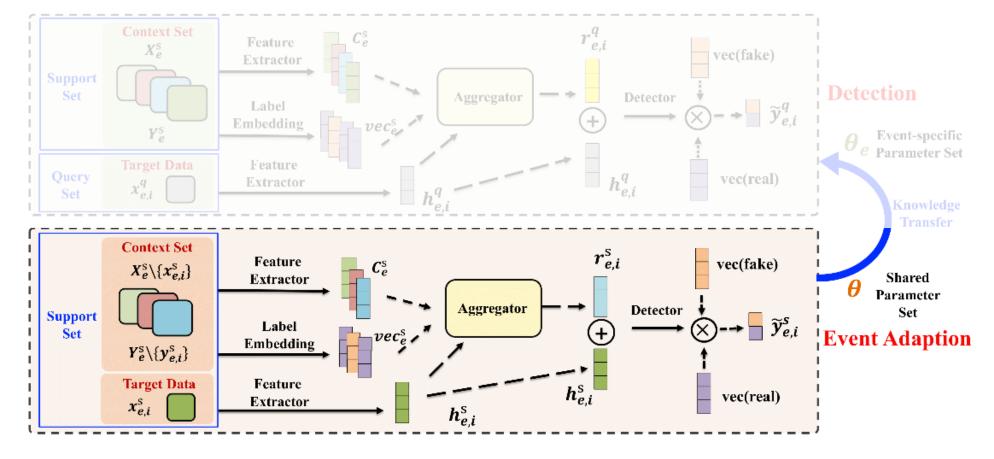
## Methodology

## Event adaption stage



- The loss can be calculated between the prediction  $\hat{y}_{e,i}^s$  and the corresponding label  $y_{e,i}^s$ .
- $\theta$ : all parameters included in the proposed model.
- The event adaption objective function on the support set can be represented as

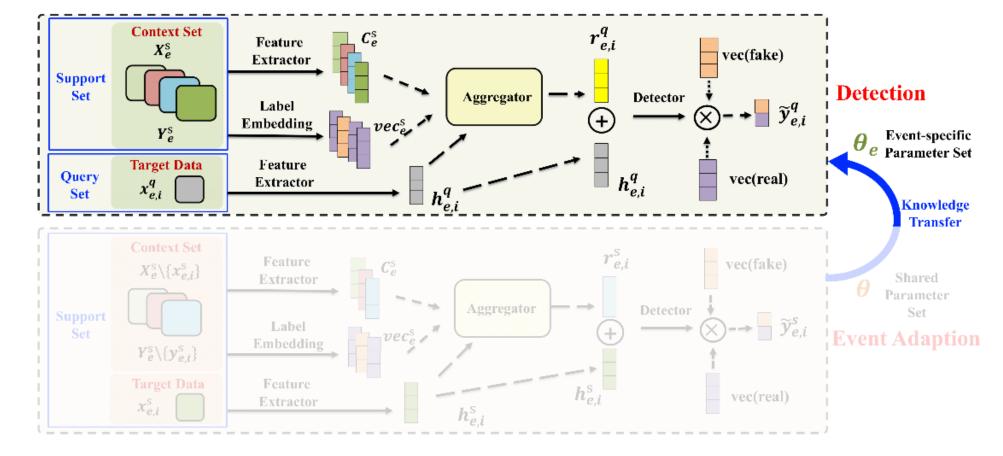
$$\mathscr{L}_{e}^{s} = \sum_{i} \log p_{\theta} \left( y_{e,i}^{s} \mid \left\{ \mathbf{X}_{e}^{s}, Y_{e}^{s} \right\} \backslash \left\{ x_{e,i}^{s}, y_{e,i}^{s} \right\}, x_{e,i}^{s} \right)$$

• Then update parameters heta one gradient descent updates on  $\mathscr{L}_e^s$  for event e.

• 
$$\theta_e = \theta - \alpha \nabla_{\theta} \mathcal{L}_e^s$$

## Methodology

## **Detection stage**



- The proposed model with event-specific parameter set  $\theta_e$  takes query set  $\mathbf{X}_e^q$  and entire support set  $\{\mathbf{X}_e^s, \mathbf{Y}_e^s\}$  as input and outputs predictions  $\mathbf{Y}_e^q$  for query set  $\mathbf{X}_e^q$ .
- The loss function in the detection stage can be represented as

• 
$$\mathcal{L}_e^q = \log p_{\theta_e} \left( Y_e^q \mid X_e^s, Y_e^s, X_e^q \right)$$

• Through this meta neural process, we can learn an initialization parameter set  $\theta$  which can rapidly learn to use given context input-outputs as conditioning to detect fake news on newly arrived events.