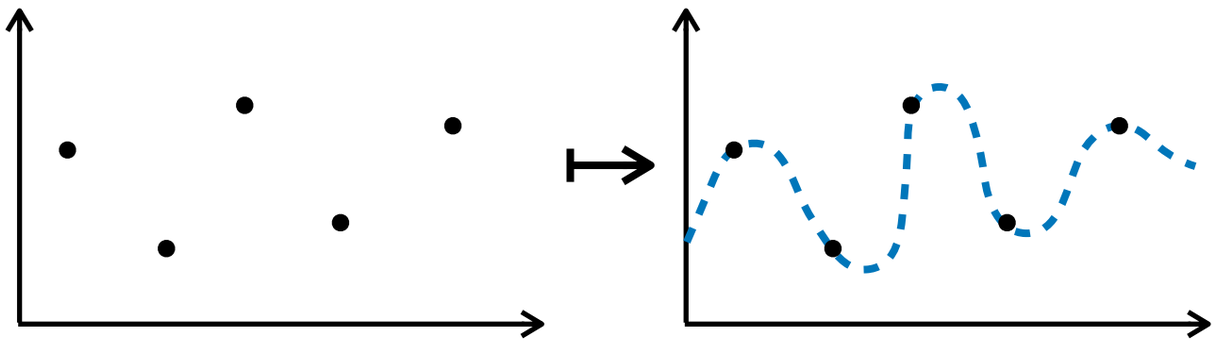
Training

Neural processes

What are neural processes?

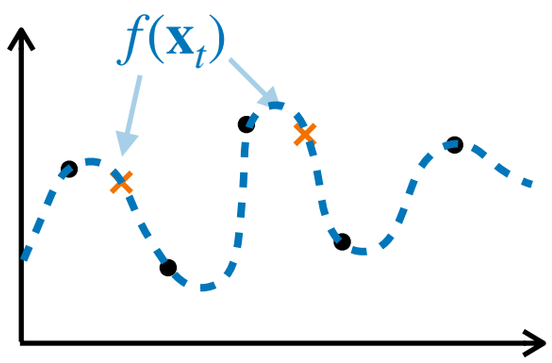
Traditional supervised learning

Dataset $\mathscr{D}$ $\longmapsto$ Prediction $f(x)$



Training

Test

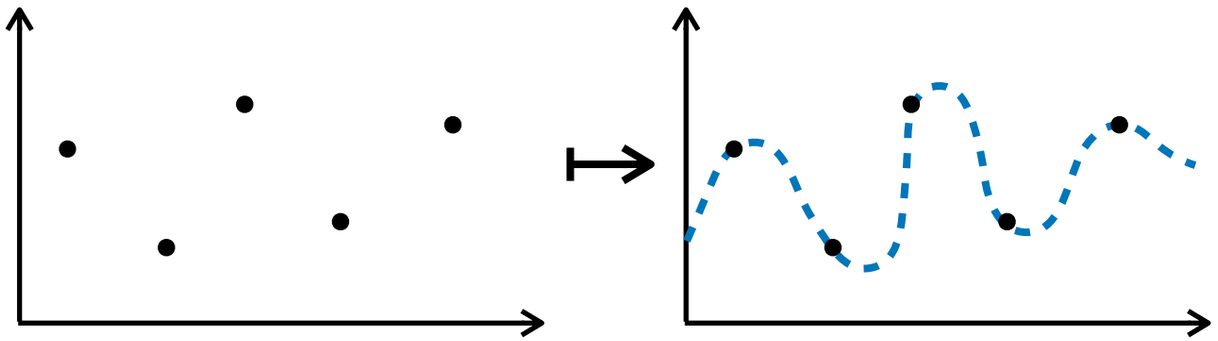


Neural processes

What are neural processes?

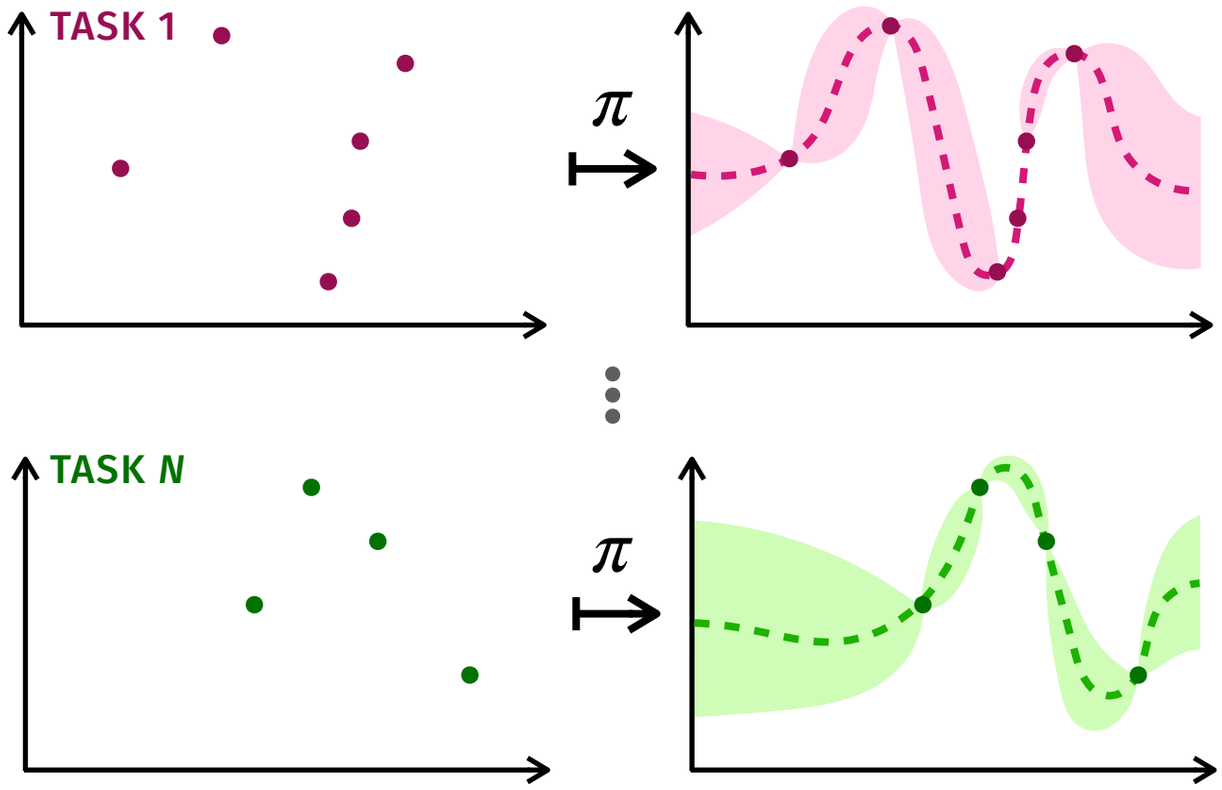
Traditional supervised learning

Dataset $\mathscr{D}$ $\longmapsto$ Prediction $f(x)$



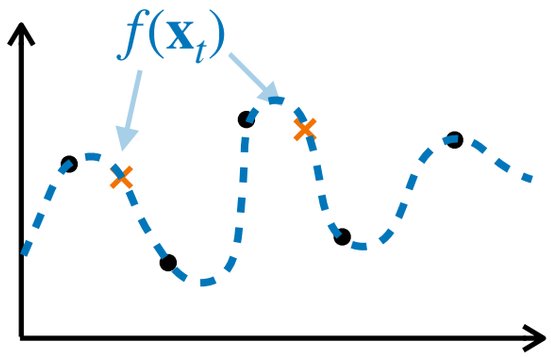
Neural processes

Dataset $\mathscr{D}$ $\overset{\pi}{\longmapsto}$ Prediction $p(\mathbf{y}_t \mid \mathbf{x}_t; \mathscr{D})$



Training

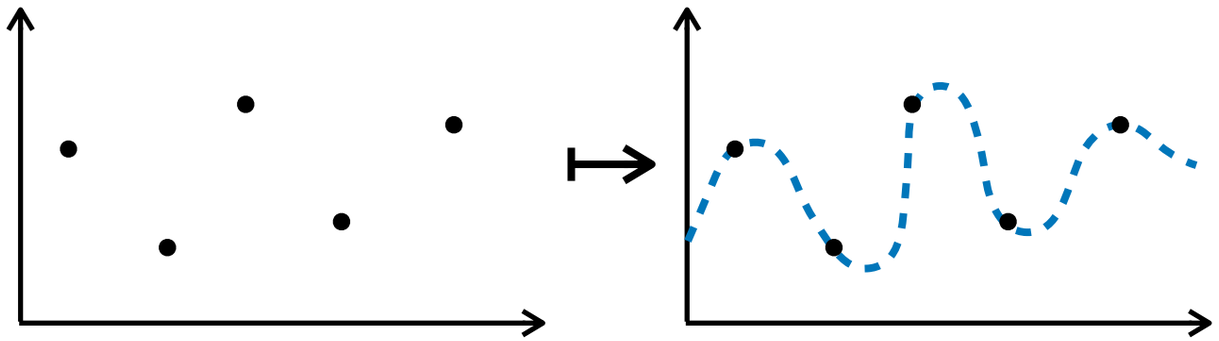
Test



What are neural processes?

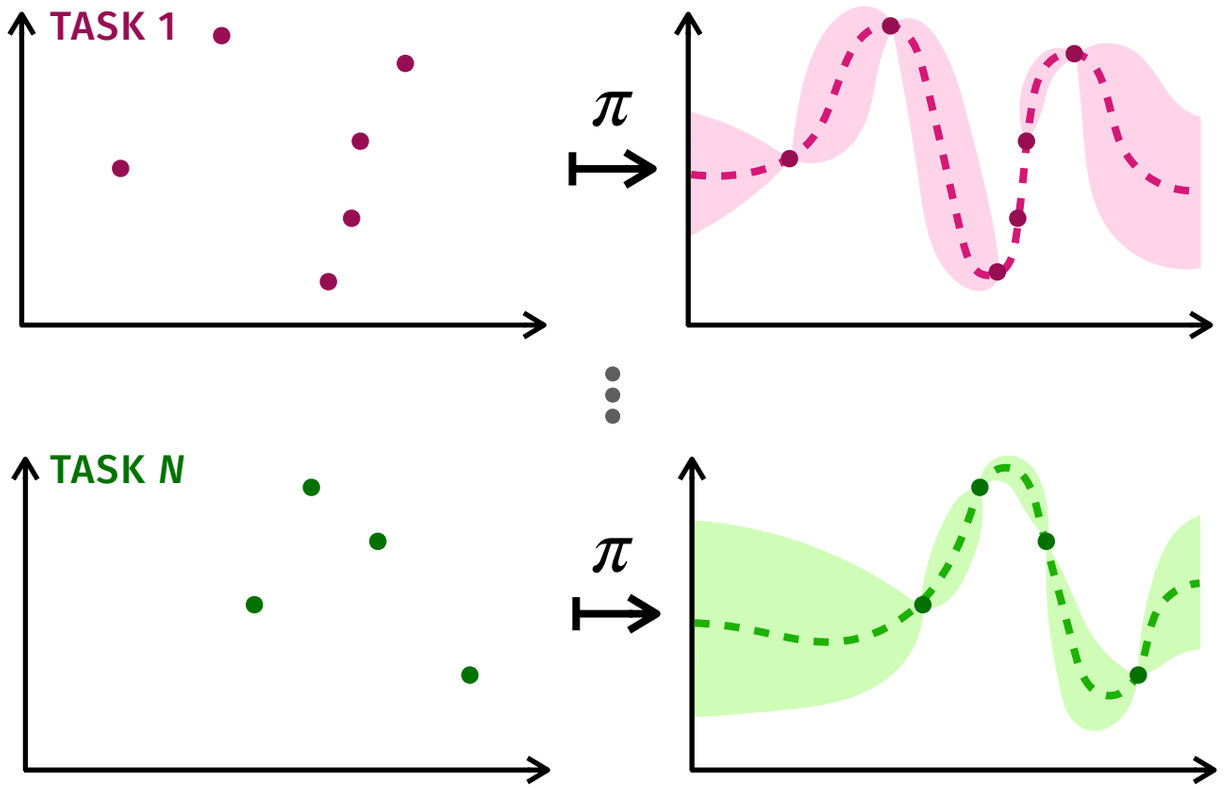
Traditional supervised learning

Dataset $\mathscr{D}$ $\longmapsto$ Prediction $f(x)$



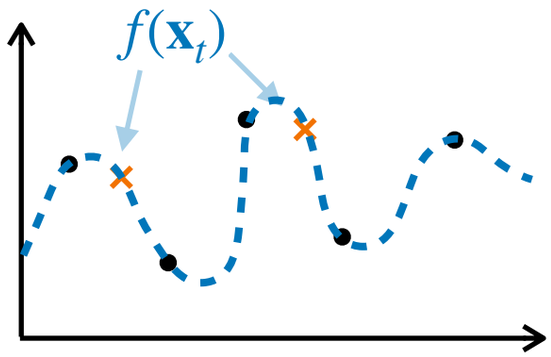
Neural processes

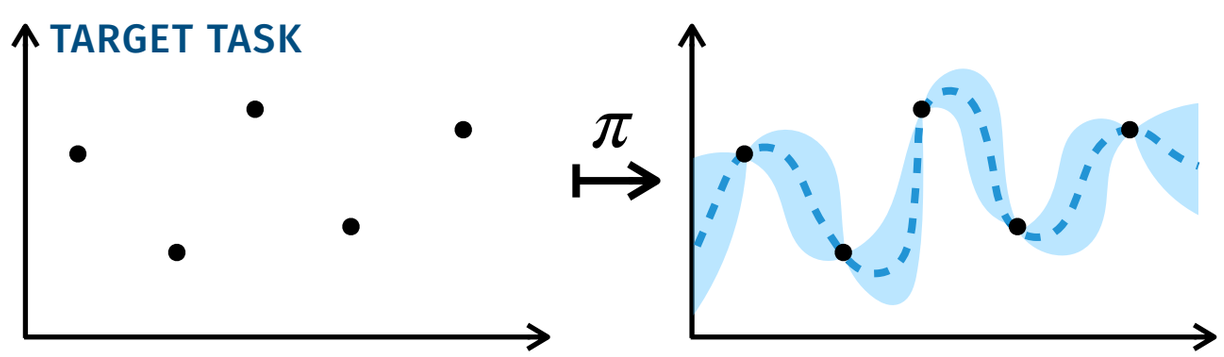
Dataset $\mathscr{D}$ $\overset{\pi}{\longmapsto}$ Prediction $p(\mathbf{y}_t \mid \mathbf{x}_t; \mathscr{D})$



Training

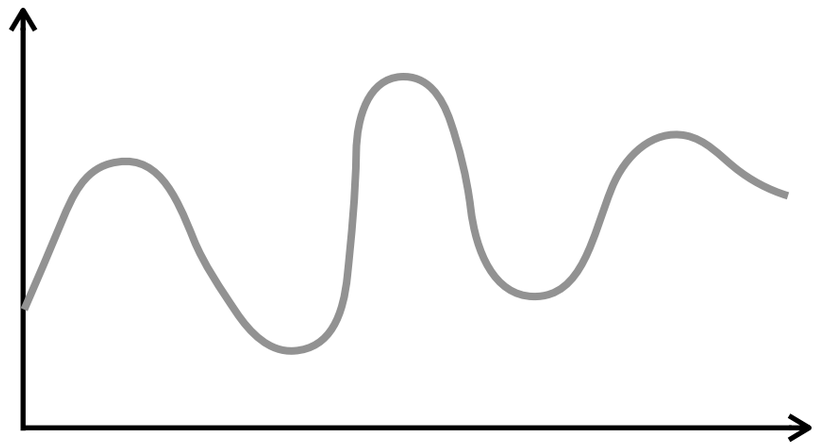
Test





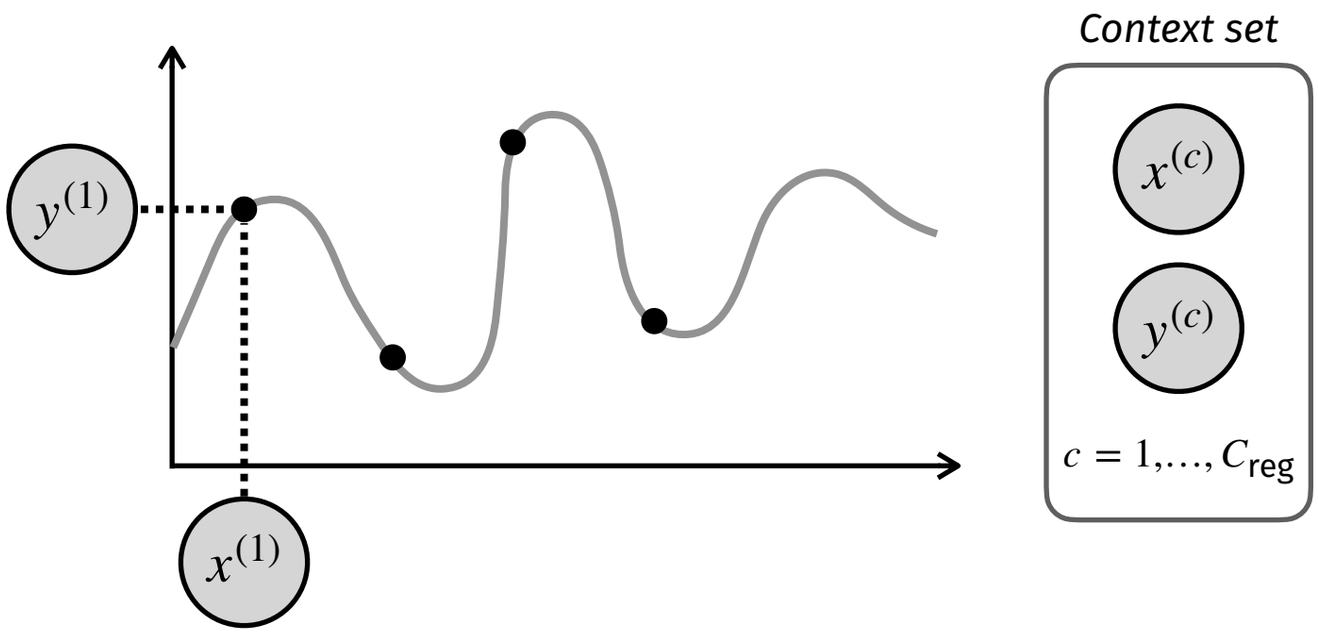
How do neural processes work?

● Context ✕ Target —— Underlying function



How do neural processes work?

● Context ✕ Target —— Underlying function



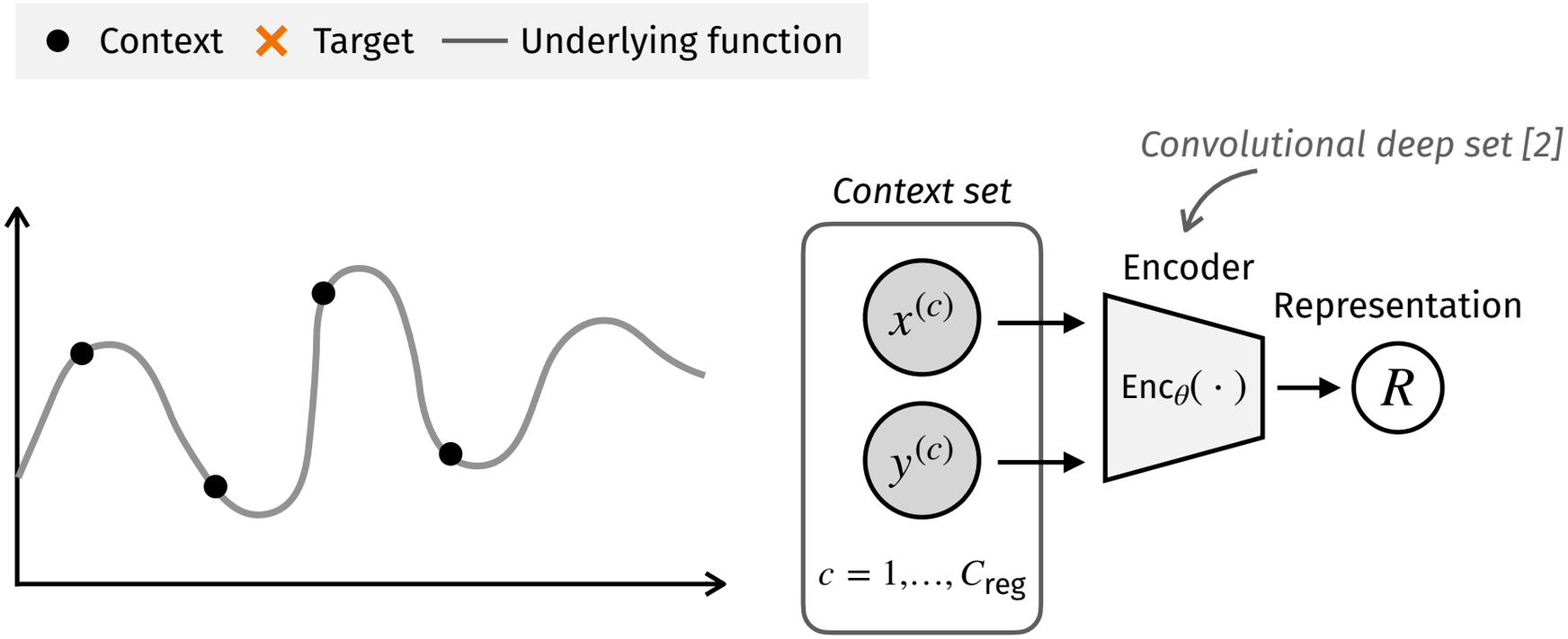
Context set

$x^{(c)}$

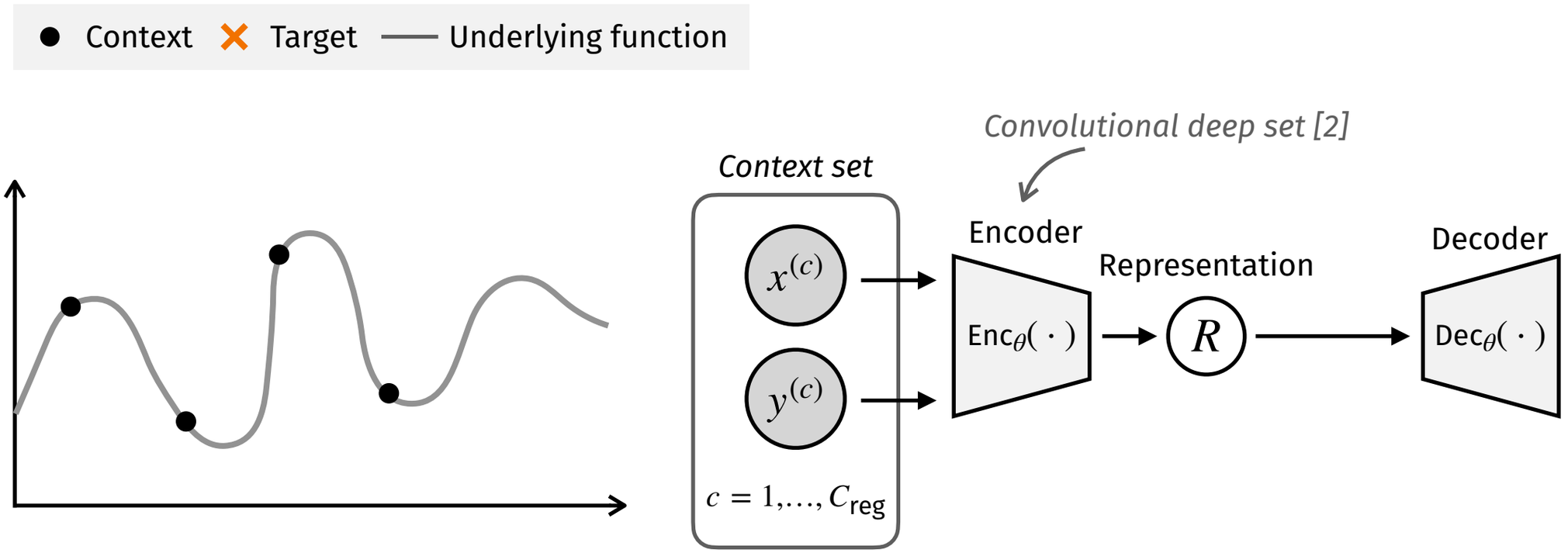
$y^{(c)}$

$c = 1, \ldots, C_{\text{reg}}$

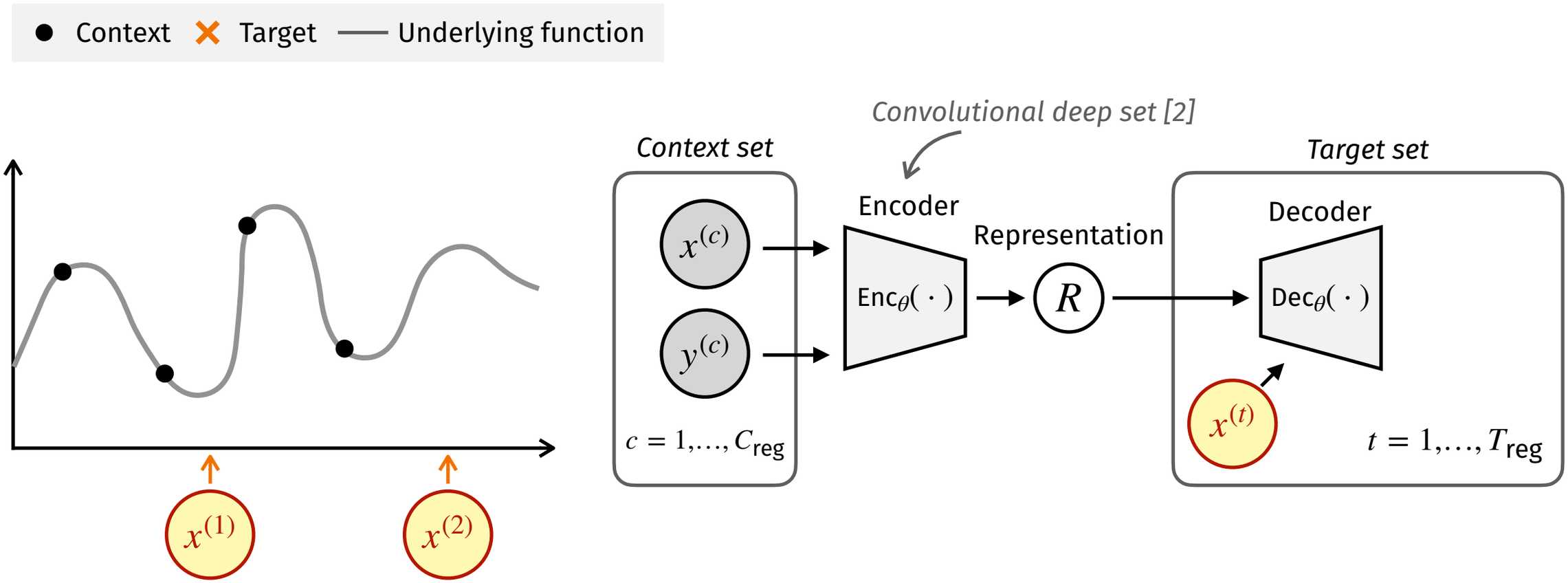
How do neural processes work?



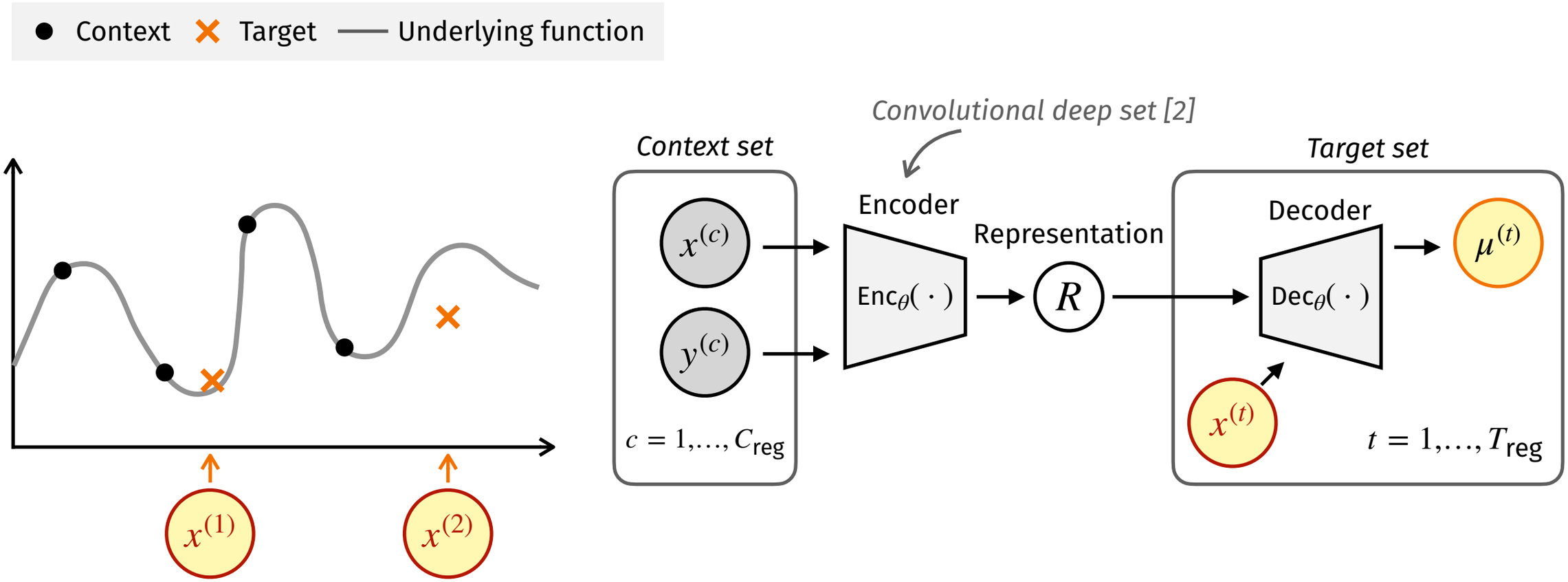
How do neural processes work?



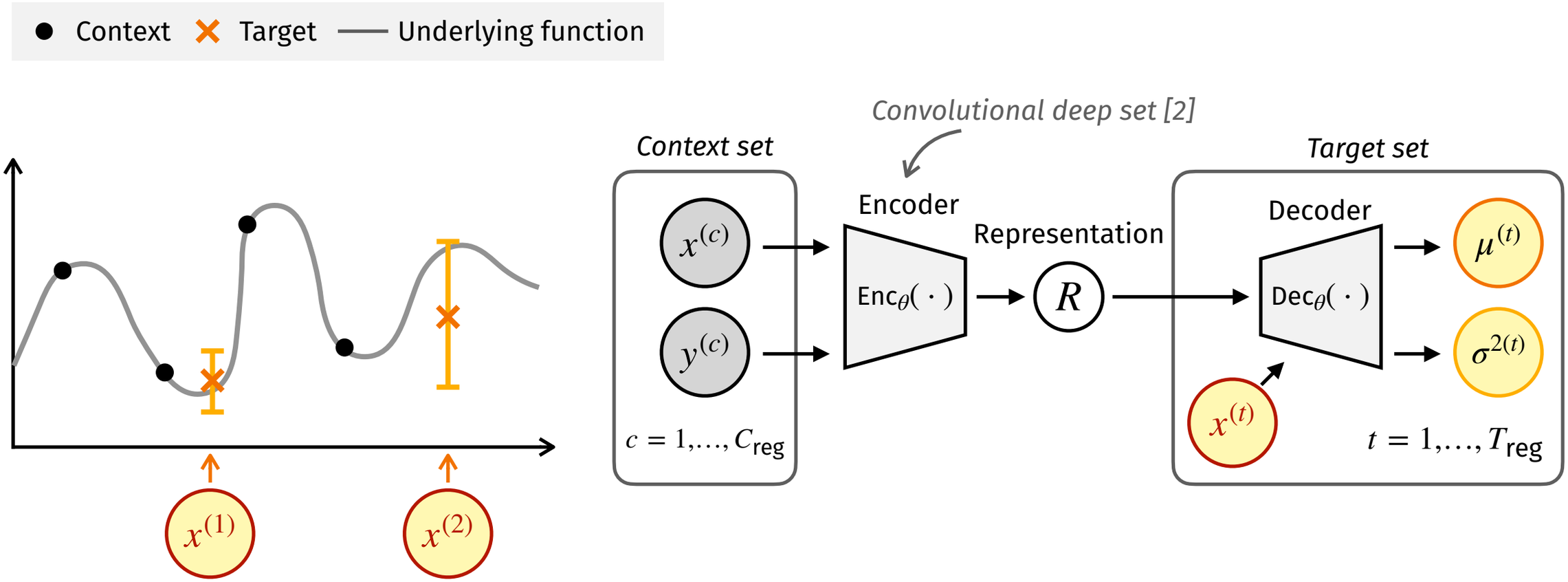
How do neural processes work?



How do neural processes work?



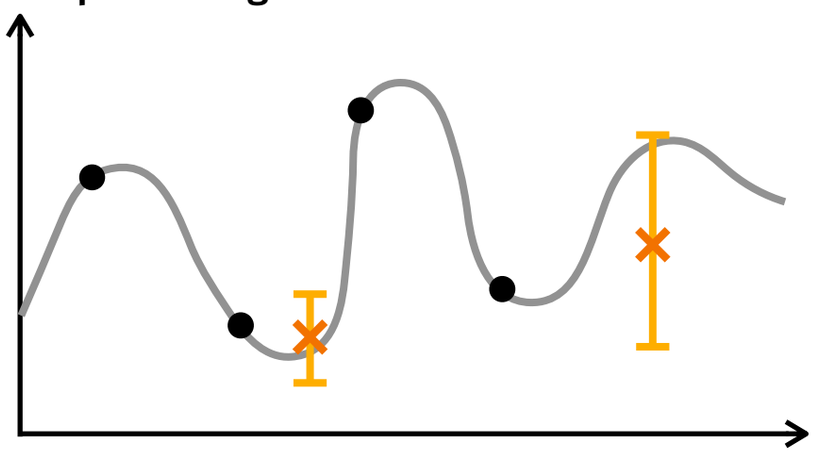
How do neural processes work?



How do neural processes work?

● Context ✕ Target —— Underlying function

Output 1: Regression



Convolutional deep set [2]

Context set

$x^{(c)}$

$y^{(c)}$

$c = 1, \ldots, C_{\text{reg}}$

Encoder

$\text{Enc}_\theta(\cdot)$

Representation

R

Target set

Decoder

$\text{Dec}_\theta(\cdot)$

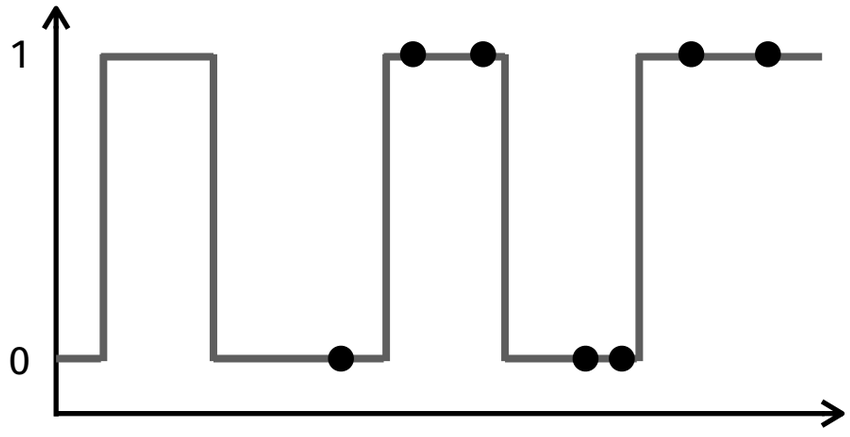
$\mu^{(t)}$

$\sigma^{2(t)}$

$x^{(t)}$

$t = 1, \ldots, T_{\text{reg}}$

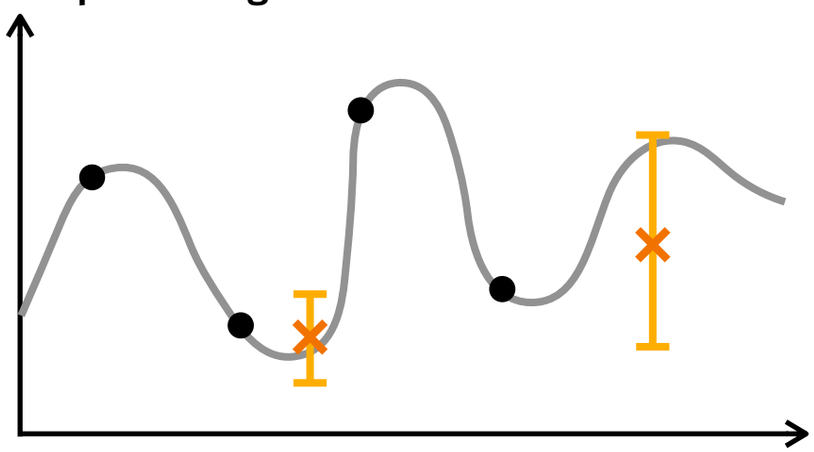
Output 2: Classification



How do neural processes work?

● Context ✕ Target —— Underlying function

Output 1: Regression



Convolutional deep set [2]

Context set

$x^{(c)}$

$y^{(c)}$

$c = 1,\ldots,C_{\text{reg}}$

Encoder

$\text{Enc}_\theta(\cdot)$

Representation

R

Target set

Decoder

$\text{Dec}_\theta(\cdot)$

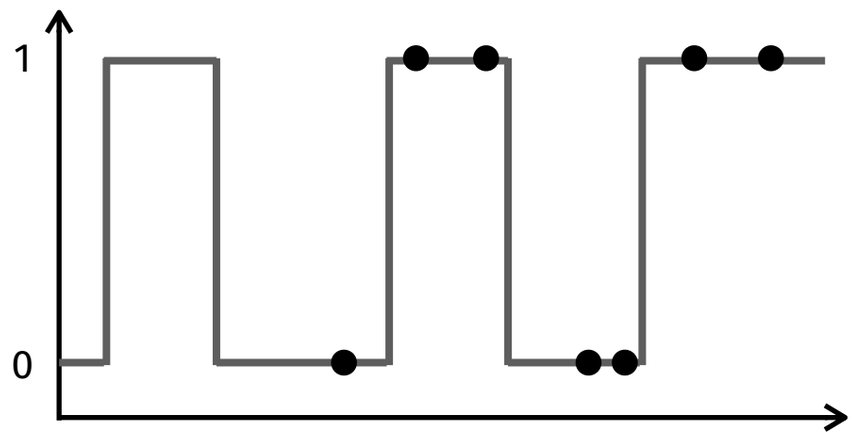
$\mu^{(t)}$

$\sigma^{2(t)}$

$x^{(t)}$

$t = 1,\ldots,T_{\text{reg}}$

Output 2: Classification



$x^{(c)}$

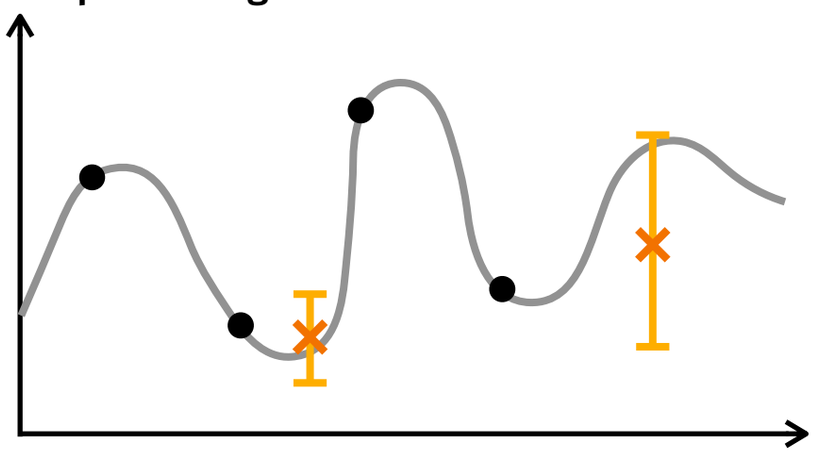
$y^{(c)}$

$c = 1,\ldots,C_{\text{class}}$

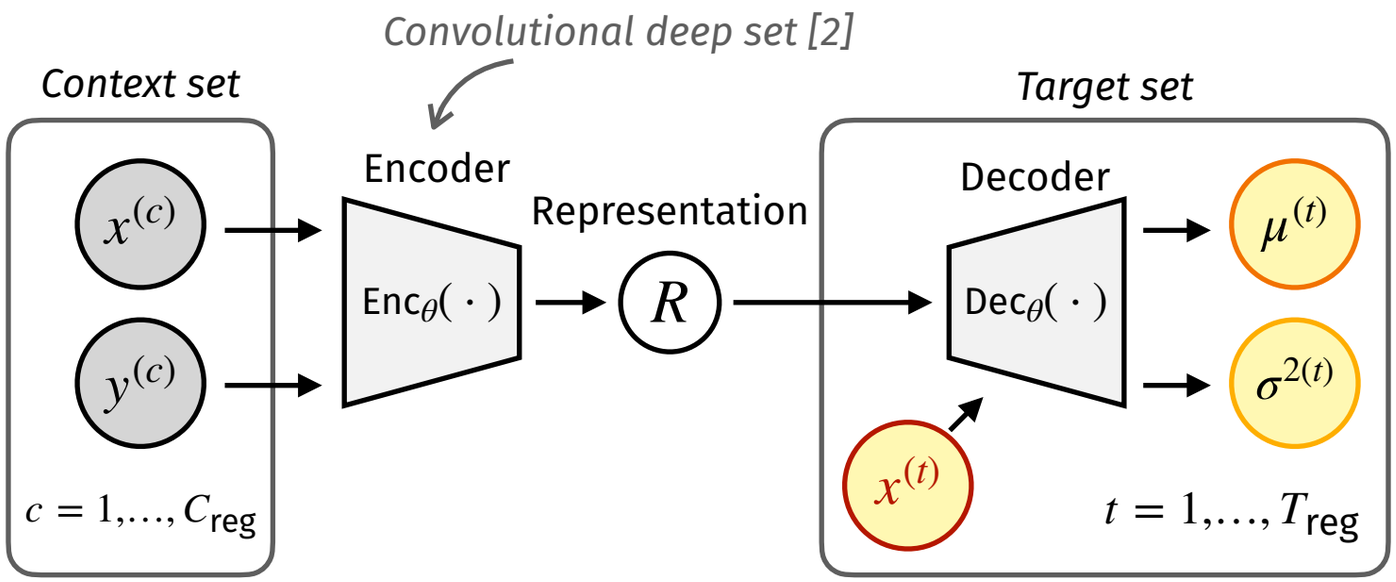
How do neural processes work?

● Context ✕ Target —— Underlying function

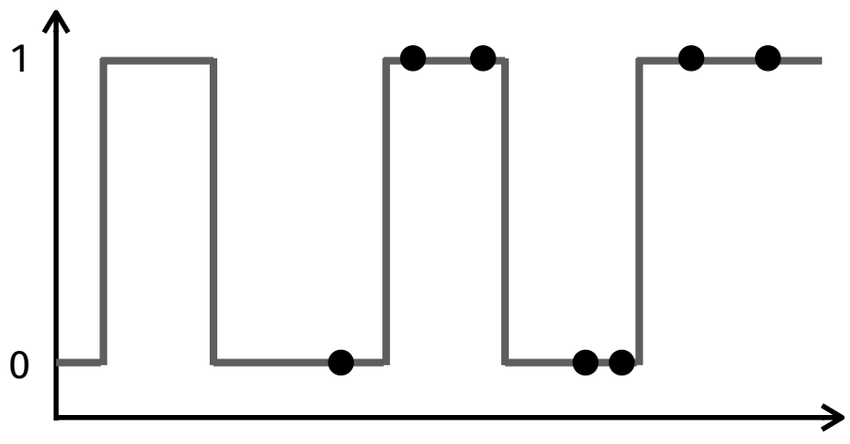
Output 1: Regression

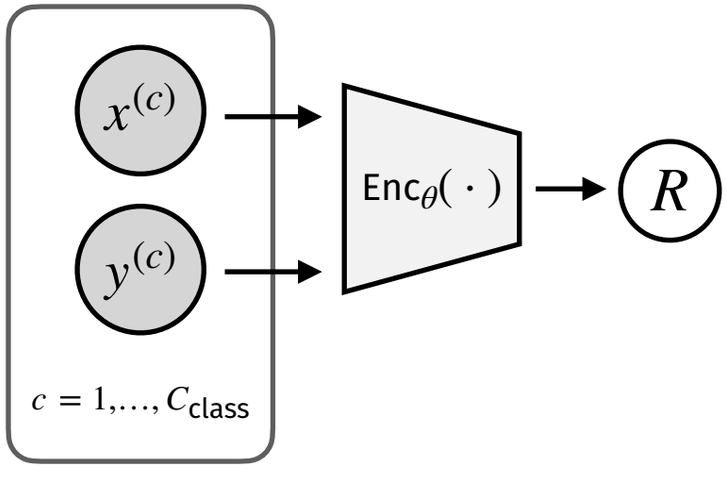


Convolutional deep set [2]



Output 2: Classification

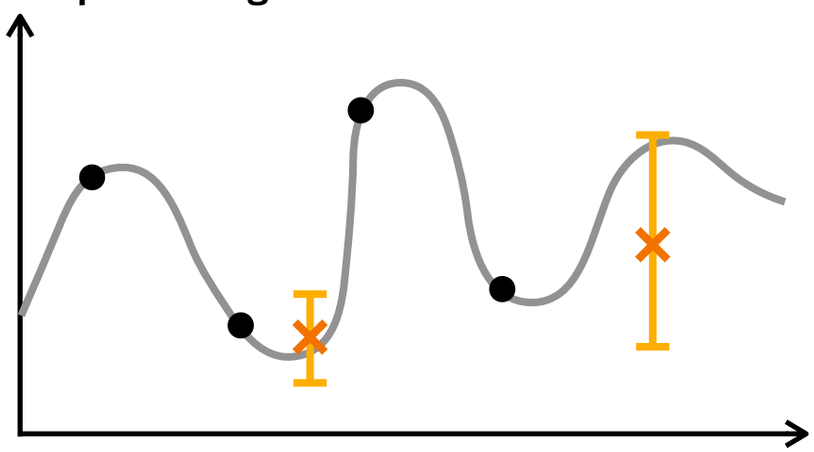




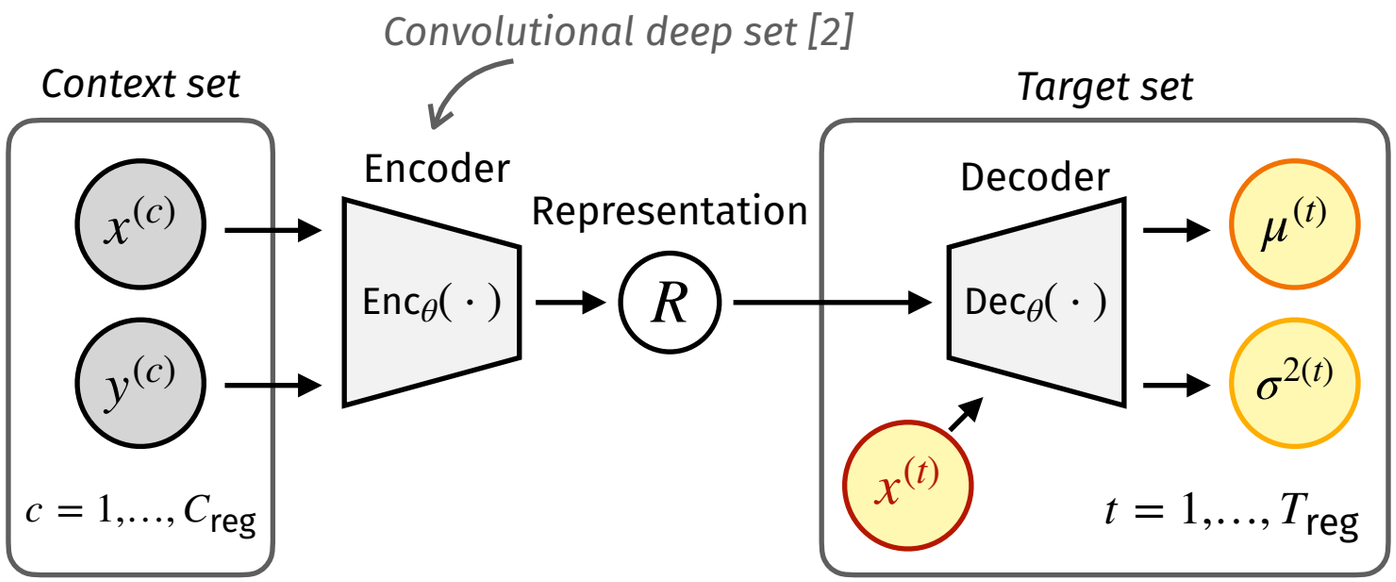
How do neural processes work?

● Context ✕ Target —— Underlying function

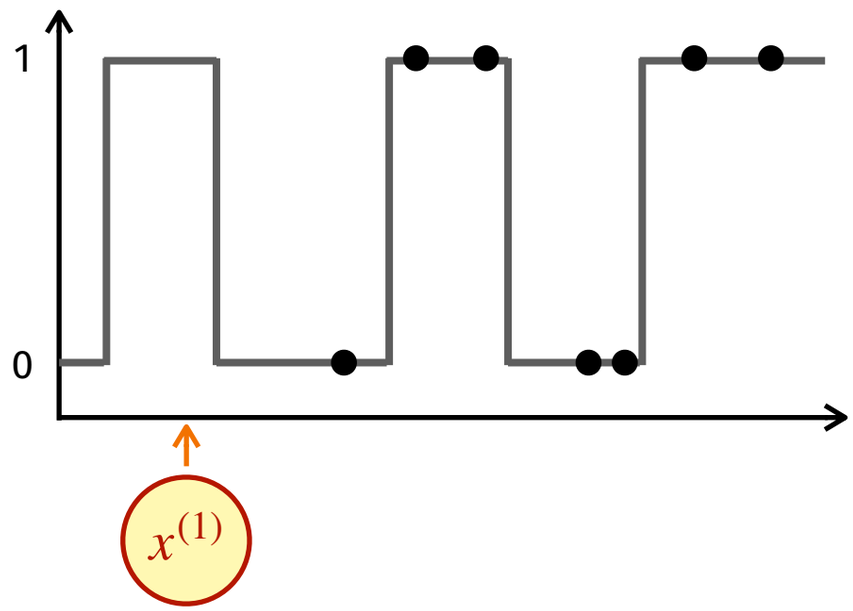
Output 1: Regression

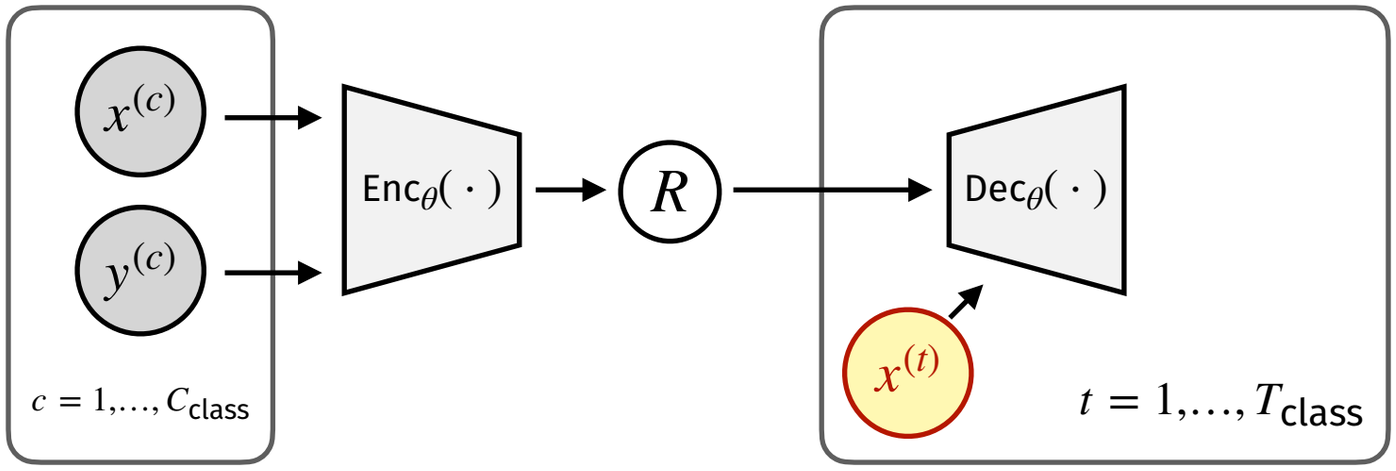


Convolutional deep set [2]



Output 2: Classification

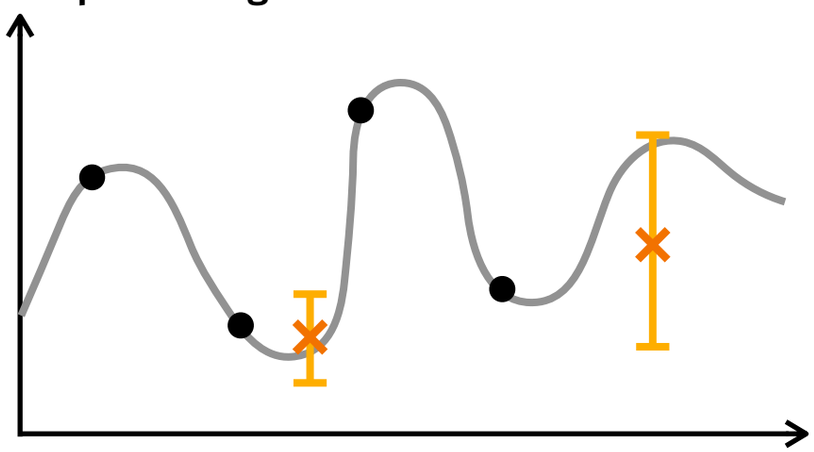




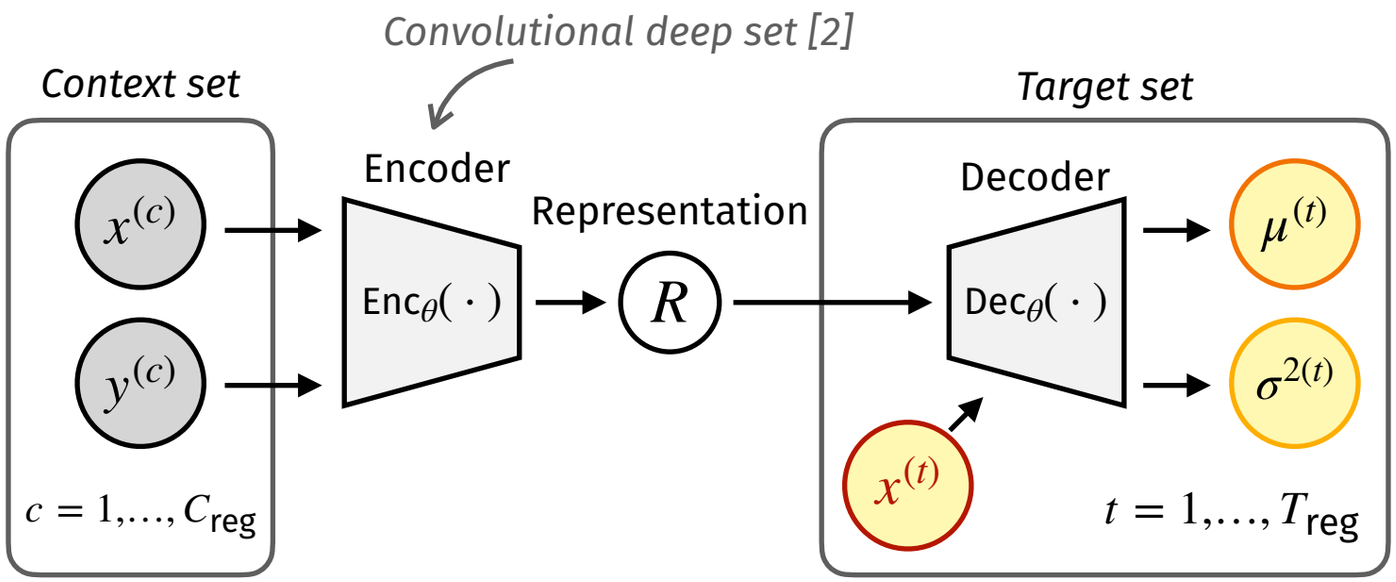
How do neural processes work?

● Context ✕ Target — Underlying function

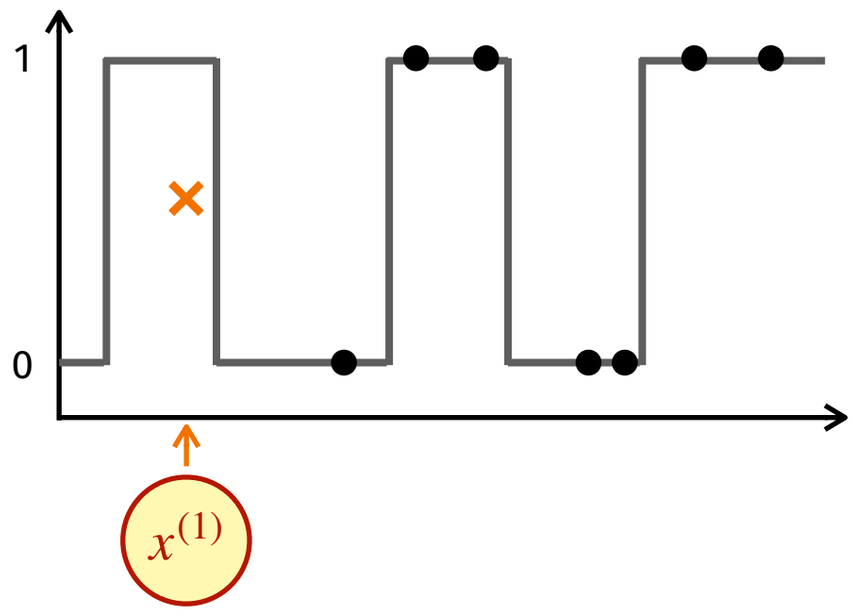
Output 1: Regression

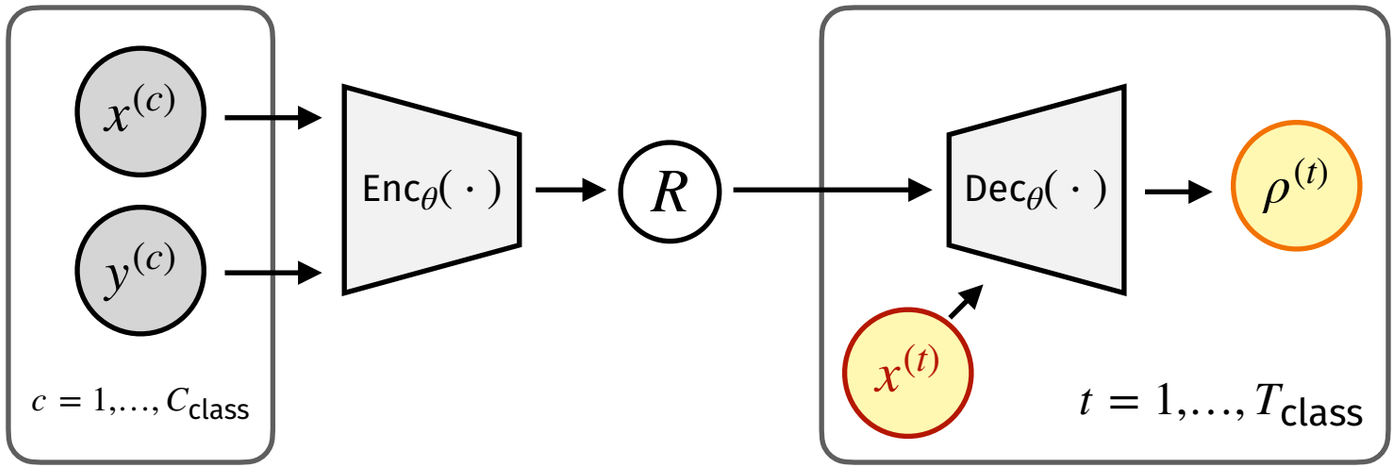


Convolutional deep set [2]



Output 2: Classification

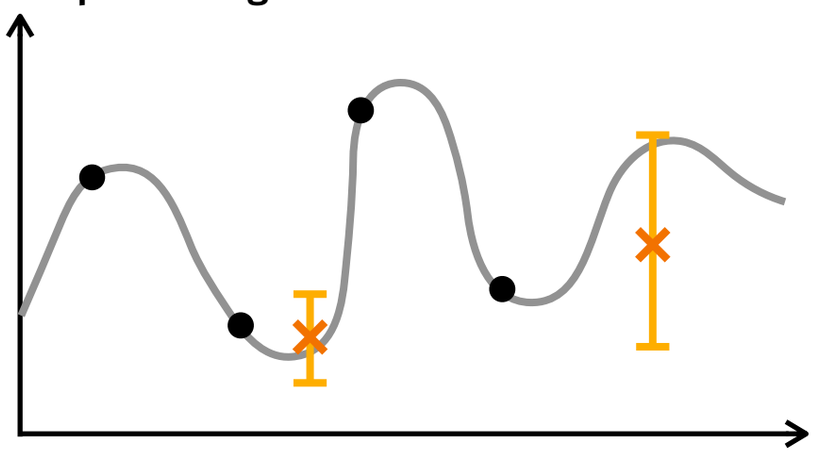




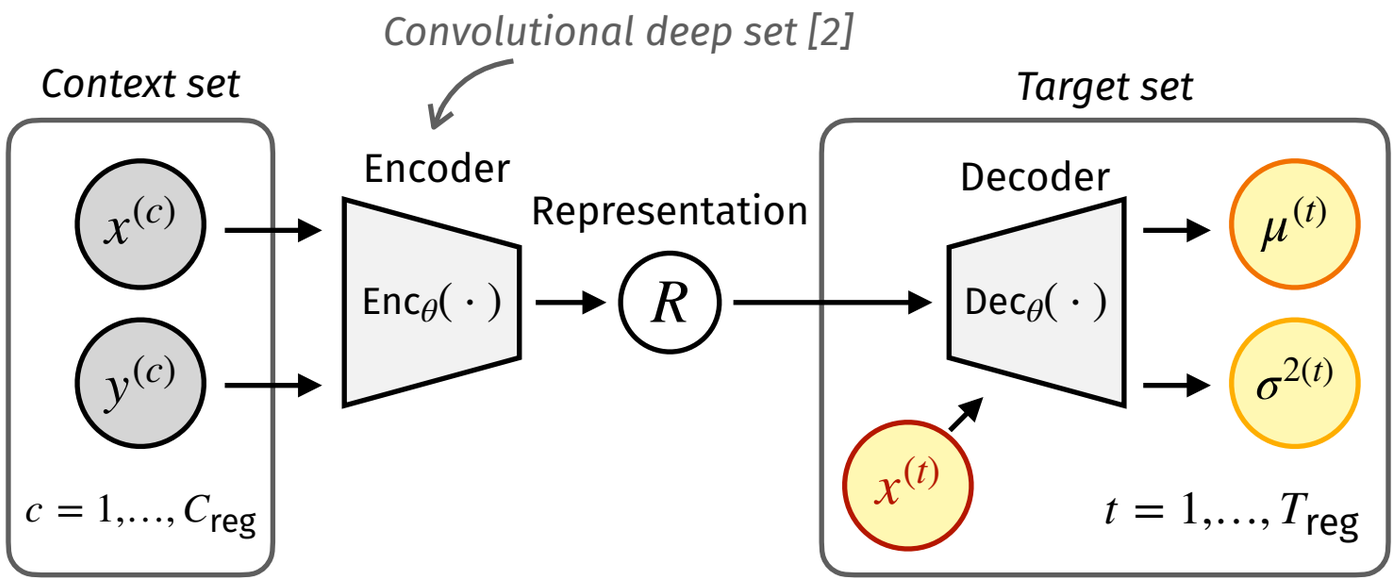
How do neural processes work?

● Context ✕ Target — Underlying function

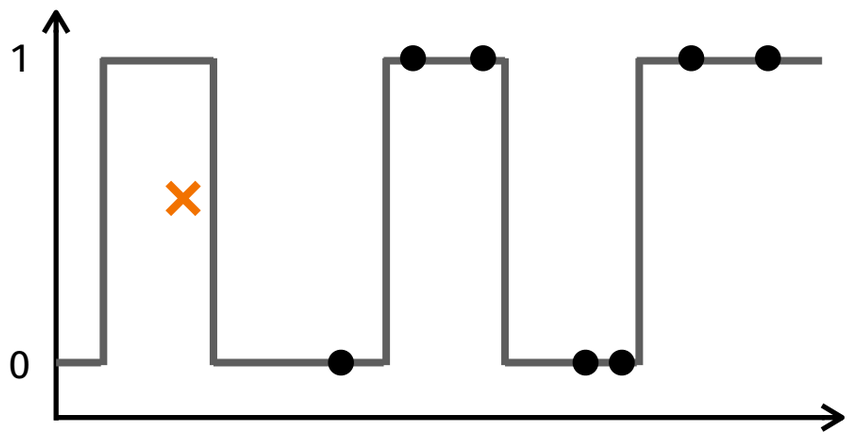
Output 1: Regression



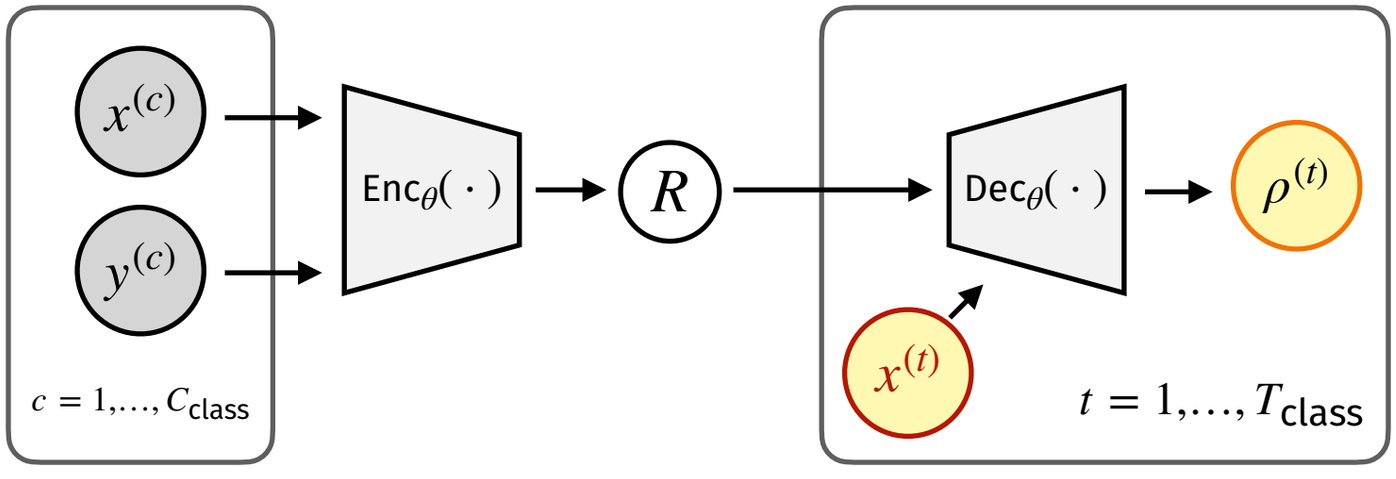
Convolutional deep set [2]



Output 2: Classification

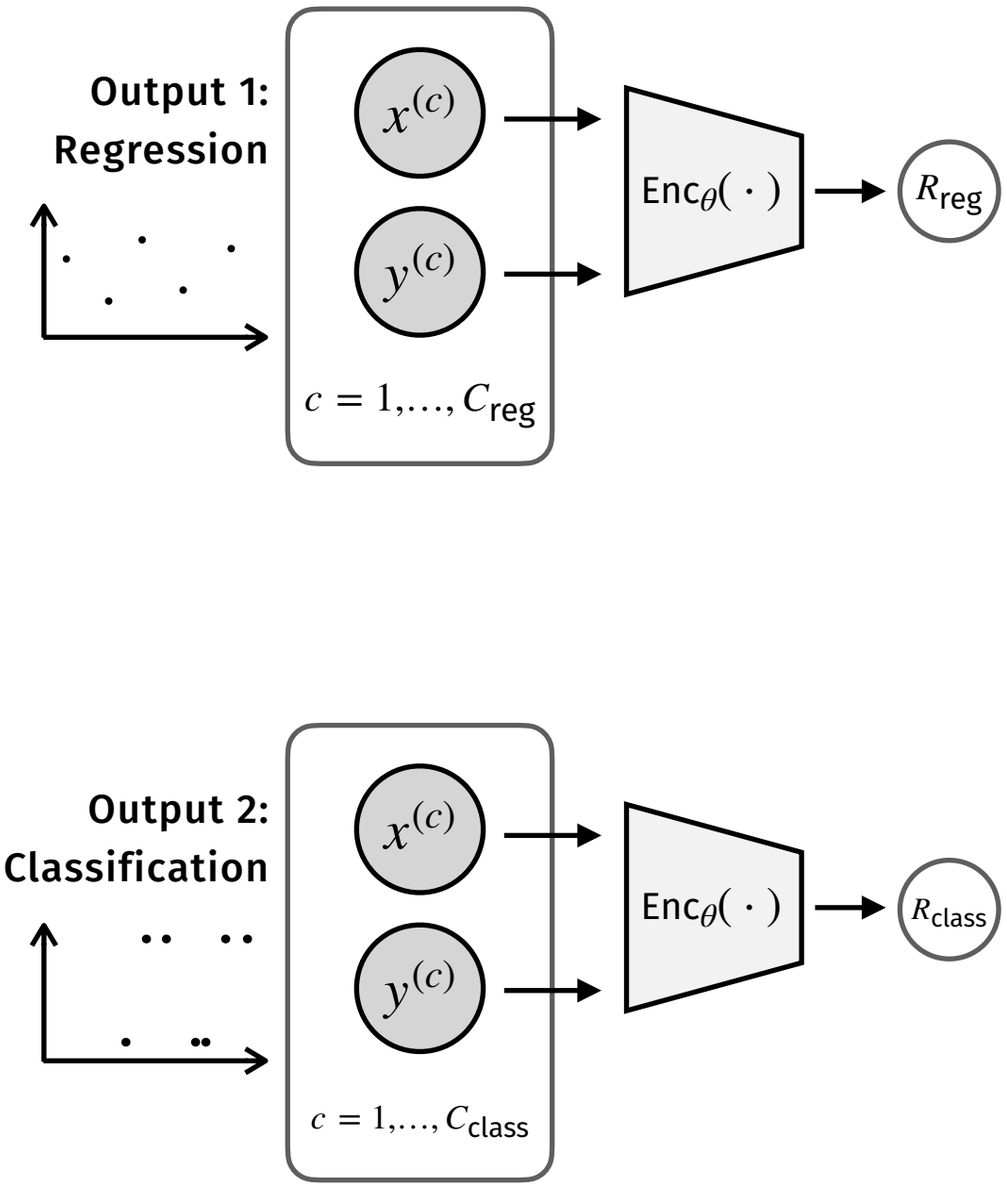


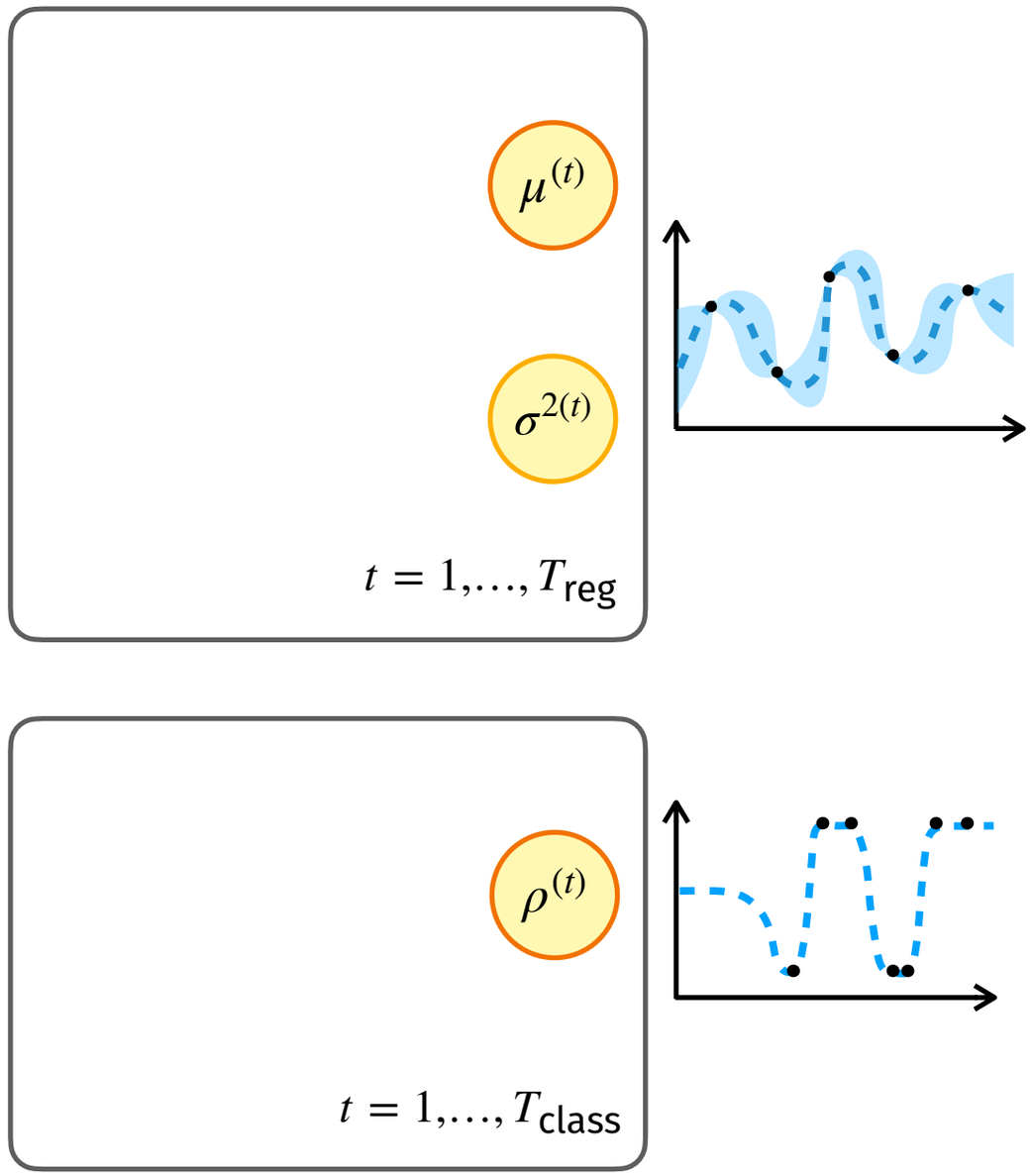
How can we connect these architectures so that our predictions are informed by both outputs?



How can they handle multi-output data?

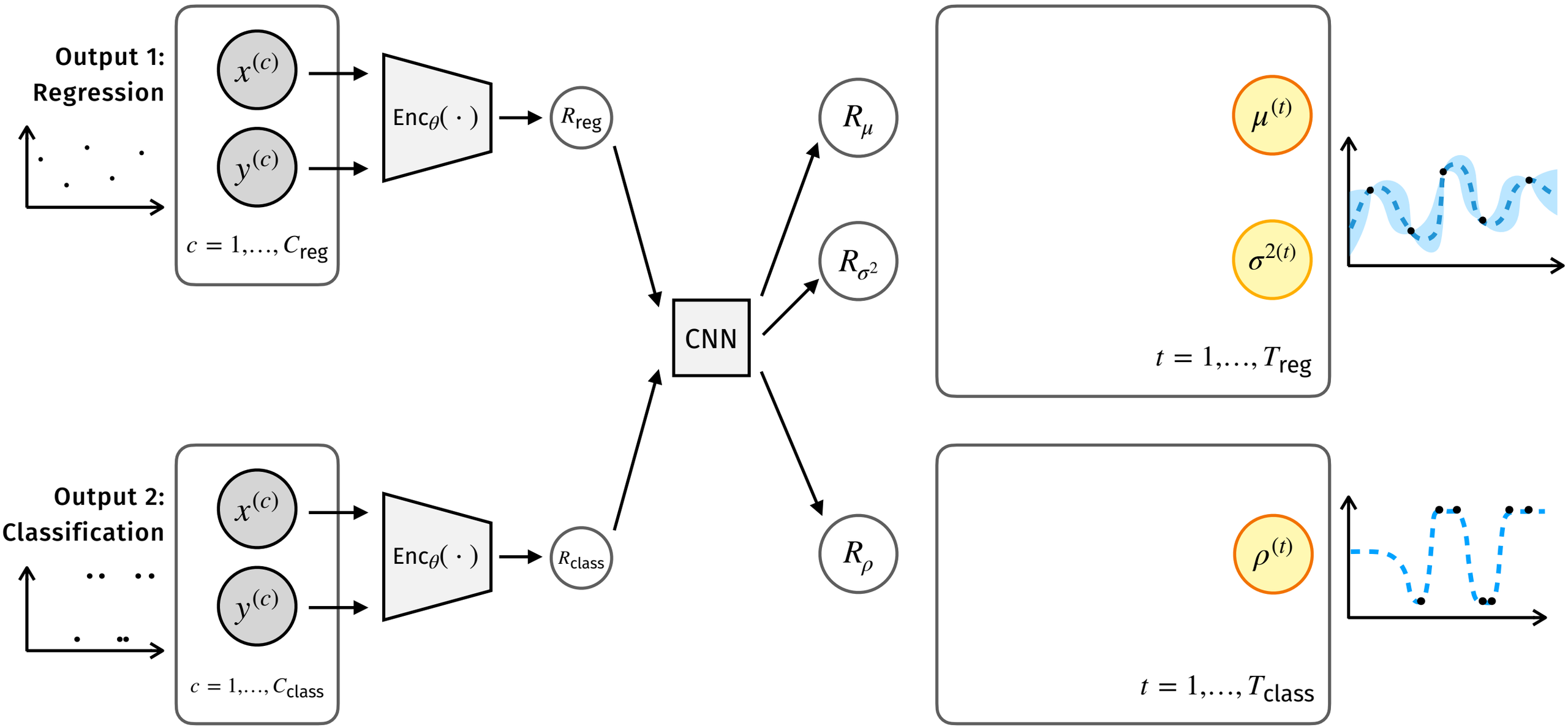
Dual-output architecture





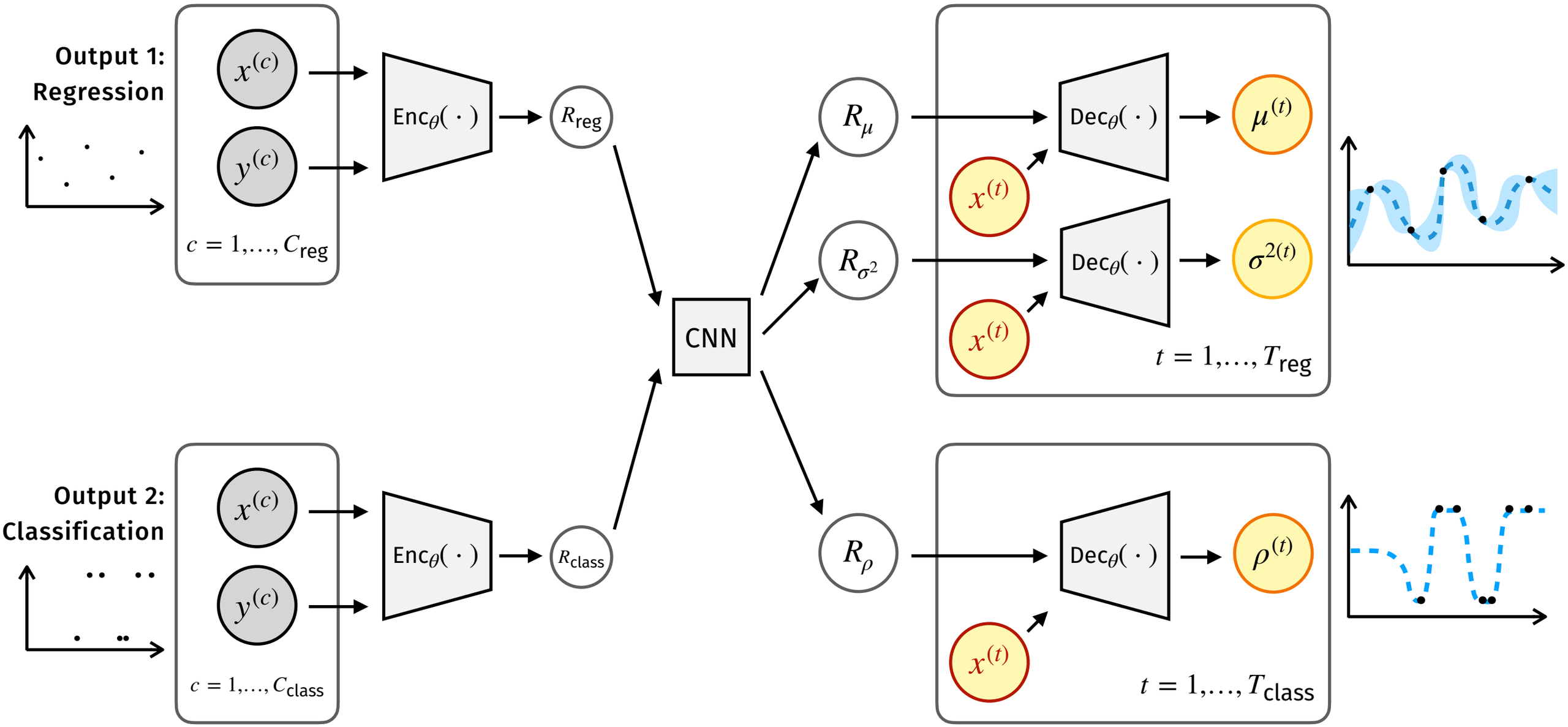
How can they handle multi-output data?

Dual-output architecture



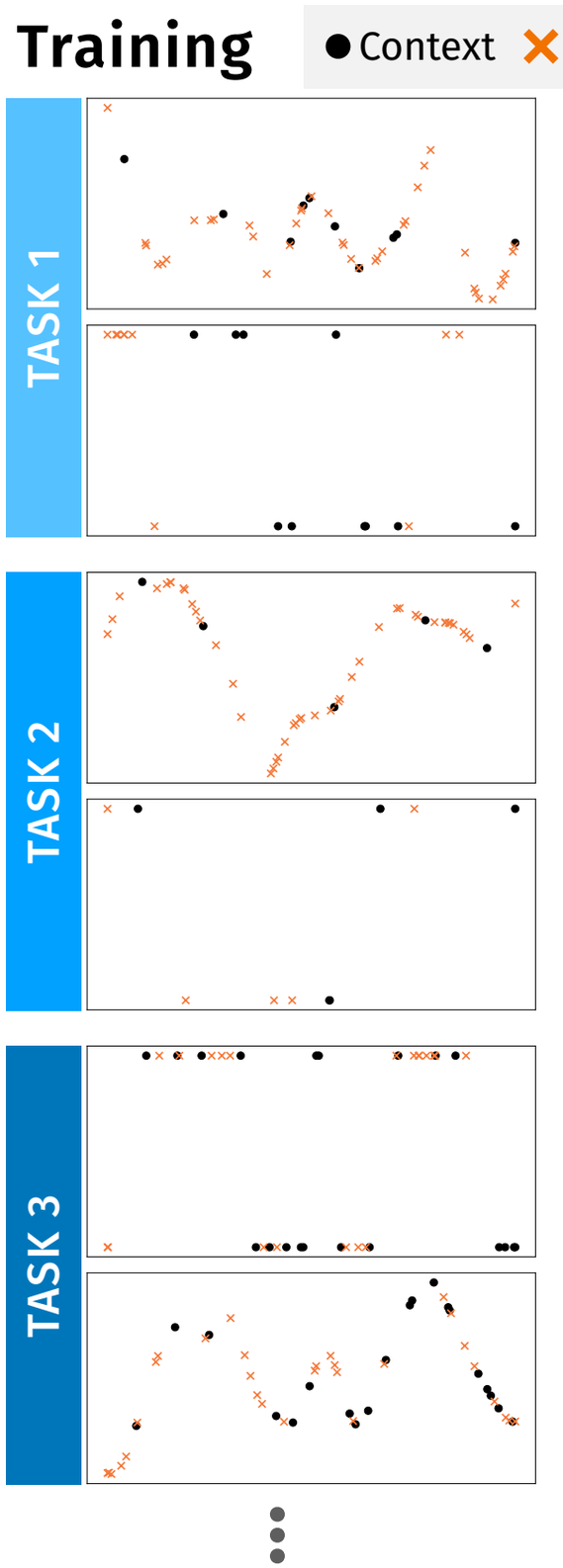
How can they handle multi-output data?

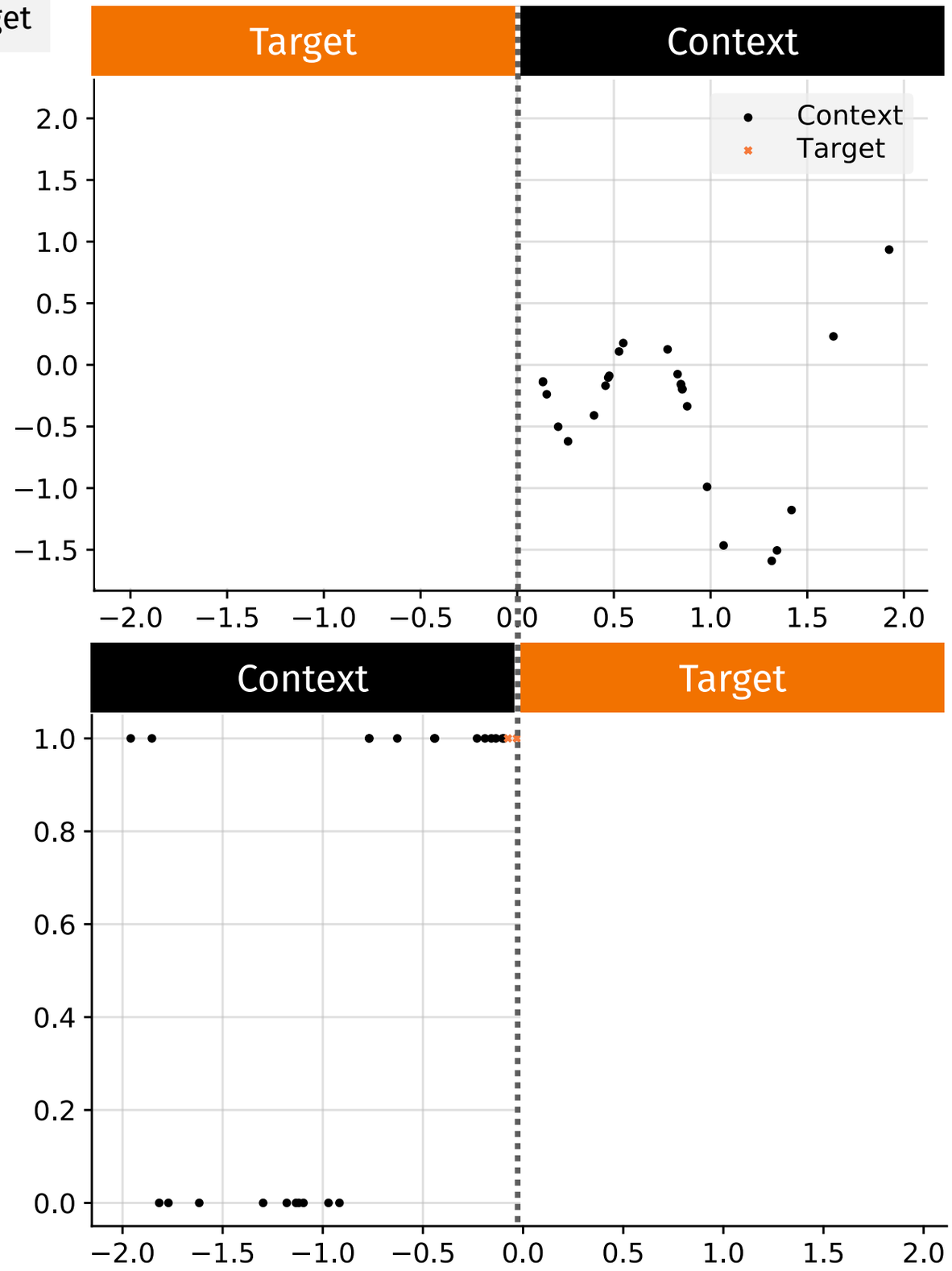
Dual-output architecture



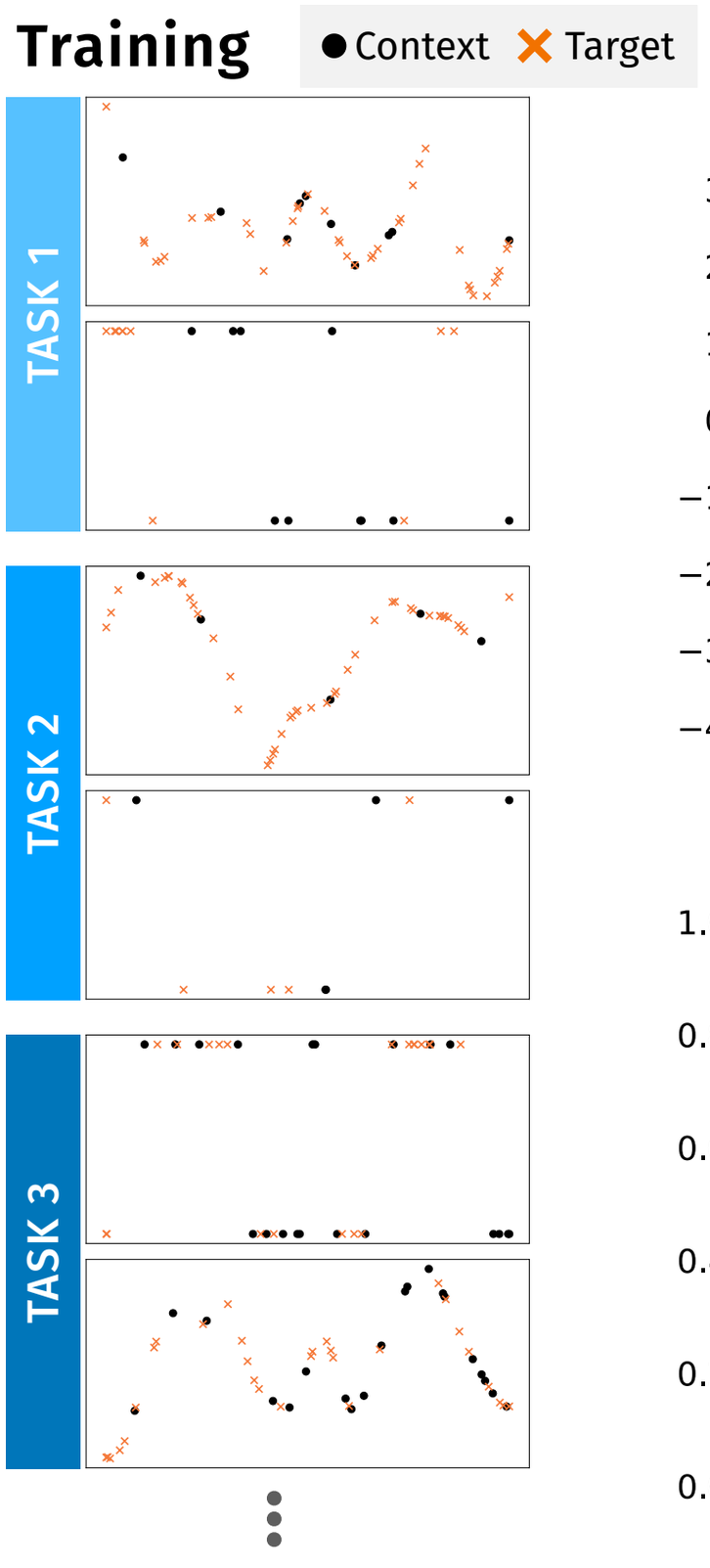
How does the dual architecture perform?

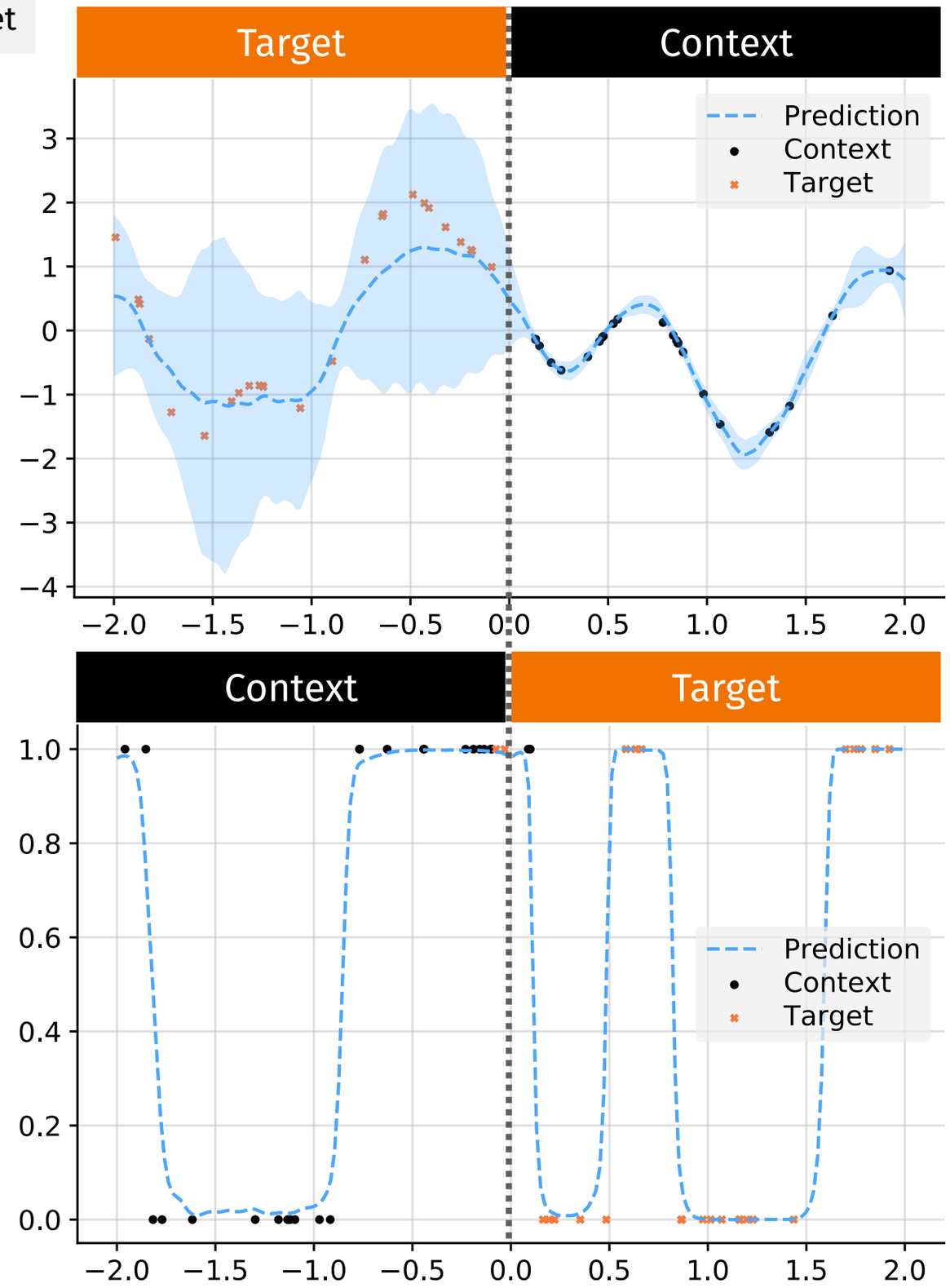
How does the dual architecture perform?





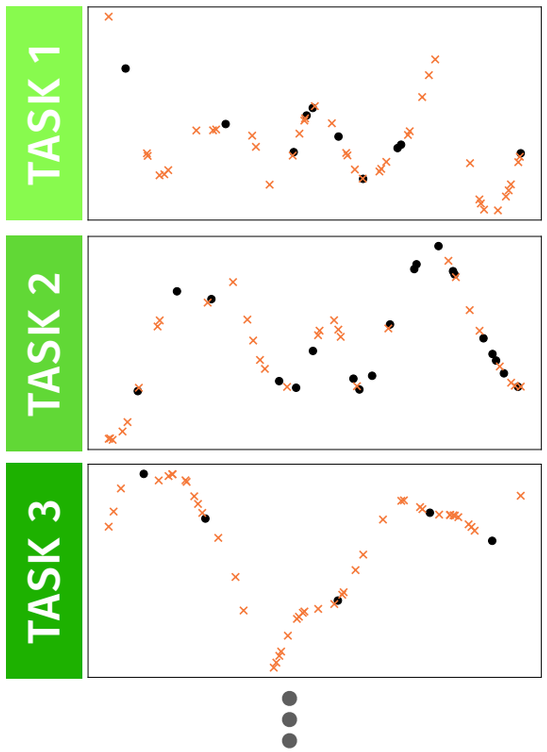
How does the dual architecture perform?

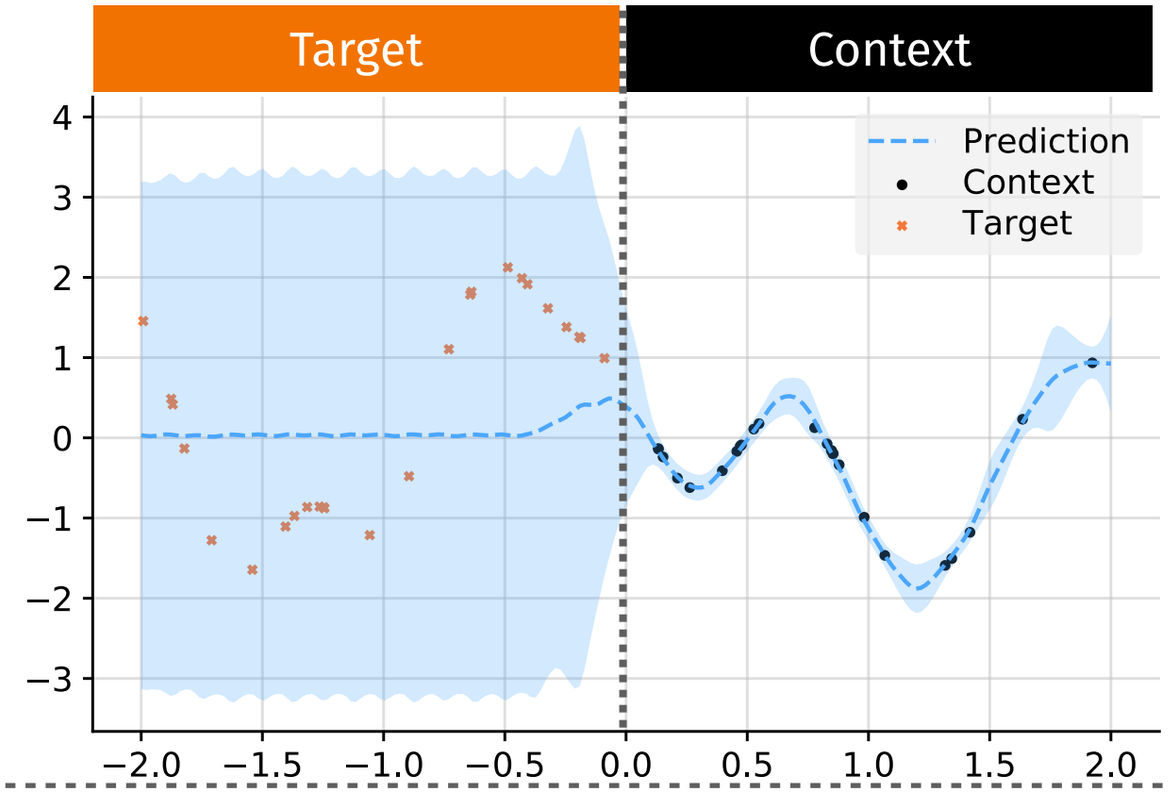




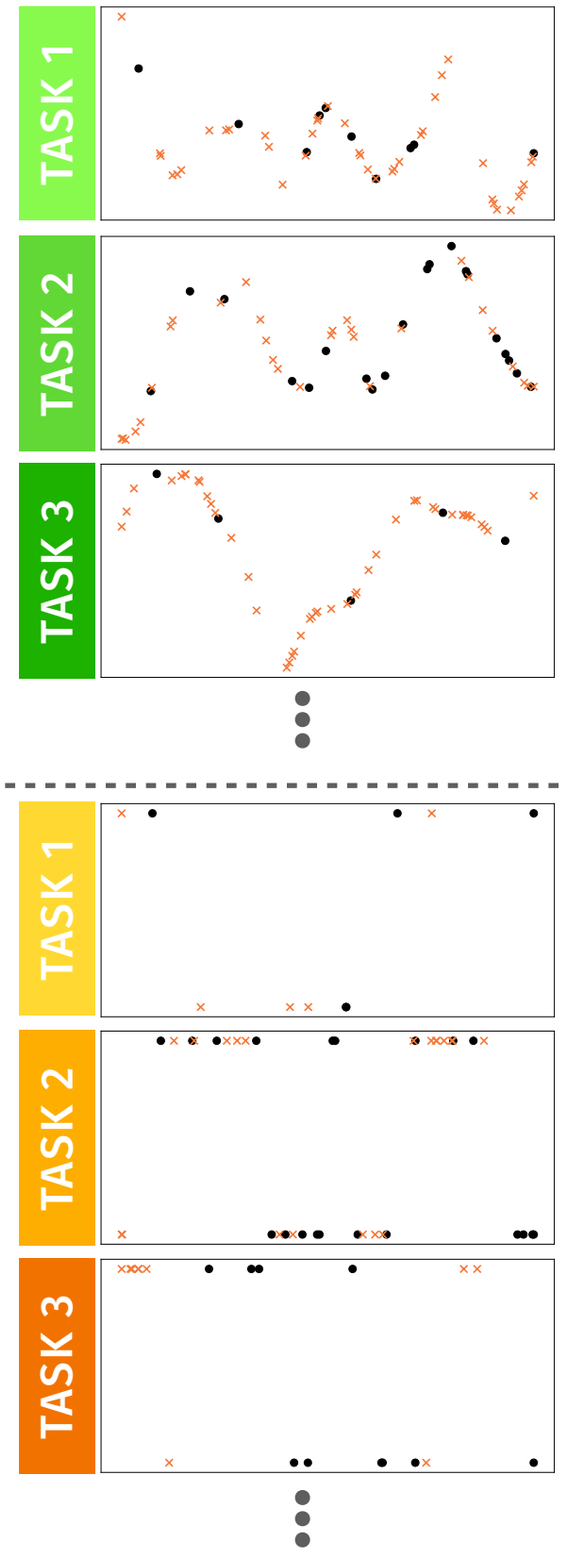
How does the single architecture perform?

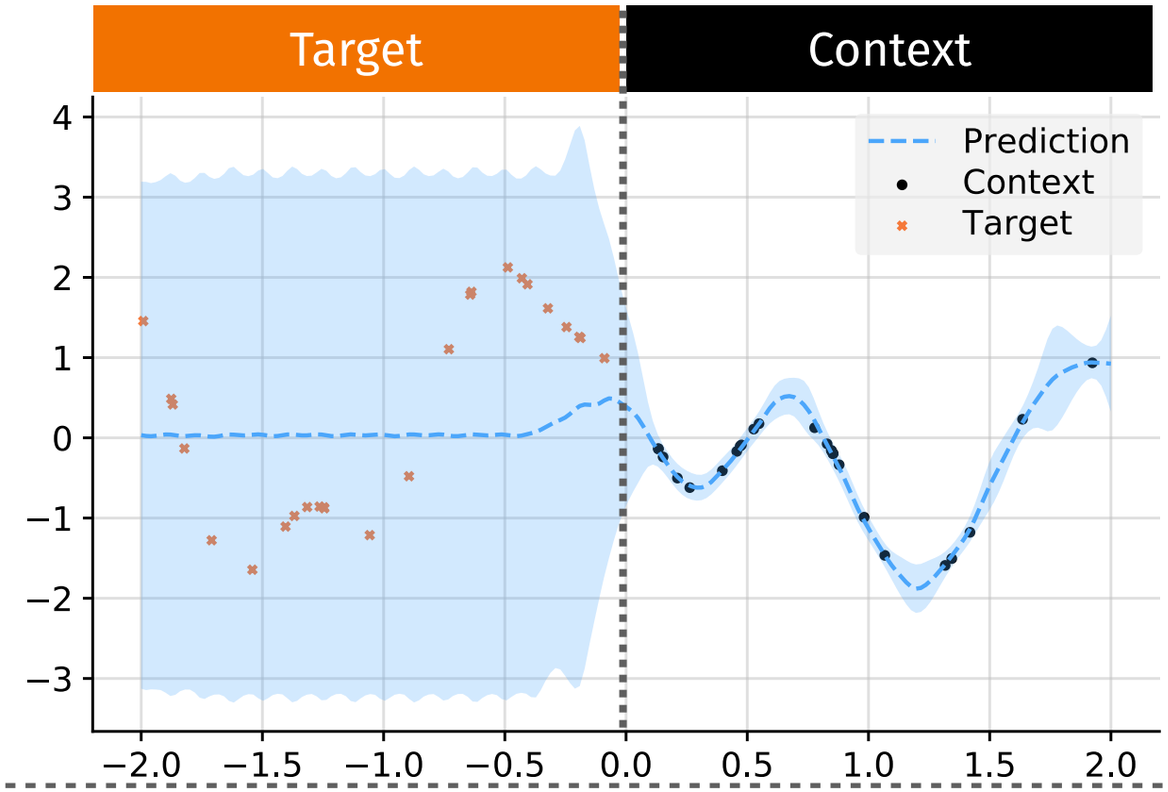
How does the single architecture perform?

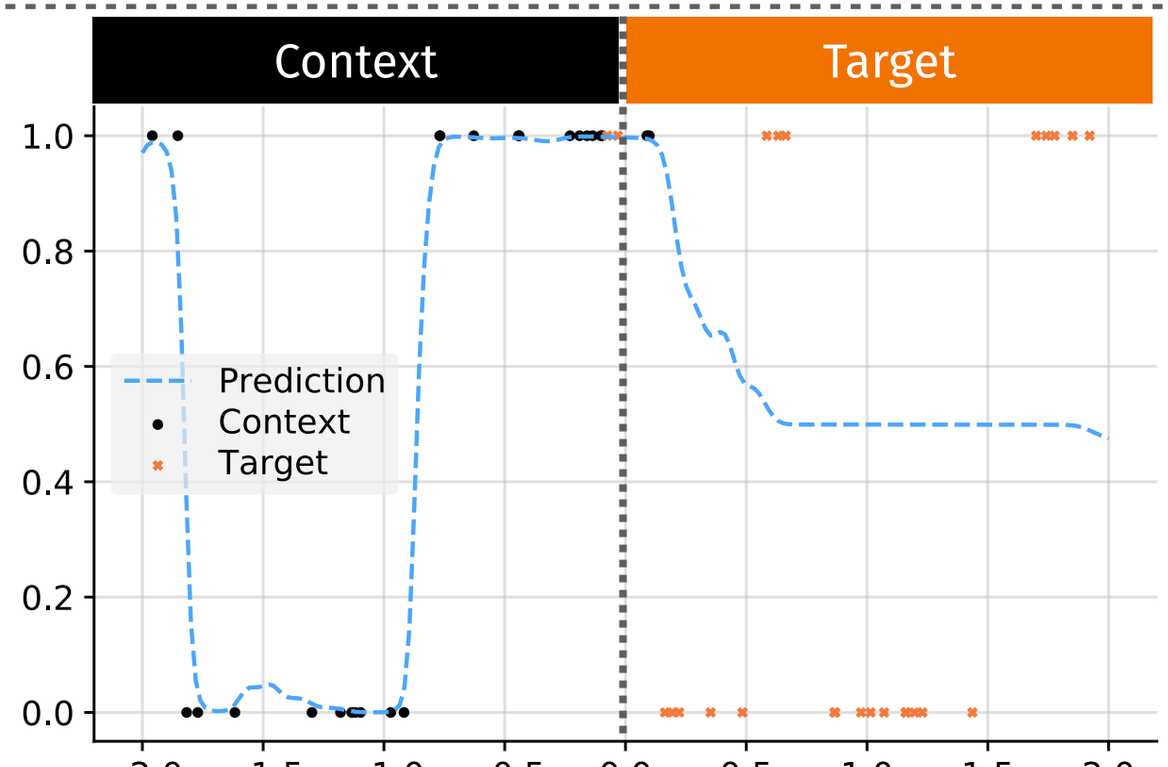




How does the single architecture perform?







How does the dual architecture perform? 5/7

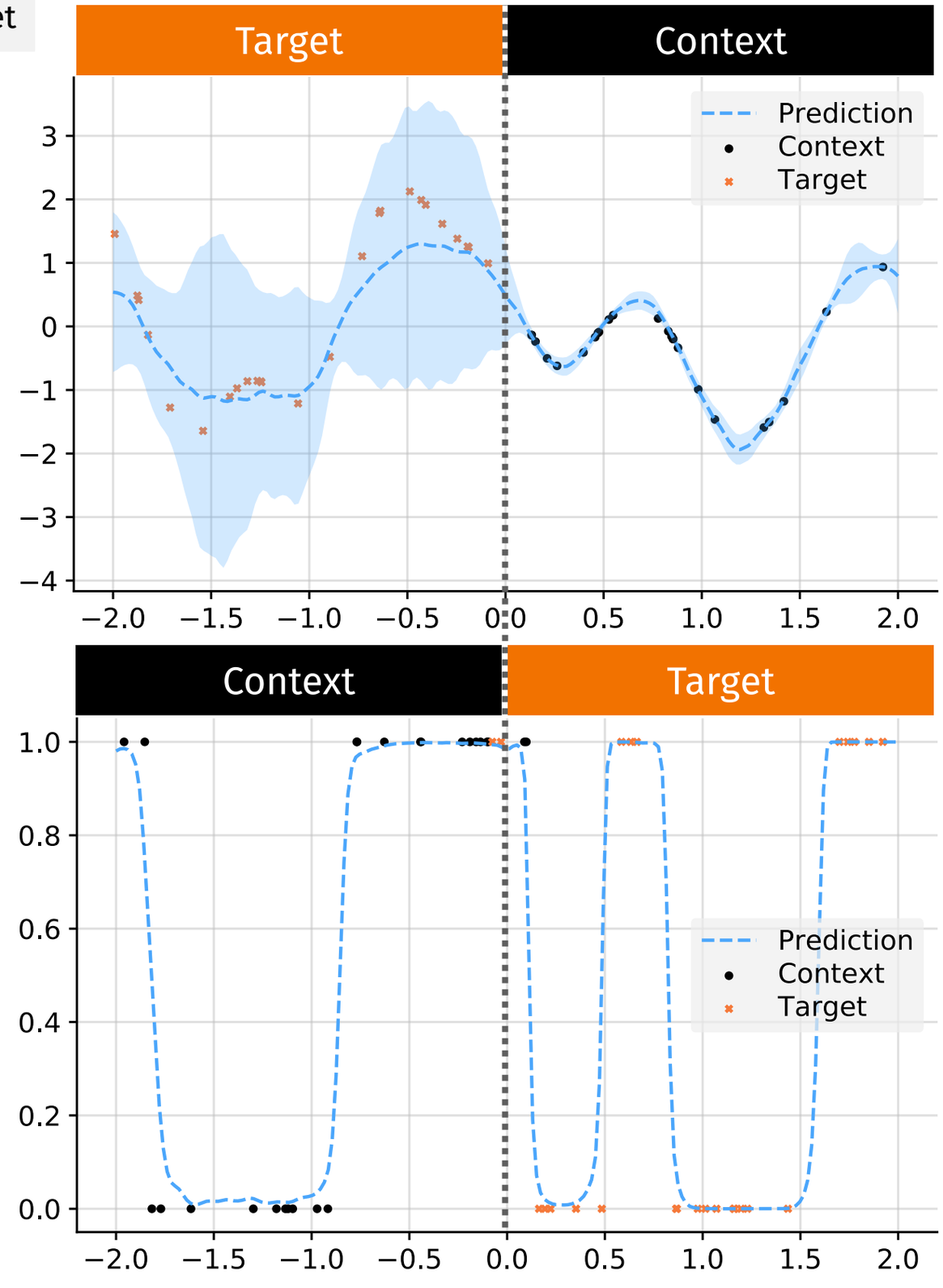
Training

● Context ✕ Target

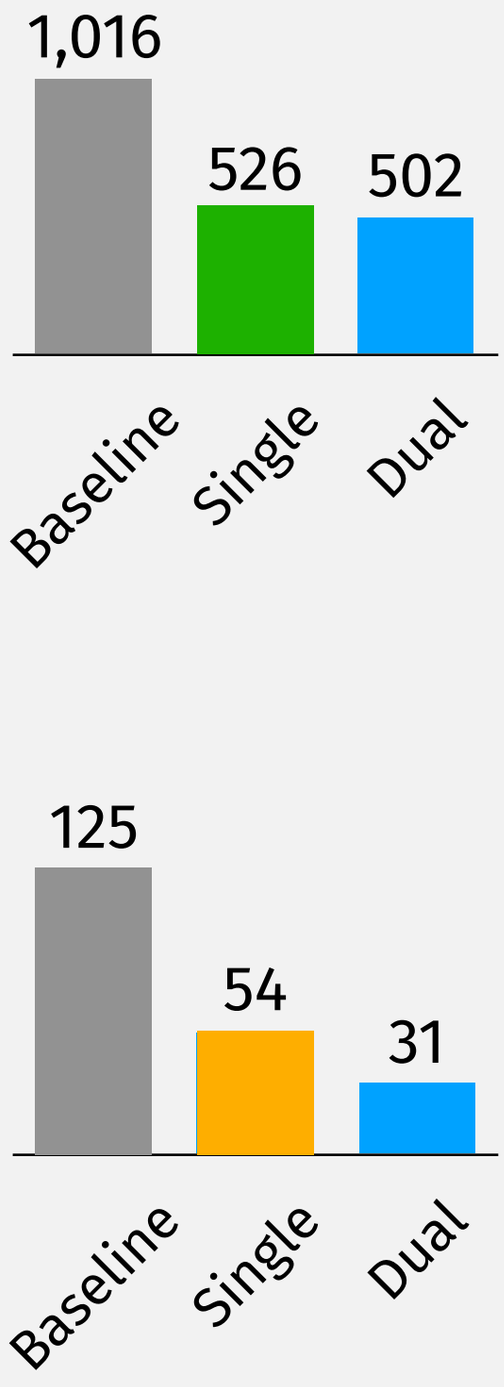
TASK 1

TASK 2

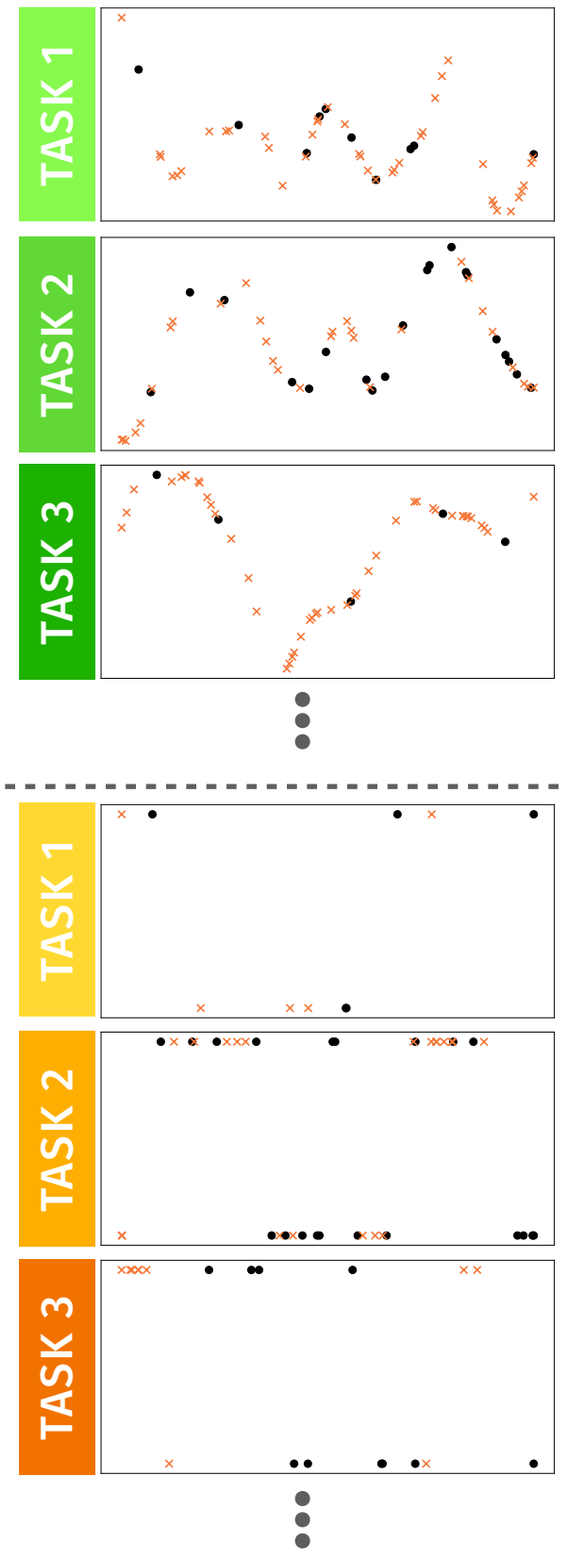
TASK 3

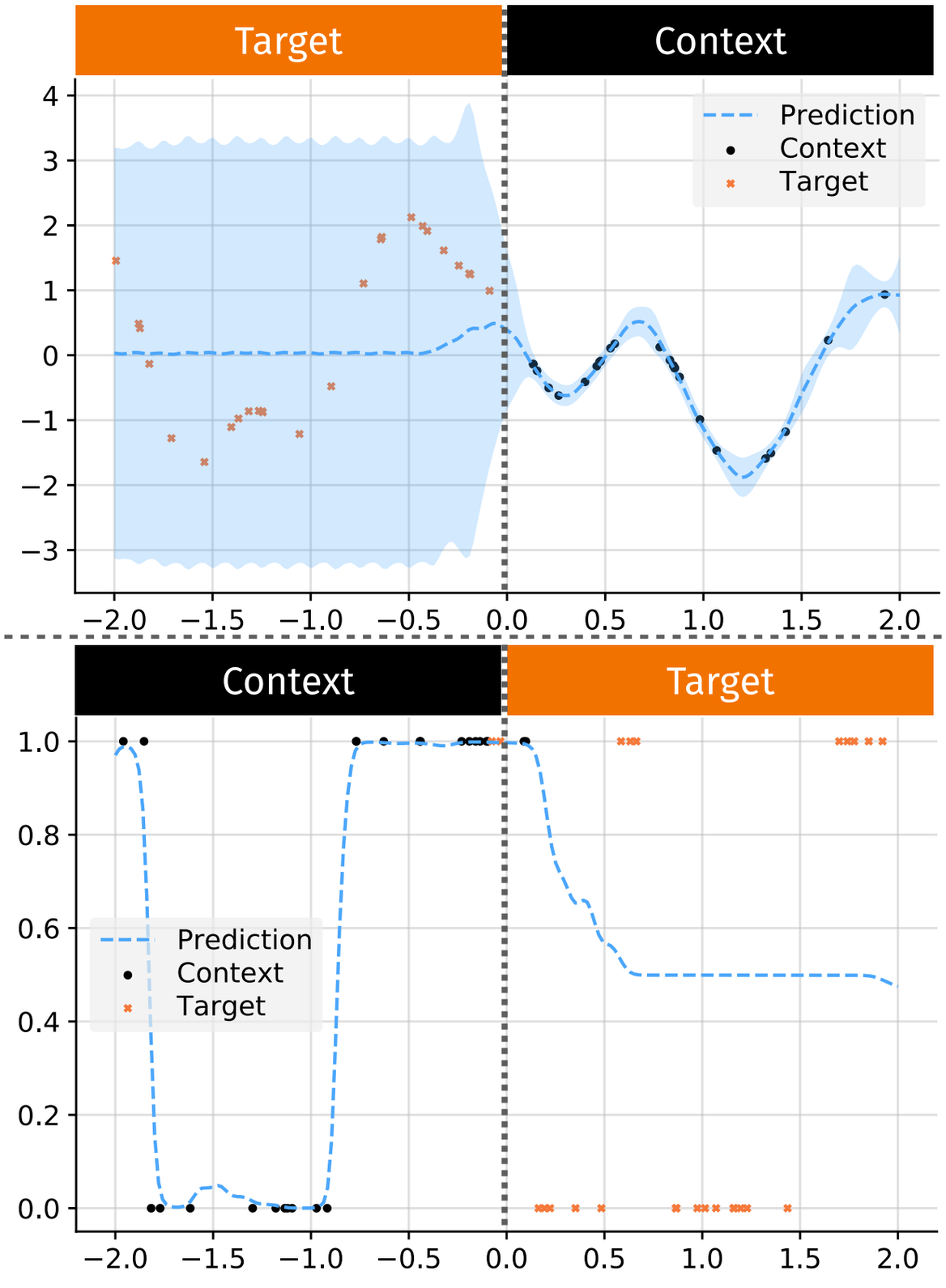


Losses

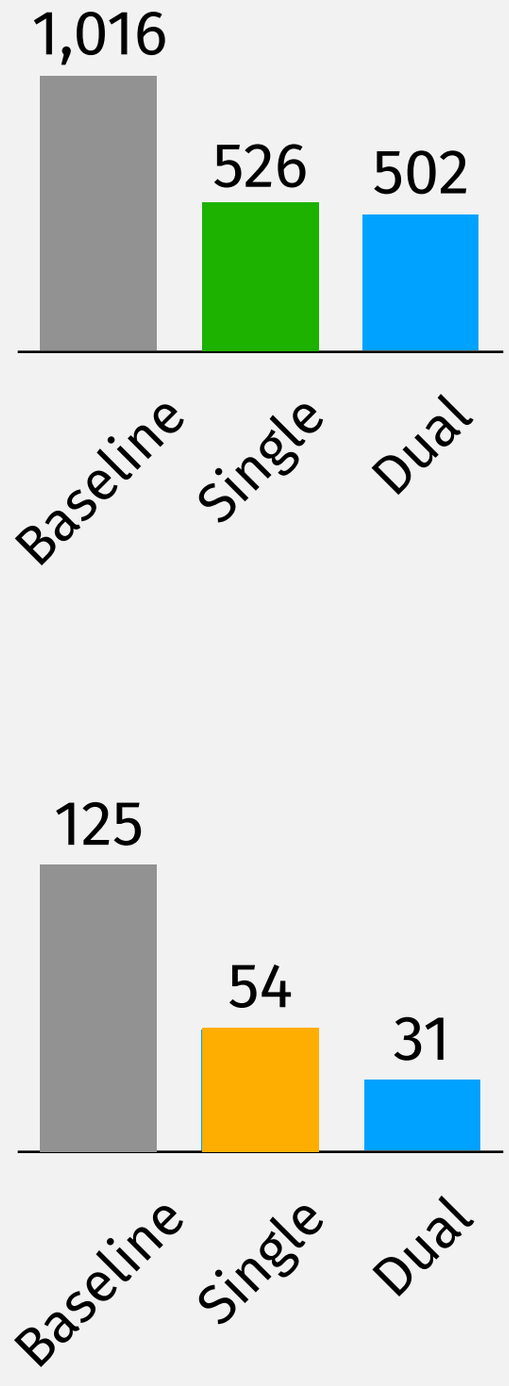


How does the single architecture perform? 6/7

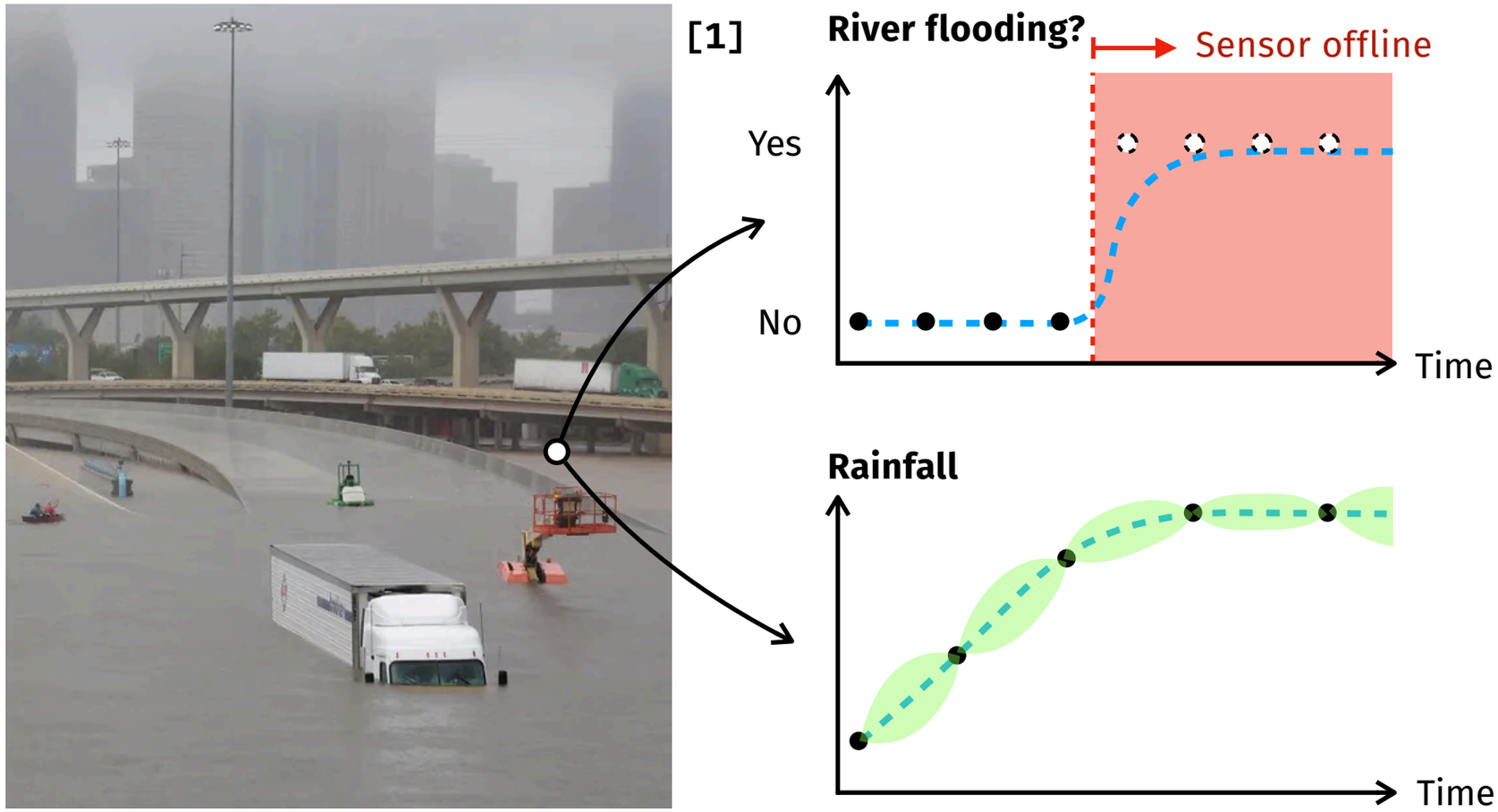




Losses



Conclusion



PLAN

Summer

- ☑ Understand **literature** on neural processes
- ☑ Build **dual architecture** from provided codebase

Term 1

- ☑ Train and test on **synthetic data**

Conclusion

PLAN

Summer

- [x] Understand **literature** on neural processes
- [x] Build **dual architecture** from provided codebase

Term 1

- [x] Train and test on **synthetic data**
- [ ] Deploy on **real-world dataset** from environmental science

Term 2

- [ ] Explore different **computational architectures**

References

1. Livsey, Anna. ‘Before and after: Images Show How Hurricane Harvey Swamped Houston’. *The Guardian*, 29 Aug. 2017. *The Guardian*, https://www.theguardian.com/us-news/2017/aug/29/before-and-after-images-show-how-hurricane-harvey-swamped-houston.

2. Gordon, Jonathan, et al. *Convolutional Conditional Neural Processes*. 2019. openreview.net, https://openreview.net/forum?id=Skey4eBYPS.

3. Dubois, Yann, et al. *Neural Process Family*. 2020, https://yanndubs.github.io/Neural-Process-Family/.