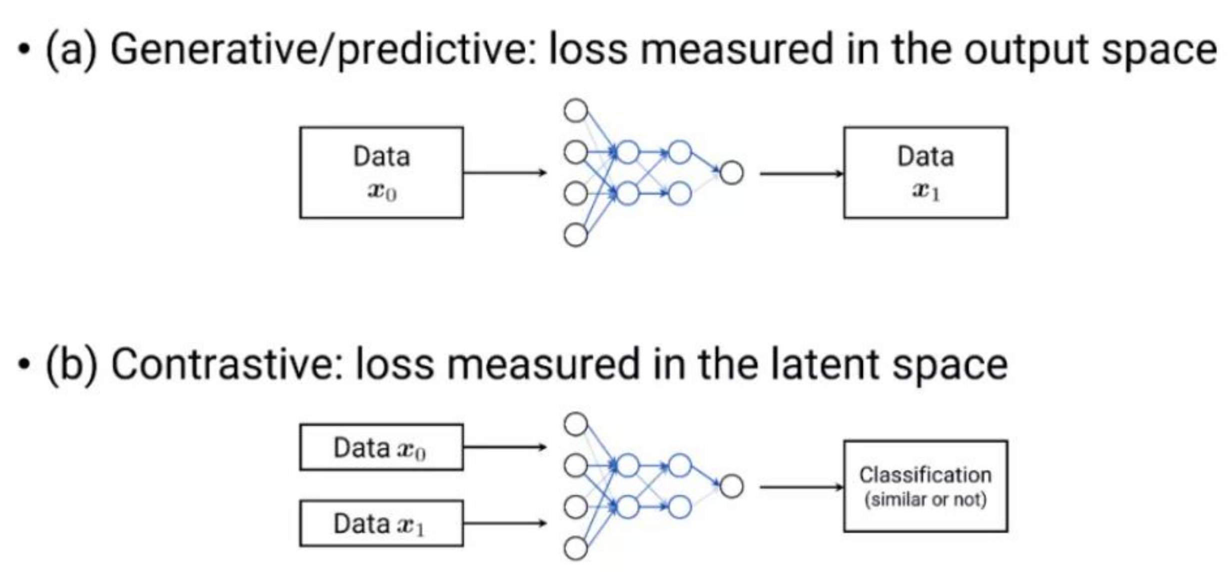
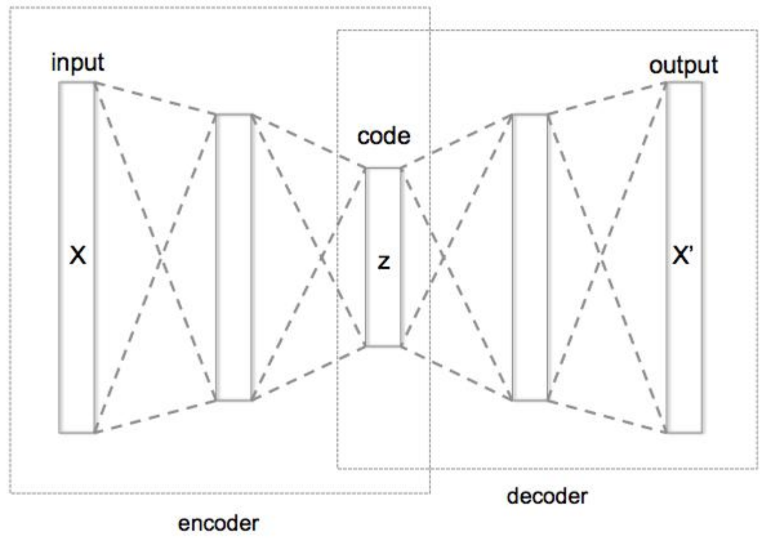
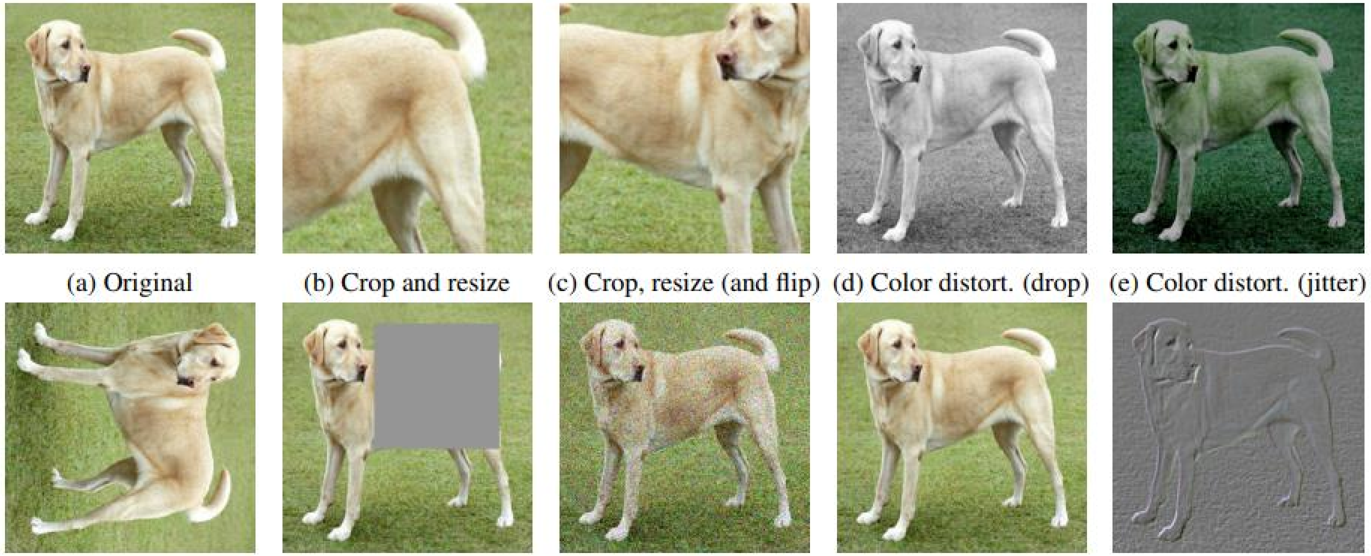
Graph Contrastive Learning

有监督学习存在以下两个问题：

- 收集大量高质量的标签往往费时费力，尤其是图这一类数据，经常出现很多节点没有标签的情况
- 往往聚焦于标签，模型学习不到迁移性较强的、有共性的知识

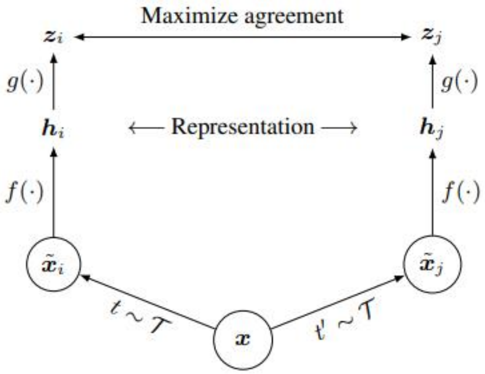




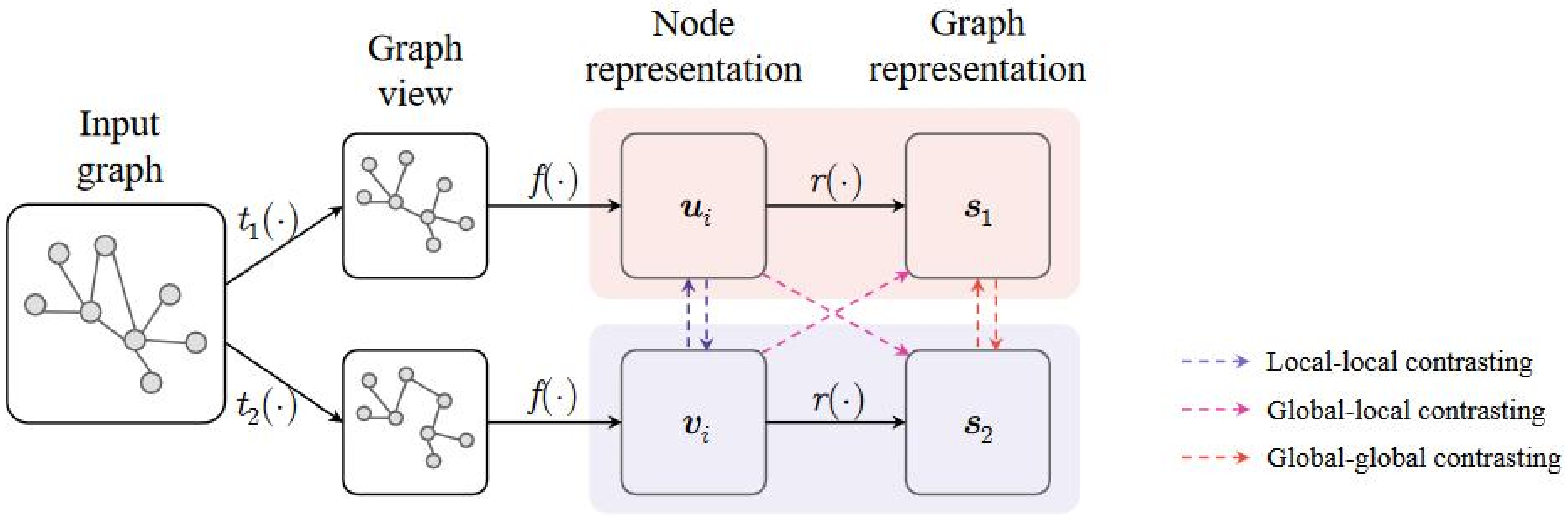


(a) Original (b) Crop and resize (c) Crop, resize (and flip) (d) Color distort. (drop) (e) Color distort. (jitter)

(f) Rotate {90°, 180°, 270°} (g) Cutout (h) Gaussian noise (i) Gaussian blur (j) Sobel filtering



$$\ell_{i,j} = -\log \frac{\exp(\mathrm{sim}(z_i, z_j)/\tau)}{\sum_{k=1}^{2N} \mathbb{1}_{[k \neq i]} \exp(\mathrm{sim}(z_i, z_k)/\tau)}$$



An Empirical Study of Graph Contrastive Learning. NeurIPS 2021

(a) data augmentation functions

(b) contrastive mode

(c) contrastive objective

(d) negative mining strategies

An Empirical Study of Graph Contrastive Learning. NeurIPS 2021

Topology augmentations:

(1) Edge Removing (ER)

(2) Edge Adding (EA)

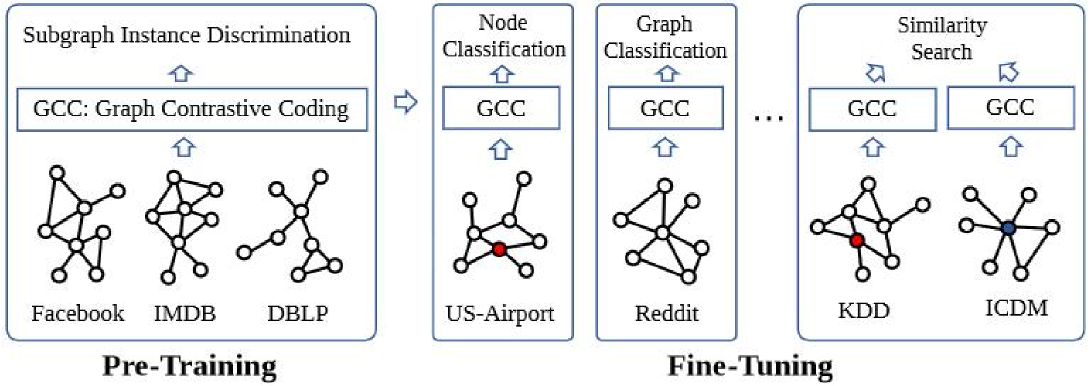
(3) Node Dropping (ND)

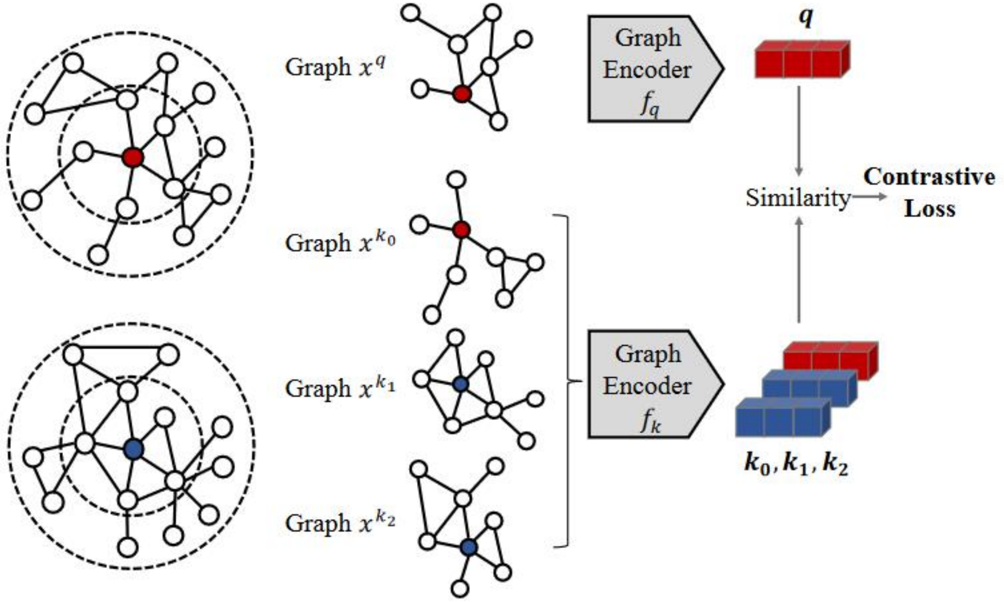
(4) Subgraph induced by Random Walks (RWS)

Feature augmentations:

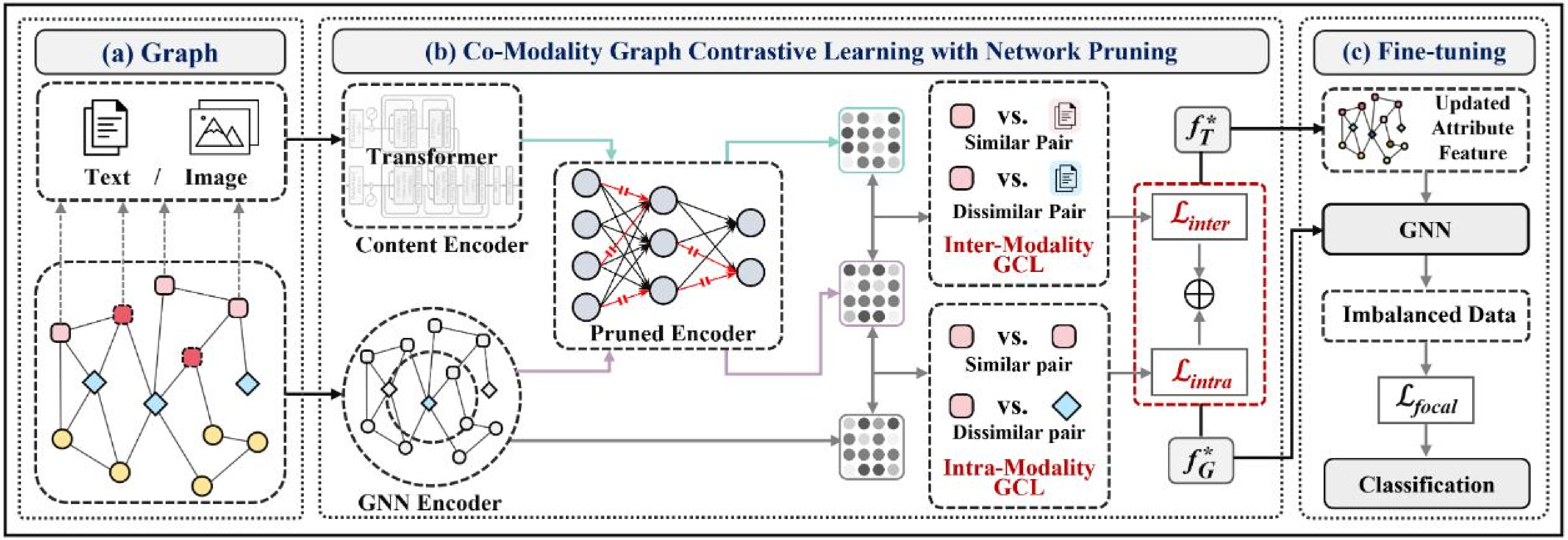
(1) Feature Masking (FM)

An Empirical Study of Graph Contrastive Learning. NeurIPS 2021





GCC: Graph Contrastive Coding for Graph Neural Network Pre-Training. KDD 2020



$$\mathcal{L}_{\text{inter}} = -\log \sum_{v_i \in \mathcal{V}} \frac{\exp\left[\text{sim}\left(\tilde{z}_G^i, \tilde{z}_T^i\right)/\tau_{\text{inter}}\right]}{\sum_{p=1}^{2n} \mathbb{1}_{[i \neq p]} \exp\left[\text{sim}\left(\tilde{z}_G^i, \tilde{z}_T^p\right)/\tau_{\text{inter}}\right]},$$

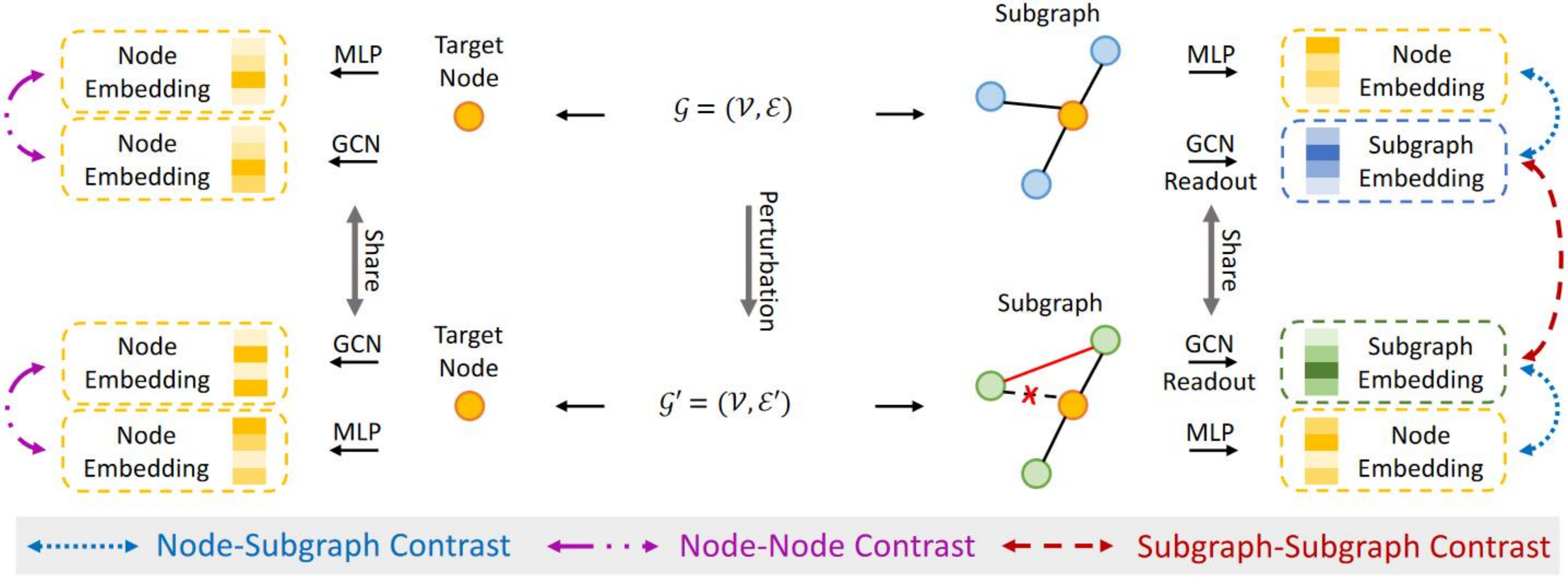
$$\mathcal{L} = \lambda \mathcal{L}_{\text{inter}} + (1-\lambda)\mathcal{L}_{\text{intra}},$$

$$\mathcal{L}_{\text{intra}} = -\log \sum_{v_i \in \mathcal{V}} \frac{\sum_{(v_i, v_p) \in \mathcal{S}} \exp\left[\text{sim}\left(\tilde{z}_G^i, \tilde{z}_G^p\right)/\tau_{\text{intra}}\right]}{\sum_{(v_i, v_q) \notin \mathcal{S}} \exp\left[\text{sim}\left(\tilde{z}_G^i, \tilde{z}_G^q\right)/\tau_{\text{intra}}\right]}.$$

$$\mathcal{S} = \{(v_i, v_p) \mid \text{sim}(x_G^i, x_G^p) \text{ in top-R of } [\text{sim}(x_G^i, x_G^p)]_{p=1}^N, \ \forall v_i \in \mathcal{V}\},$$

$$\mathcal{L}_{\text{focal}} = -\frac{1}{|\mathcal{V}_l|} \sum_{i \in \mathcal{V}_l} \sum_{c=0}^{C} \alpha_c \, y_{ic} \, (1-\hat{y}_{ic})^{\gamma} \log(\hat{y}_{ic}),$$

Co-Modality Graph Contrastive Learning for Imbalanced Node Classification. NeurIPS 2022



$$\mathbf{H}_i^{(\ell+1)} = \sigma\left(\widetilde{\mathbf{D}}_i^{-\frac{1}{2}}\widetilde{\mathbf{A}}_i\widetilde{\mathbf{D}}_i^{-\frac{1}{2}}\mathbf{H}_i^{(\ell)}\mathbf{W}^{(\ell)}\right)$$

$$s_i^1 = Bilinear\left(\boldsymbol{z}_i, \boldsymbol{e}_i\right)$$

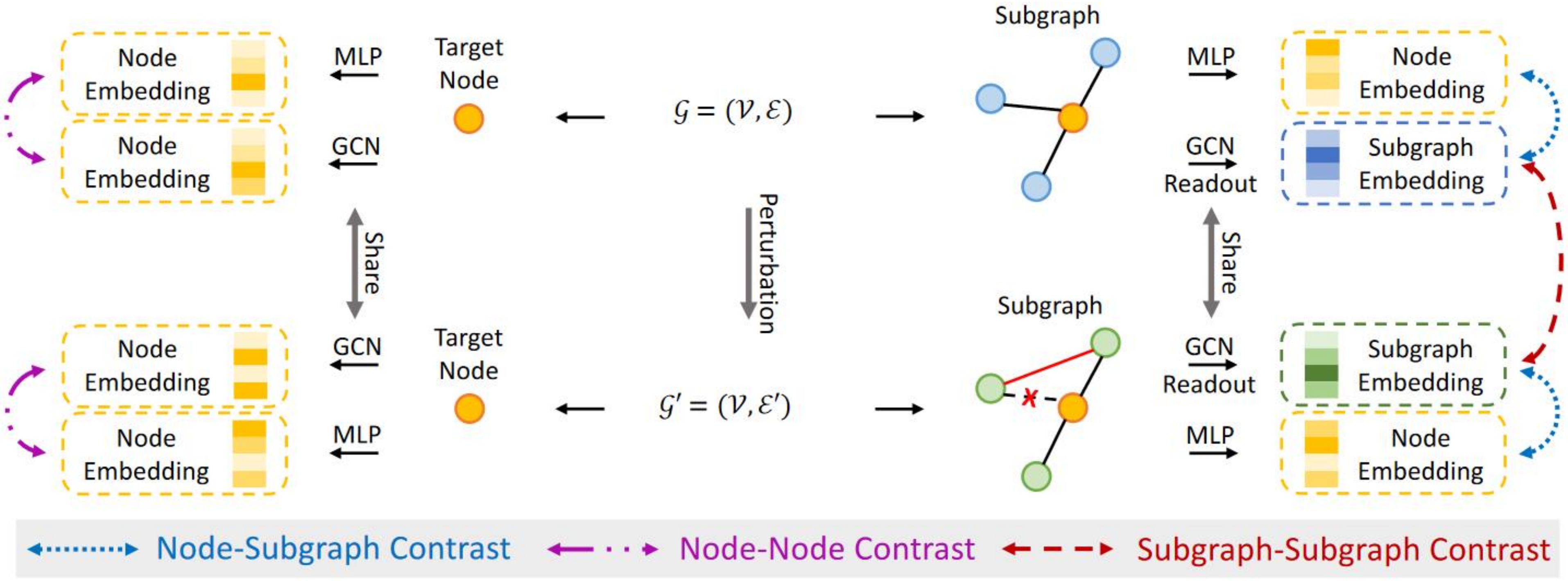
$$\boldsymbol{z}_i = Readout\left(\mathbf{Z}_i\right) = \sum_{j=1}^{n_i}\frac{\left(\mathbf{Z}_i\right)_j}{n_i}.$$

$$\mathcal{L}_{NS}^1 = -\sum_{i=1}^{N}\left(y_i \log\left(s_i^1\right) + \left(1-y_i\right)\log\left(1-s_i^1\right)\right)$$

$$\boldsymbol{h}_i^{(\ell+1)} = \sigma\left(\boldsymbol{h}_i^{(\ell)}\mathbf{W}^{(\ell)}\right),$$

$$\mathcal{L}_{NS} = \alpha\mathcal{L}_{NS}^1 + (1-\alpha)\,\mathcal{L}_{NS}^2,$$

Graph Anomaly Detection via Multi-Scale Contrastive Learning Networks with Augmented View. AAAI 2023



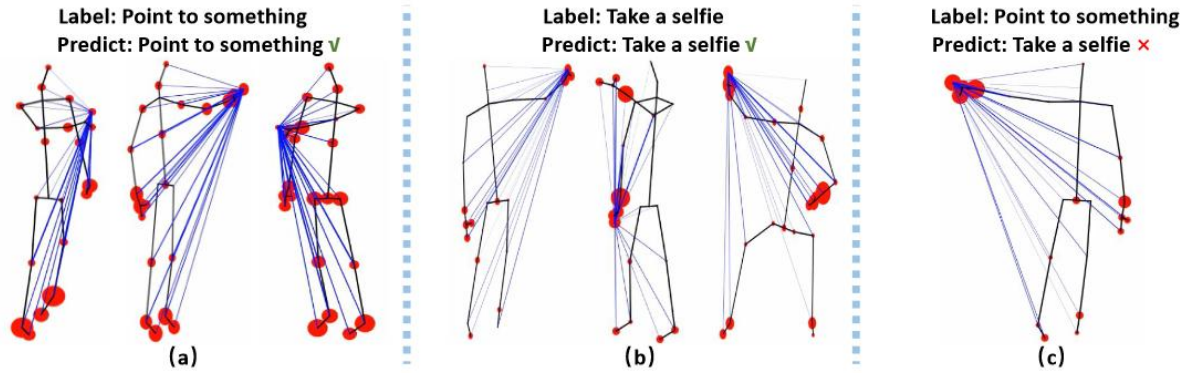
$$s_i = s_i^n - s_i^p,$$

$$\bar{S}_i = \frac{1}{R}\sum_{r=1}^{R} S_i^{(r)},$$

$$s_i = \alpha s_i^1 + (1-\alpha)\, s_i^2,$$
$$\hat{s}_i = \alpha \hat{s}_i^1 + (1-\alpha)\, \hat{s}_i^2,$$
$$S_i = \beta s_i + (1-\beta)\, \hat{s}_i,$$

$$S_i = \bar{S}_i + \sqrt{\frac{1}{R}\sum_{r=1}^{R}\left(S_i^{(r)} - \bar{S}_i\right)^2},$$

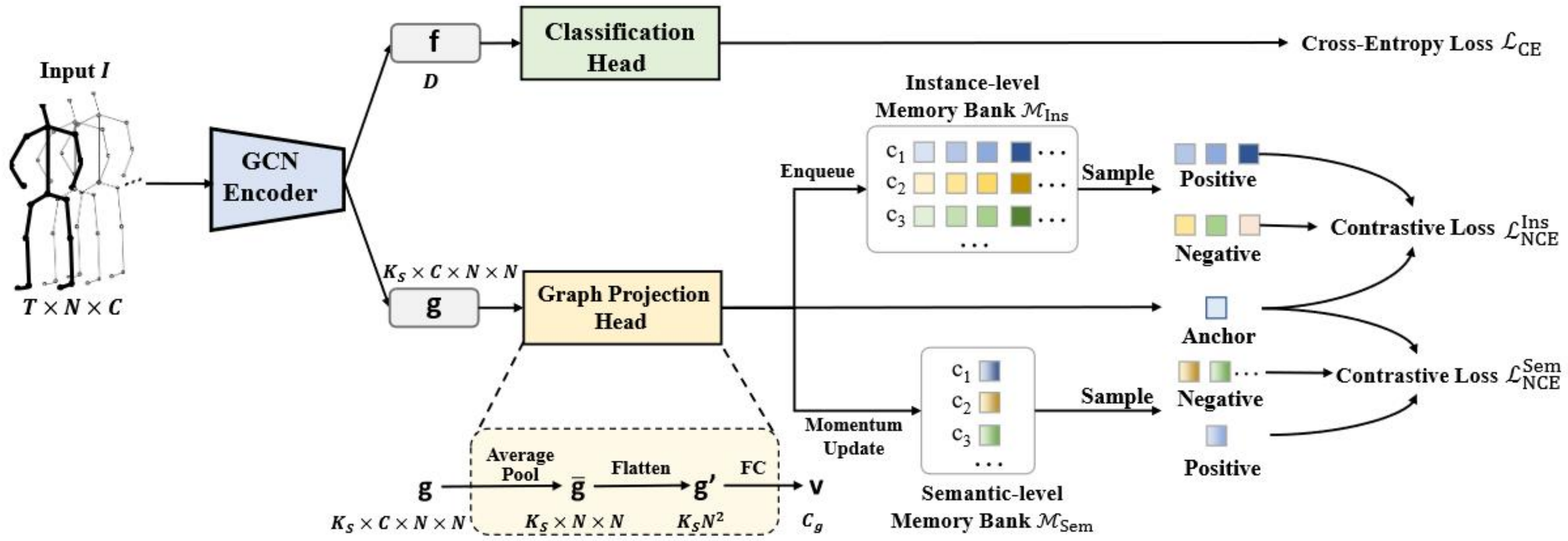
Graph Anomaly Detection via Multi-Scale Contrastive Learning Networks with Augmented View. AAAI 2023



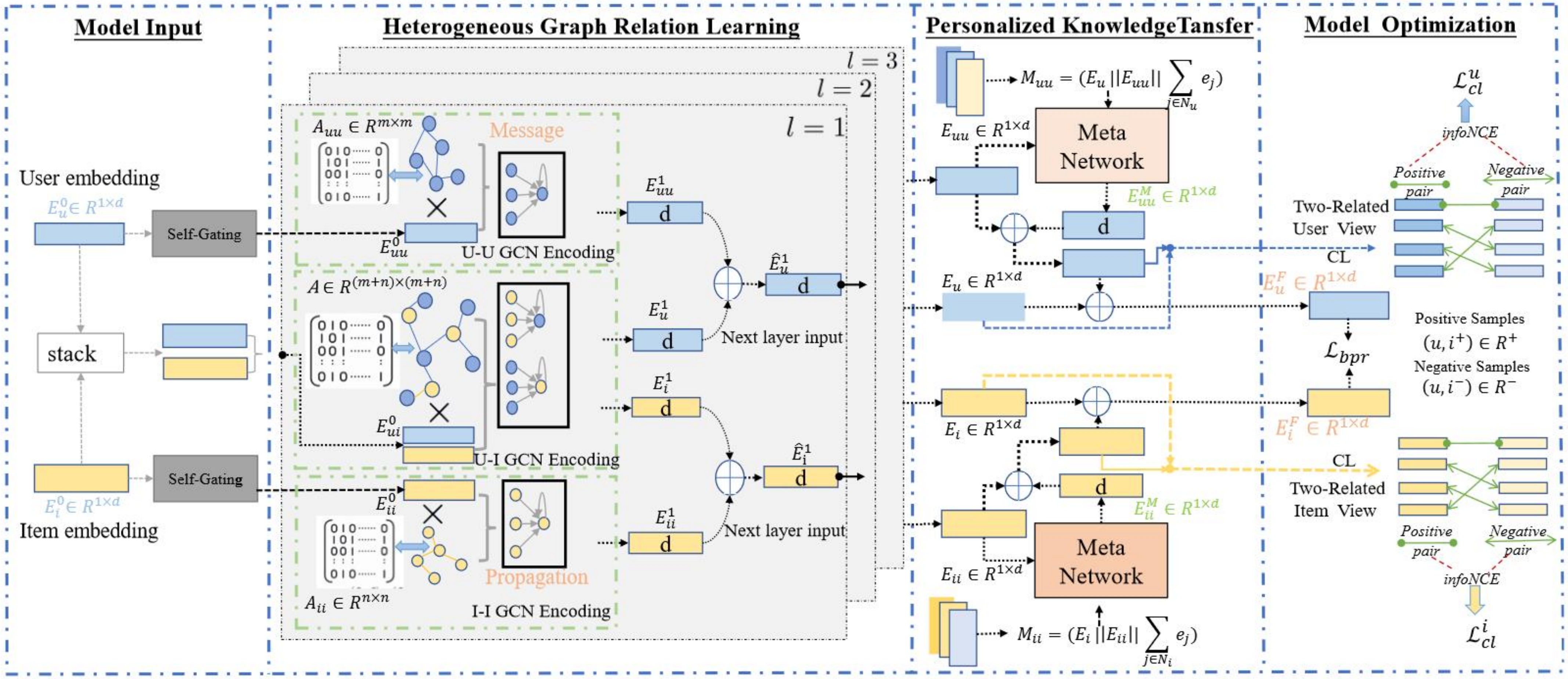
$$\mathcal{L}_{\text{CE}} = -\sum_i \mathbf{y_i} \log \hat{\mathbf{y}}_\mathbf{i}$$

$$\mathcal{L}_{\text{NCE}}^{\text{Ins}} = -\sum_{\mathbf{v}^+ \in \mathcal{N}_{\text{Ins}}^+} \log \frac{sim(\mathbf{v}, \mathbf{v}^+)/\tau}{sim(\mathbf{v}, \mathbf{v}^+)/\tau + \sum_{\mathbf{v}^- \in \mathcal{N}_{\text{Ins}}^-} sim(\mathbf{v}, \mathbf{v}^-)/\tau},$$

$$\mathcal{L}_{\text{NCE}}^{\text{Sem}} = -\sum_{\mathbf{v}^+ \in \mathcal{N}_{\text{Sem}}^+} \log \frac{sim(\mathbf{v}, \mathbf{v}^+)/\tau}{sim(\mathbf{v}, \mathbf{v}^+)/\tau + \sum_{\mathbf{v}^- \in \mathcal{N}_{\text{Sem}}^-} sim(\mathbf{v}, \mathbf{v}^-)/\tau}.$$



Graph Contrastive Learning for Skeleton-based Action Recognition. ICLR 2023



$$\mathbf{E}_{uu}^0 = \mathbf{E}_u^0 \odot \sigma(\mathbf{E}_u^0 \mathbf{W}_g + \mathbf{b}_g); \quad \mathbf{E}_{ii}^0 = \mathbf{E}_i^0 \odot \sigma(\mathbf{E}_i^0 \mathbf{W}_g + \mathbf{b}_g)$$

$$\mathbf{E}_u = \mathbf{E}_u^0 + \sum_{l=1}^{L} \frac{\mathbf{E}_u^l}{||\mathbf{E}_u^l||}; \quad \mathbf{E}_i = \mathbf{E}_i^0 + \sum_{l=1}^{L} \frac{\mathbf{E}_i^l}{||\mathbf{E}_i^l||}$$

$$\mathbf{e}_u^{l+1} = \sum_{i\in\mathcal{N}_u} \frac{1}{\sqrt{|\mathcal{N}_u|}\sqrt{|\mathcal{N}_i|}} \mathbf{e}_i^l; \quad \mathbf{e}_i^{l+1} = \sum_{u\in\mathcal{N}_i} \frac{1}{\sqrt{|\mathcal{N}_i|}\sqrt{|\mathcal{N}_u|}} \mathbf{e}_u^l$$

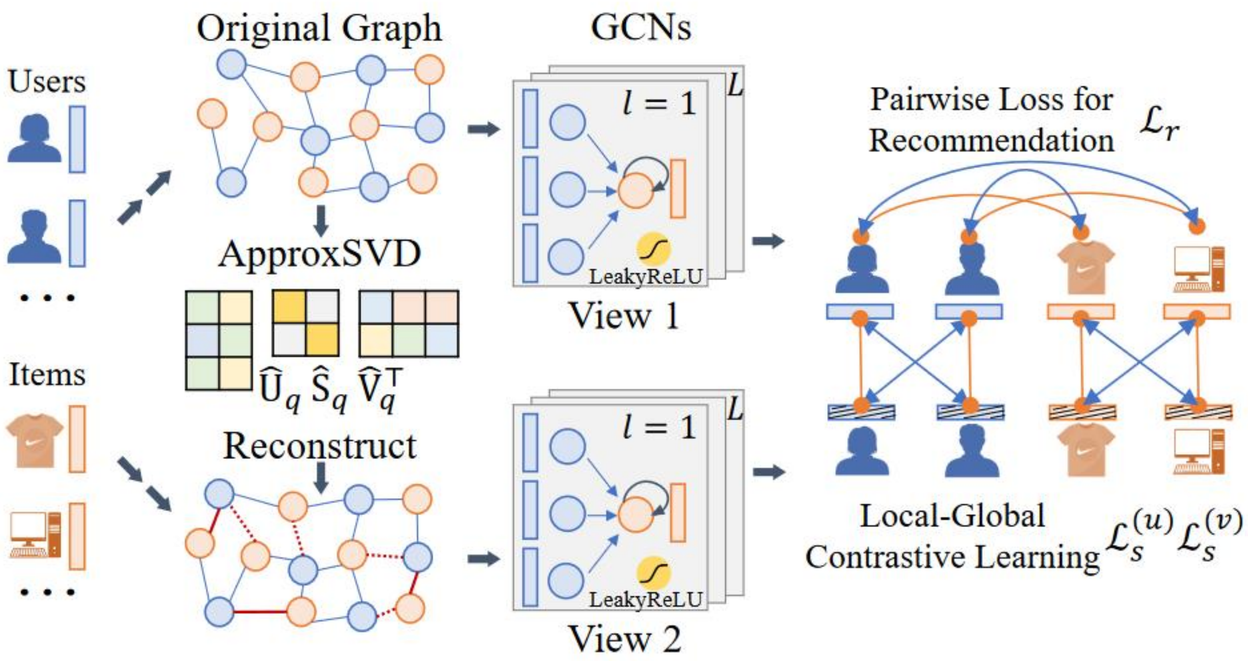
$$\mathbf{E}_{uu}^M = \sigma(\mathbf{W}_{uu}^{M1} \mathbf{W}_{uu}^{M2} \mathbf{E}_{uu})$$

$$\mathcal{L}_{cl}^u = \sum_{u\in\mathcal{V}_u} -\log \frac{\exp\left(s(\mathbf{e}_{uu}^M + \mathbf{e}_{uu}, \mathbf{e}_u)/\tau\right)}{\sum_{u'\in\mathcal{V}_u} \exp\left(s(\mathbf{e}_{uu}^M + \mathbf{e}_{uu}, \mathbf{e}_u')/\tau\right)}$$

$$\widehat{\mathbf{E}}_u^{l+1} = f(\mathbf{E}_u^{l+1}, \mathbf{E}_{uu}^{l+1}); \quad \widehat{\mathbf{E}}_i^{l+1} = f(\mathbf{E}_i^{l+1}, \mathbf{E}_{ii}^{l+1})$$

$$\mathbf{E}_u^F = \alpha_u * \mathbf{E}_u + (1 - \alpha_u) * (\mathbf{E}_{uu} + \mathbf{E}_{uu}^M);$$

Heterogeneous Graph Contrastive Learning for Recommendation. WSDM 2023



$$\mathcal{A} = \boldsymbol{U}\boldsymbol{S}\boldsymbol{V}^\top$$

$$\hat{\boldsymbol{U}}_q, \hat{\boldsymbol{S}}_q, \hat{\boldsymbol{V}}_q^\top = \text{ApproxSVD}(\mathcal{A}, q), \quad \hat{\mathcal{A}}_{SVD} = \hat{\boldsymbol{U}}_q \hat{\boldsymbol{S}}_q \hat{\boldsymbol{V}}_q^\top$$

LightGCL: Simple Yet Effective Graph Contrastive Learning for Recommendation. ICLR 2023