FinSIR: Financial SIR-GCN for Stock Recommendation

Brian Godwin Lim
Division of Information Science
Nara Institute of Science and Technology

Background

- Efficient market hypothesis
  - Numerous studies demonstrate predictability of the stock market
    - Model stocks independently
    - Performance based on a few stocks

- Goal: Recommend top-K profitable stocks in a market
  - Maximize cumulative investment rate of return (IRR)

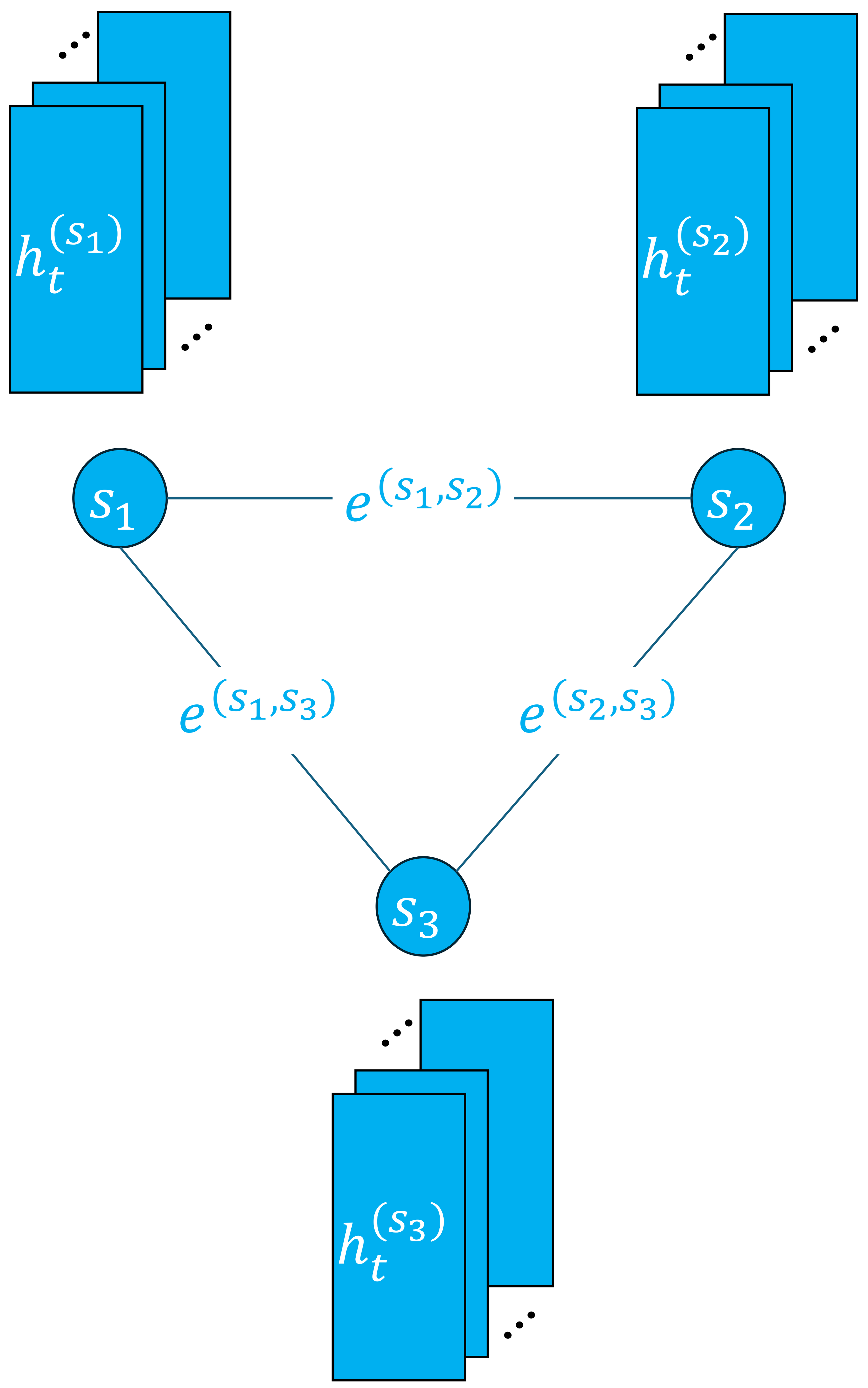
$$\mathrm{IRR}_K = \sum_{t=1}^{T} \frac{1}{K} \sum_{k=1}^{K} r_t^{\left(s_{\mathrm{argmax}_k(\hat{\boldsymbol{r}}_{\boldsymbol{t}})}\right)},$$

where $\mathrm{argmax}_k(\hat{\boldsymbol{r}}_{\boldsymbol{t}})$ represents the index of the kth largest value in $\hat{\boldsymbol{r}}_{\boldsymbol{t}}$ and $\boldsymbol{r}_{\boldsymbol{t}}$ $\hat{\boldsymbol{r}}_{\boldsymbol{t}}$ are the vector of 1-day actual and predicted returns, respectively

Data

- Historical price (node features)
  - Close, 5-, 10-, 20-, 30-day Moving Average

- Relational graph (edge features)
    - $e_i^{(s,s')}$ indicates the presence/absence of the ith relationship
  - Wiki-relation
  - Industry-relation
  - Price-correlation

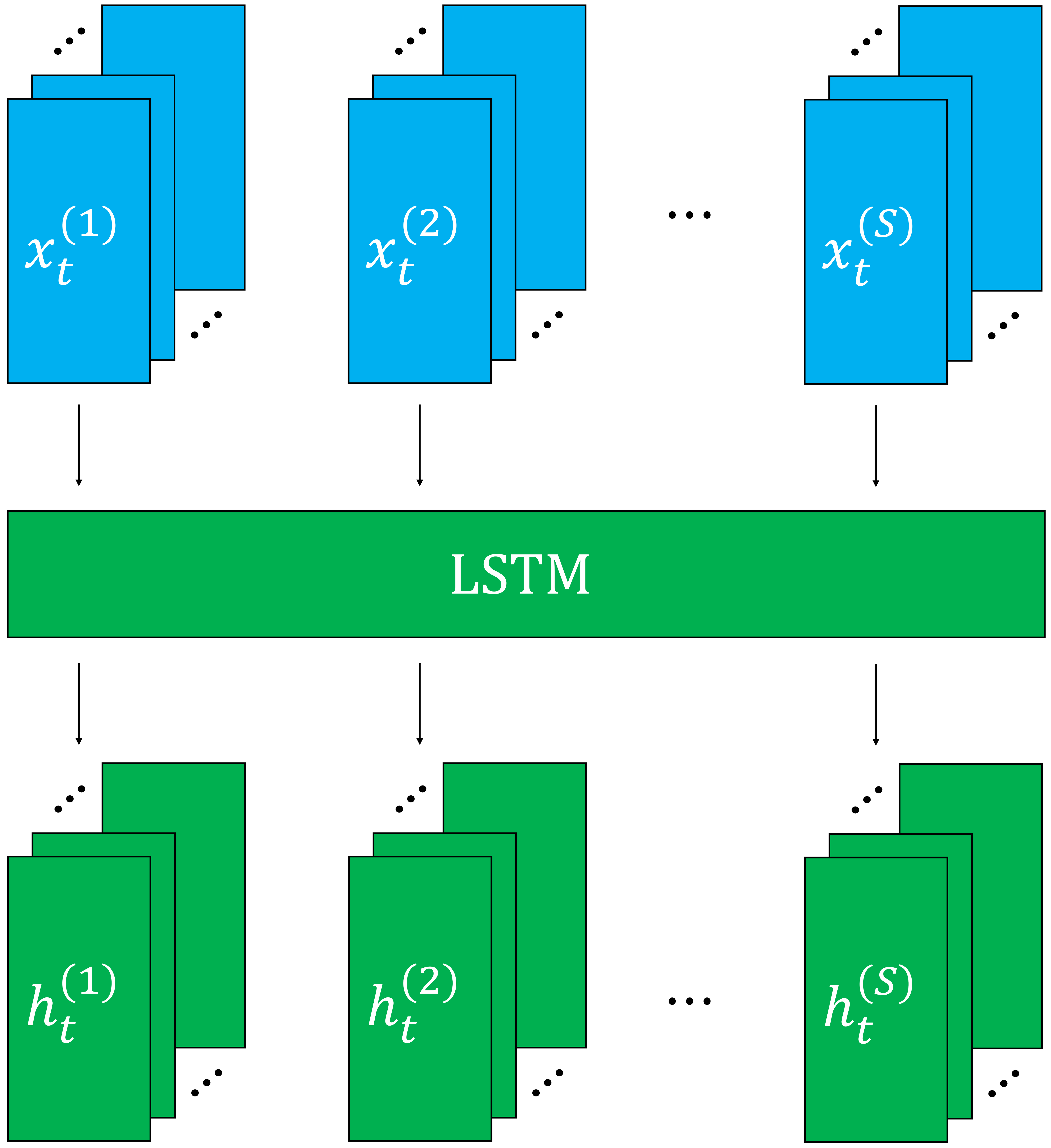


FinSIR Model

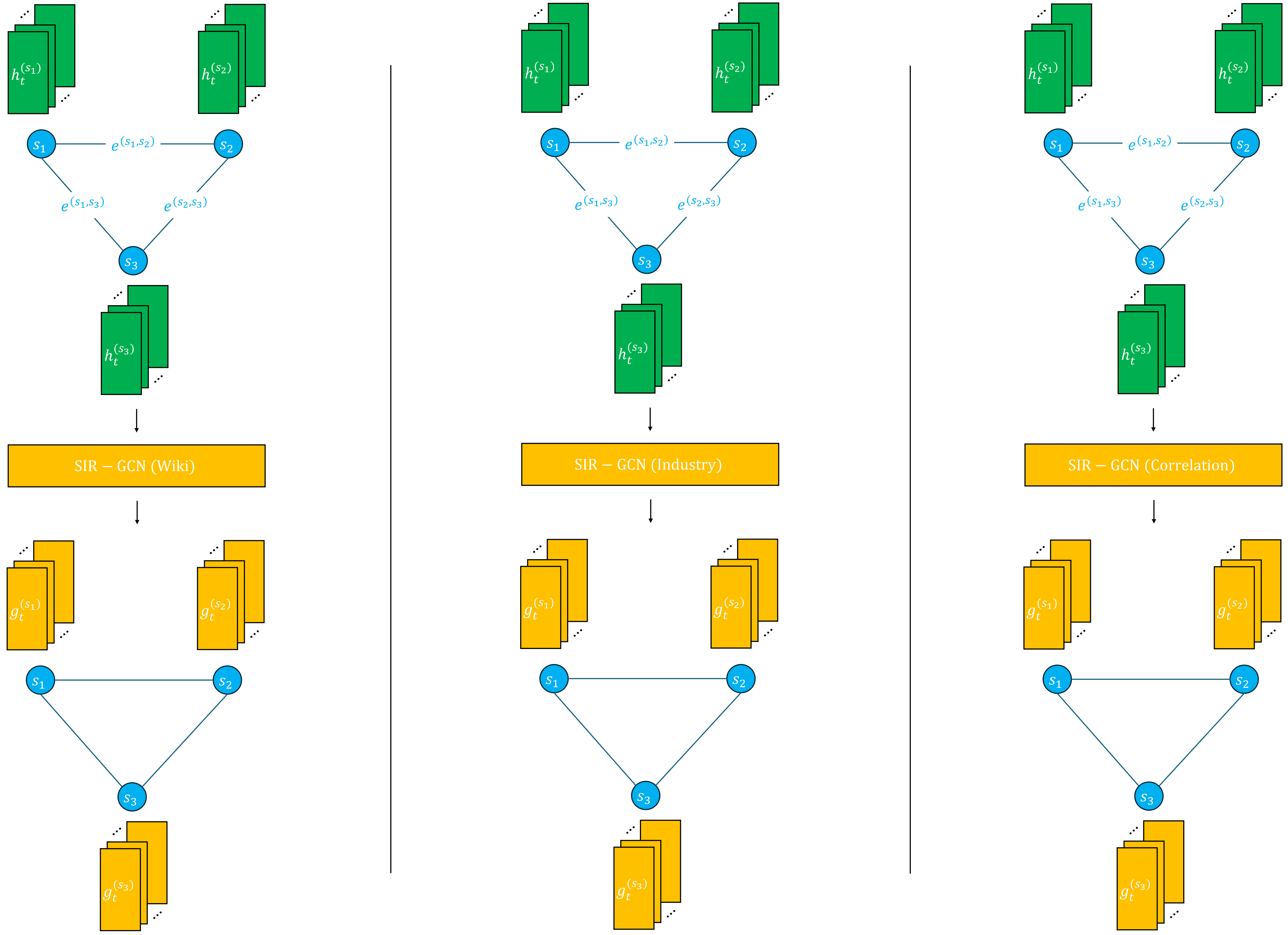
- Feature Pre-processing
  - LSTM (All)
- Relational Graph Convolution
  - SIR-GCN*
- Feature Post-processing
  - LSTM (All)
  - Temporal Attention
  - MLP

*Novelty: Explore the recently proposed SIR-GCN in the context of stock recommender systems

Feature Pre-processing



Relational Graph Convolution



SIR-GCN Variants

- LSTM

$$h_t^{(s,s')}, c_t^{(s,s')} = \mathrm{LSTM}\left(\left[h_t^{(s)}; h_t^{(s')}; e^{(s,s')}\right], h_{t-1}^{(s,s')}, c_{t-1}^{(s,s')}\right)$$

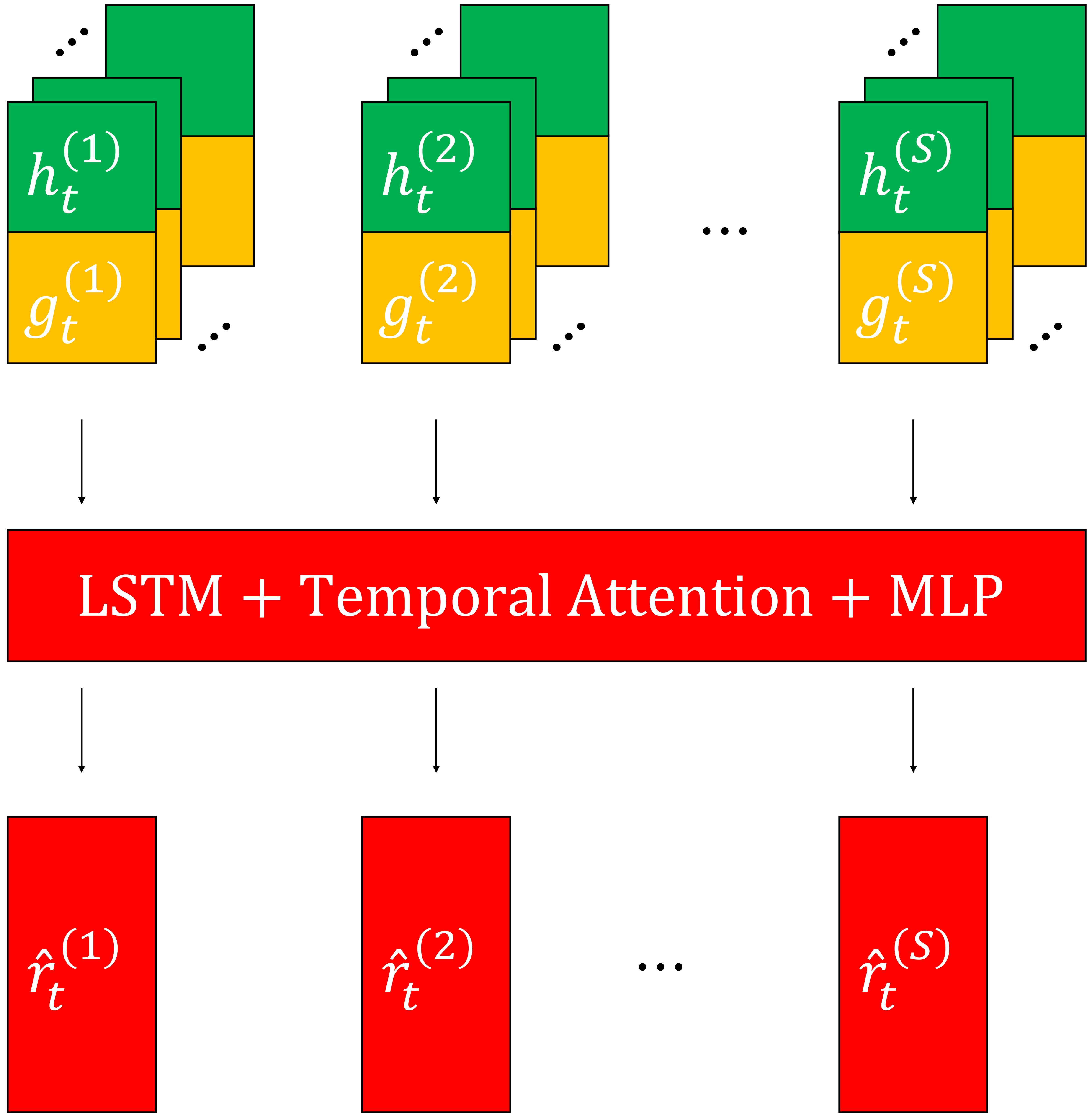
$$g_t^{(s)} = \bigoplus_{s' \in \mathcal{N}(s)} h_t^{(s,s')}$$

- Vanilla (original)

$$g_t^{(s)} = \bigoplus_{s' \in \mathcal{N}(s)} W_R\, \sigma\left(W_Q h_t^{(s)} + W_K h_t^{(s')} + W_E e^{(s,s')}\right)$$

$\bigoplus$ represents any aggregation function: sum, mean, max, symmetric mean, gated sum

Feature Post-processing



Assume $g_t^{(s)}$ is the concatenated features from the three graph convolutions

Simple FinSIR Model

- Feature Pre-processing
  - LSTM (Last)
- Relational Graph Convolution
  - SIR-GCN
- Feature Post-processing
  - MLP

Model Training

- Point-wise regression loss

$$l_{\mathrm{MSE}}(\boldsymbol{r_t}, \boldsymbol{\hat{r}_t}) = \frac{1}{S}\sum_{s=1}^{S}\left(r_t^{(s)} - \hat{r}_t^{(s)}\right)^2$$

- Pairwise rank-aware loss

$$l_{\mathrm{Rank}}(\boldsymbol{r_t}, \boldsymbol{\hat{r}_t}) = \sum_{s=1}^{S}\sum_{s'=1}^{S}\max\left(0, -\left(r_t^{(s)} - r_t^{(s')}\right)\left(\hat{r}_t^{(s)} - \hat{r}_t^{(s')}\right)\right)$$

- Training loss

$$l(\boldsymbol{r_t}, \boldsymbol{\hat{r}_t}) = l_{\mathrm{MSE}}(\boldsymbol{r_t}, \boldsymbol{\hat{r}_t}) + \alpha \cdot l_{\mathrm{Rank}}(\boldsymbol{r_t}, \boldsymbol{\hat{r}_t})$$