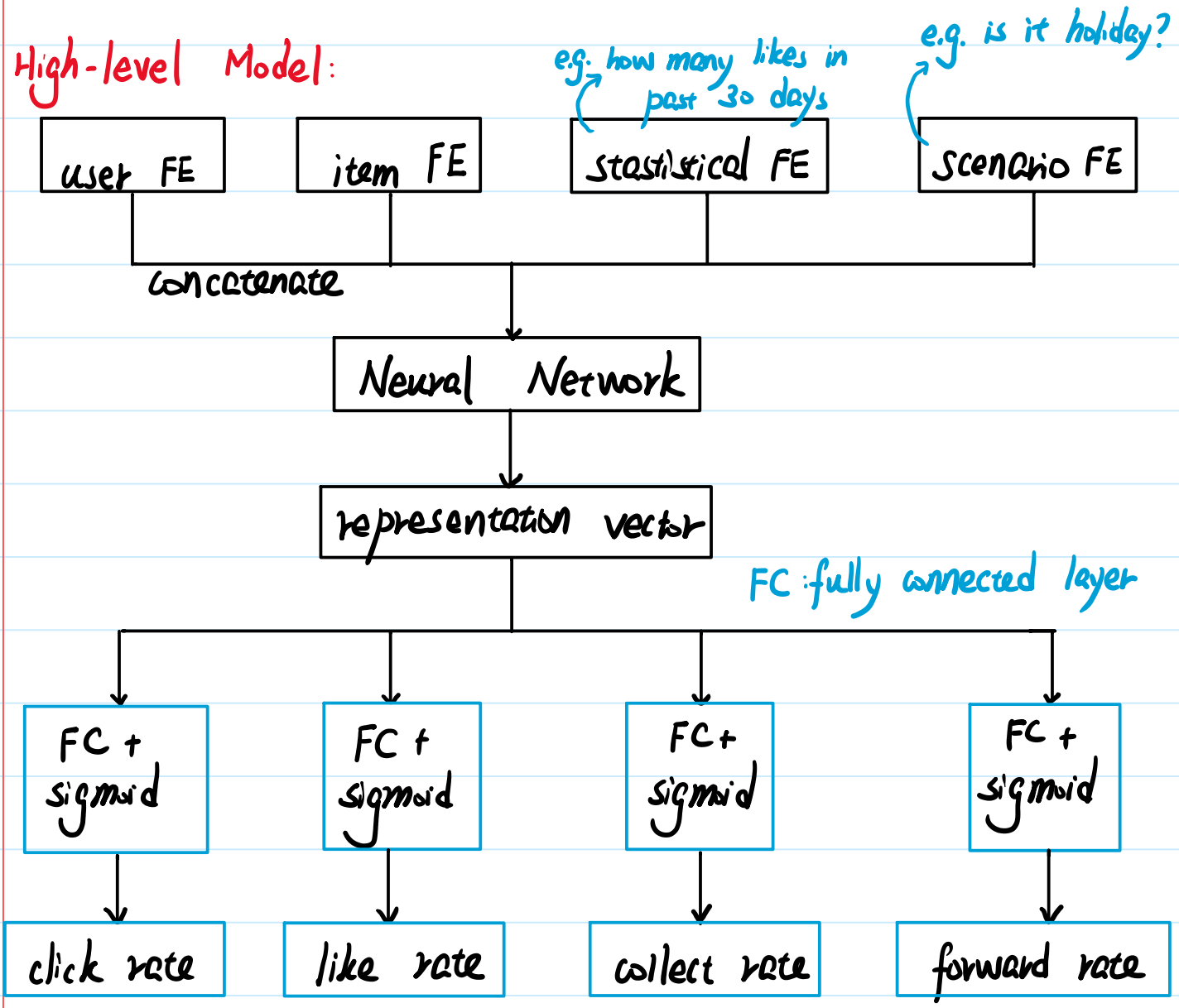
Multi-objective ranking model

Friday, March 29, 2024 9:35 AM

Ranking:

rank items generated from candidate retrieval.

High-level Model:



Multi-objective ranking model

Friday, March 29, 2024 9:35 AM

Model Training:

	click rate	like rate	collect rate	forward rate
estimate:	p_1	p_2	p_3	p_4
	↕			
target:	y_1	y_2	y_3	y_4

loss function: $\sum_{i=1}^{4} \alpha_i \, CE(y_i, p_i)$

$$CE(y_i, p_i) = \text{Cross Entropy}(y_i, p_i)$$

$$= -\left[y_i \ln p_i + (1 - y_i) \cdot \ln(1 - p_i) \right]$$

each prediction is a "binary classification"

① whether click? ② whether like? ③ whether collect?

④ whether forward/share?

Too many negative samples:

downsample the negative samples.

make balance between positive and negative samples.

Multi-objective ranking model

Friday, March 29, 2024 9:35 AM

Calibrate Estimate:

① # of positive and negative samples: n_+ n_-

② down sample negative samples: $\alpha \cdot n_-$ $\alpha \in [0, 1]$

③ overestimate click rate because $\alpha n_- << n_-$

$$p_{true} = \frac{n_+}{n_+ + n_-} \text{①} \qquad p_{pred} = \frac{n_+}{n_+ + \alpha \cdot n_-} \text{②}$$

From ①: $n_+ = \frac{p_{true}\, n_-}{1 - p_{true}}$ plug into ②

$$p_{true} = \frac{\alpha \cdot p_{pred}}{(1 - p_{pred}) + \alpha p_{pred}}$$

① use model to obtain p_{pred}

② then calibrate to obtain p_{true}

③ use p_{true} to rank items