

Оценка неопределённости для графовых свёрточных сетей

Чупахин Андрей
Соболева Евгения
Лобанов Валерий

Студенческая образовательная программа
Университета «Сириус»
"Современные методы теории информации,
оптимизации и управления"

6 августа 2021 г.

Графовые свёрточные сети



- ▶ На слое l каждая вершина v_i^l имеет свой вектор свойств h_i^l
- ▶ Каждый слой GCN агрегирует свойства из u_i^l , где u_i - соседние вершины v , в представление следующего слоя v_i^{l+1}

$$h_j^{\text{new}} = f \left(\sum_{j=j+1}^n h_j \right)$$

Зачем нужна оценка неопределённости для графов

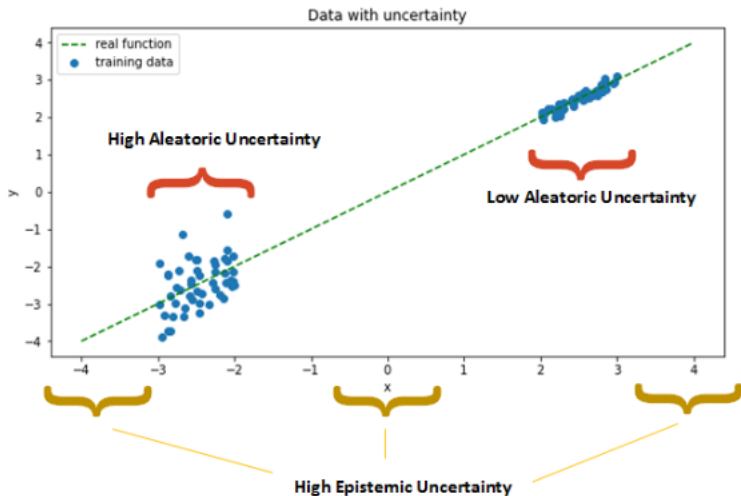
- ▶ Оценка уверенности модели в жизненно важных предсказаниях (медицинские прогнозы, социальные графы для детекции фрода, улучшение рекомендательных систем)
- ▶ Доразметка данных с наибольшей неопределенностью для уменьшения времени активного обучения модели

Оценки неопределенности

Предсказательную неопределенность можно разграничить двумя терминами: aleatoric и epistemic

- ▶ Aleatoric uncertainty отражает шум в данных
 - ▶ Эту часть неопределенности нельзя устранить добавлением данных
- ▶ Epistemic uncertainty отражает недостаток информации
 - ▶ Эту часть неопределенности можно снизить путем добавления новых данных

Оценки неопределенности



NUQ метод

Рассмотрим задачу бинарной классификации $(X, Y) \in \mathbb{R}^d \times \{0,1\}$. Необходимо найти правило $\hat{g}$, которое аппроксимирует оптимальное на датасете $\mathcal{D}$:

$$g^* = \arg \min_g \mathbb{P}(g(X) \neq Y)$$

Рассмотрим правило классификации $\hat{g}(x) = \hat{g}_{\mathcal{D}}(x)$ на датасете $\mathcal{D}$. Определим оценку риска в каждой точке:

$$\mathcal{R}(x) = \mathbb{P}(\hat{g}(X) \neq Y \mid X = x)$$

где

$$\begin{aligned} \mathbb{P}(\hat{g}(X) \neq Y \mid X = x) &\equiv \mathbb{P}_{\text{tr}} (\hat{g}(X) \neq Y \mid X = x) \equiv \\ &\equiv \mathbb{P}_{\text{test}} (\hat{g}(X) \neq Y \mid X = x) \end{aligned}$$

NUQ метод

Предлагается рассмотреть следующее определение оценки total uncertainty :

$$U_t(x) = \min\{\eta(x), 1 - \eta(x)\} + 2\sqrt{\frac{2}{\pi}}\tau(x)$$

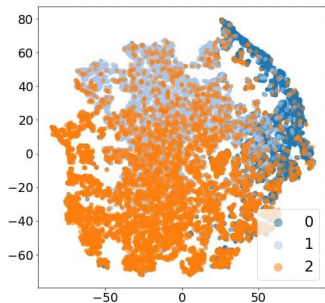
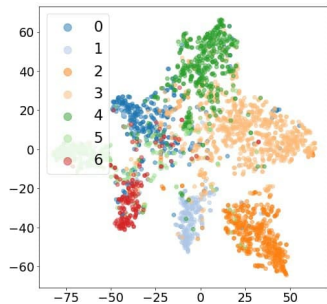
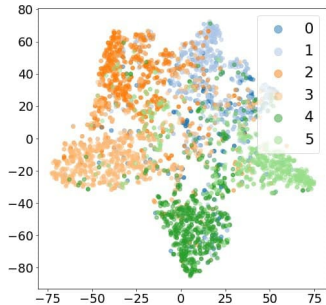
полученное оценкой асимптотической аппроксимации

$$\mathbb{E}_{\mathcal{D}}\mathcal{L}(x) = \min\{\eta(x), 1 - \eta(x)\} + 2\mathbb{E}_{\mathcal{D}}|\hat{\eta}(x) - \eta(x)|$$

где $\tau(x)$ - стандартное отклонение Гауссиана

$$\tau^2(x) = \frac{1}{nh^d} \frac{\sigma^2(x)}{p(x)} \int [K(u)]^2 du$$

t-SNE эмбеддингов : CiteSeer, Cora, PubMed



AUROC for the Misclassification detection

Data	Model	Ep.	Al.	Total	MaxProb.
Cora	NUQ	73.68	74.66	77.42	72.69
	S-BGCN-T-K	68.8	75.3	—	—
Citeseer	NUQ	60.81	64.78	64.59	74.14
	S-BGCN-T-K	60.7	67.2	—	—
PubMed	NUQ	65.18	71.23	71.09	67.64
	S-BGCN-T-K	64.2	69.3	—	—

AUROC for the Out of Distribution detection

Data	Model	Ep.	Al.	Total	MaxProb
Cora	NUQ	70.98 ± 0.02	71.45 ± 0.02	71.81 ± 0.02	71.84 ± 0.03
	S-BGCN-T-K	70.8	85.5	—	—
Citeseer	NUQ	55.46 ± 0.03	61.76 ± 0.02	61.75 ± 0.02	71.57 ± 0.05
	S-BGCN-T-K	55.1	78.4	—	—
PubMed	NUQ	58.84 ± 0.02	66.07 ± 0.03	66 ± 0.01	59.78 ± 0.05
	S-BGCN-T-K	59.2	71.8	—	—

Дальнейшая работа

- ▶ Построить PCA эмбедингов, визуализация uncertainty
- ▶ Проверить метрики на state-of-the-art архитектурах GCN
- ▶ Сравнить NUQ с deep ensembles GCN
- ▶ Применить bi-Lipschitz для GCN
- ▶ ...

Вклад участников

- ▶ Андрей — работа с эмбедингами и архитектурами GCN
- ▶ Евгения — подсчёт метрик и графики
- ▶ Валерий — NUQ и misclassification detection