

You Can Have Better Graph Neural Networks by Not Training Weights at All: Finding Untrained GNNs

Tickets

Tianjin Huang, Tianlong Chen, Meng Fang, Vlado Menkovski, Jiaxu Zhao, Lu Yin, Yulong Pei, Decebal Constantin Mocanu, Zhangyang Wang, Mykola Pechenizkiy, Shiwei Liu

Автор рецензии: Орлов Александр

Обзор

1. В последнее время графовые нейронные сети в разных своих вариациях (GCN, GIN, GAT, GraphSAGE и тп) успешно применяются для широкого круга задач. Но обучать графовые нейронные сети больно, времязатратно и ресурсозатратно. И как замечают авторы в последнее время появляются многообещающие результаты в области необучаемых сетей. В этих работах появляются модели, которые уже могут соперничать по точности со полностью обученными CNN без какого-либо обновления весов. Опираясь на этот опыт, авторы исследуют необучаемые разреженные GNNы, отвечая на вопрос возможно ли получить сносную графовую сеть вообще не обучая веса.
2. Статья была выпущена 10 сентября 2022 и представлена на конференции LOG 2022 Conference, которая прошла с 9 по 12 декабря 2022. На конференции статья получила Oral & Best Paper Award
3. Посмотрим на некоторые работы, в которых авторы статьи принимали участие:
 - "Sparse Training via Boosting Pruning Plasticity with Neuroregeneration"
 - "Sparsity May Cry: Let Us Fail (Current) Sparse Neural Networks Together!"
 - "Sparse MoE as the New Dropout: Scaling Dense and Self-Slimmable Transformers."
 - и т д

Можно заметить, что все эти работы так или иначе связаны с разреженными нейронными сетями. У многих авторов так же есть статьи по графовым нейронным сетям и, например, их применениям в биологии. Так что вполне ожидаемо видеть от этих людей работу на подобную тему, в которой в экспериментах используются биологические данные

4. У статьи пока нет цитирований, вероятно, из-за новизны работы. Но в то же время многие статьи (в количестве 45 штук) по графовым нейронным сетям ссылаются на нее.

Сильные стороны

1. Эта работа решает хорошо мотивированную проблему. Интересно, что даже для таких сложных архитектур, как GNN, можно построить хорошую модель не обучая веса. Возможно, это стоило ожидать после аналогичного результата для CNN, но все равно
2. Статья очень хорошо написана, ее легко читать и в ней хорошо прослеживаются основные мысли и шаги авторов. Интуитивно понятные объяснения даются по всему документу для каждого принятого решения.
3. Был проведен очень тщательный набор экспериментов, чтобы продемонстрировать эффективность предлагаемого подхода
4. Авторы сделали очень подробный и информативный обзор литературы, где рассказали про все статьи на смежные темы.
5. В статье показывается, что разреженные сети имеют хорошую обобщающую способность, а также устойчивы к изменениям в графе

Слабые стороны

1. Вначале описание "разреженности" показалось мне немного запутанным. Понимание того, о чем идет речь, приходит после просмотра формул и псевдокода, но некоторая интуиция на раннем этапе была бы полезна.
2. Анализ сглаживания очень интересен, однако он носит чисто эмпирический характер. Было бы неплохо увидеть какой-нибудь теоретический результат, который показывает, как UGT предотвращают неразличимость объектов узла по мере накопления большего количества слоев GNN.

Вопросы

1. Есть ли у авторов направление для какого-либо теоретического анализа по Untrained GNNs Tickets (UGTs)?
2. Хотя авторы утверждают, что цель не в производительности, все равно хочется узнать о том, как хорошо с точки зрения ресурсов и скорости работает предложенный подход. Могли бы авторы предоставить результаты по времени выполнения экспериментов?
3. В статье используется придуманный авторами способ для установления разреженности. Было бы интересно посмотреть на применение других методов (например, использовать не линейную политику, а какую-нибудь другую).