

ИССЛЕДОВАНИЕ ФУНКЦИЙ ПОТЕРЬ ДЛЯ МЕТОДОВ ОБУЧЕНИЯ ВЕКТОРНЫХ ПРЕДСТАВЛЕНИЙ НА
ГРАФОВЫХ ДАННЫХ

Андреева П. О. (Университет ИТМО)
Научный руководитель – Ваганов Д.А.
(Университет ИТМО)

В данной работе изучены существующие варианты графовых нейронных сетей, выделены основные функции потерь, используемые для обучения векторных представлений на графовых данных. Также в докладе рассматриваются возможности обобщения всех выбранных функций потерь и проведены эксперименты по выявлению оптимальных конфигураций архитектур и функций потерь для различных задач и данных.

Введение. Основной проблемой графовых данных является включение в модель машинного обучения информации из графовой структуры, являющейся намного более сложной, чем матричная. Одним из вариантов решения данной проблемы является обучение представлений графовых данных в низкоразмерном пространстве с последующим решением задач уже на обученных векторах. Представление данных в низкоразмерном пространстве является уже задачей машинного обучения само по себе, а не предобработкой, поэтому важную роль в обучении играет и архитектура нейронной сети и функция потерь. Общая идея функций потерь неконтролируемого построения представлений – устремлять ближе друг к другу представления тех вершин, которые “схожи” и в графе.

В существующих сравнительных анализах методы рассматриваются в том виде, как представлены в оригинальных статьях, то есть в своих конкретных сочетаниях архитектуры с функциями потерь, показавшими наибольшую эффективность для конкретного метода. Тогда как целью данной работы является абстрагирование от архитектуры и рассмотрение в деталях функций потерь.

Основная часть. По результатам обзора источников были выбраны в качестве элементов архитектуры нейронной сети только графовые свёрточные слои, так как методы на их основе являются хорошо масштабируемыми, показывают одни из лучших результатов, а также позволяют переводить в векторы данные, которые не участвовали в обучении.

В качестве функций потерь были выбраны те, что позволяют обучать представления только на топологии графа, не используя размеченные данные. Именно функция потерь определяет то, как будут расположены относительно друг друга представления узлов графа в векторном пространстве. В векторном пространстве достаточно просто определить понятие близости или схожести – разница или скалярное произведение векторов. Поэтому отличия функций потерь скорее ограничиваются различиями в определении того же понятия между вершинами в самом графе.

Выделены основные варианты схожести вершин в графе:

- вероятность появления двух вершин в одном случайном блуждании (случайное блуждание может быть построено различными способами: фиксированной или бесконечной длины с некоторой вероятностью остановки или вероятностью повторного старта и т.д.);
- вес ребра между вершинами (или его наличие/отсутствие в невзвешенном графе);
- длина самого короткого пути между вершинами;
- количество общих соседей;
- геометрическая схожесть вершин (по числу соседей).

Выбранные функции потерь можно привести к общему виду на основе следующих суждений:

- возможно сделать такую замену переменных, после которой аналитический поиск оптимума окажется проще;

- в случае обучения представлений в качестве такой замены выбирается произведение старых аргументов;
- при получении значения произведения представлений, при котором достигается минимум исходной функции, возможно составить новую функцию потерь, устремляющую произведение векторов к найденному значению.

В качестве экспериментов, показывающих оптимальные конфигурации сверточных сетей и функций потерь были построены нейронные сети, состоящие из графовых сверточных слоев. После построения представлений на задачах классификации и предсказания была оценена средняя точность каждой из конфигураций. Кроме того, проведено сравнение с точностью аналогичной задачи в контролируемых условиях.

Выводы. В работе были рассмотрены существующие функции потерь неконтролируемого обучения представлений графовых данных, выделены основные типы понятий схожести вершин в графе, а также предложен подход к обобщению выбранных функций потерь. Кроме того были проведены эксперименты по поиску оптимальных конфигураций архитектуры нейронной сети и функций потерь, выявлены различия и сходства между функциями потерь, оценены преимущества при решении конкретных задач. Дальнейшая работа будет заключаться в построении экспериментов и оценке точности с учетом особенностей данных (структура графа, количество признаков и тп).

Андреева П.О. (автор)

Ваганов Д.А. (научный руководитель)