

Графовые свёрточные нейронные сети

Александр Колодезный БПМИ192

Национальный исследовательский университет
«Высшая школа экономики» (Москва)

23 ноября 2021 г.

Graph Convolutional Network

Вид одного слоя

$$\mathbf{h}_v^{\ell+1} = \phi^{\ell+1} \left(\mathbf{h}_v^{\ell}, \Psi(\{\psi^{\ell+1}(\mathbf{h}_u^{\ell}) \mid u \in \mathcal{N}_v\}) \right)$$

- ▶ Ψ — permutation-invariant функция
- ▶ ϕ^{l+1} и ψ^{l+1} — некоторый функция на l -ом слое

Graph Convolutional Network

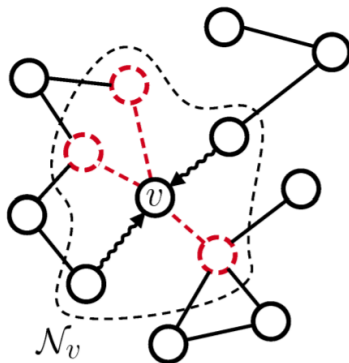
$$\mathbf{h}_v^{\ell+1} = \sigma(\mathbf{W}^{\ell+1} \sum_{u \in \mathcal{N}(v)} \mathbf{L}_{uv} \mathbf{h}_u^{\ell}),$$

- ▶ σ — сигмоида
- ▶ $\mathbf{W}^{\ell+1}$ — обучаемая матрица
- ▶ \mathbf{L}_{uv} — нормированный Лапласиан

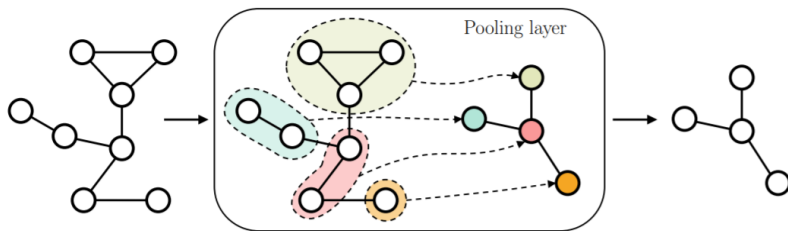
Sampling

Виды

- ▶ Для каждой вершины выбираем случайное подмножество соседних, которые будут влиять на новое значение h_v . (GraphSAGE).
- ▶ Выбирать случайное множество вершин всего графа и обучаться только на полученном ими подграфе (FastGCN).



Pooling



- ▶ Сжатие кластеров вершин в одну
- ▶ Уменьшает размеры графа
- ▶ Уменьшает вычислительную стоимость

Differentiable Pooling

- ▶ Обучаем для каждой вершины попадание в кластера

$$\mathbf{S}^{\ell+1} = \text{softmax}(\text{GNN}(\mathbf{A}^{\ell}, \mathbf{H}^{\ell})),$$

- ▶ Пересчитываются embedding для новых вершин, и новая матрица смежности

$$\mathbf{H}^{\ell+1} = \mathbf{S}^{\ell+1T} \mathbf{H}^{\ell} \quad \text{and} \quad \mathbf{A}^{\ell+1} = \mathbf{S}^{\ell+1T} \mathbf{A}^{\ell} \mathbf{S}^{\ell+1}.$$

- ▶ Новая матрица смежности оказывается полной

Top-k Pooling

- ▶ Для каждой вершины посчитаем attention как проекция его эмбединга на обучаемый вектор p

$$s^{\ell+1} = \frac{\mathbf{H}^{\ell} p^{\ell+1}}{\|p^{\ell+1}\|}.$$

- ▶ Выбираем k вершин с наибольшим полученным attention
- ▶ Оставляем m графе только найденные вершины
- ▶ Расширение метода Self-attention pooling

$$s^{\ell+1} = \sigma(\text{GCN}(\mathbf{A}^{\ell}, \mathbf{H}^{\ell})).$$

Другие методы pooling-a

- ▶ Edge pooling — выбираем ребро, которое сжимаем

$$s^{\ell+1}((v, u) \in \mathcal{E}_g) = \sigma(\mathbf{w}^T [\mathbf{h}_v^\ell, \mathbf{h}_u^\ell] + \mathbf{b}).$$

- ▶ Топологические виды pooling-a
 - ▶ GRACLUS — алгоритм основанный на спектральной кластеризации
 - ▶ Non-negative Factorization Matrix Pooling — pooling основанный на NFM факторизации матрицы смежности.

Graph embedding

- ▶ Хотим построить embedding для всего графа
- ▶ Нужно агрегировать embedding для нод
- ▶ Можно делать rolling графа, пока не останется одна вершина
- ▶ Применить перестановочно-инвариантную функцию от всех нод

$$\mathbf{h}_g^\ell = \Psi\left(\{f(\mathbf{h}_v^\ell) \mid v \in \mathcal{V}_g\}\right),$$

- ▶ Выбрать Ψ как сумму, максимум или минимум, а f — тождественную функцию.

Список литературы

- ▶ <https://arxiv.org/pdf/1912.12693.pdf>
- ▶ <https://arxiv.org/pdf/1609.02907.pdf>
- ▶ <https://arxiv.org/pdf/1812.04202.pdf>
- ▶ <https://arxiv.org/pdf/1801.10247.pdf>
- ▶ <https://arxiv.org/pdf/1706.02216.pdf>