

Рецензия на рукопись
Дифференцируемый алгоритм поиска архитектуры модели с контролем её
сложности
Автор Яковлев К.Д.

Данная работа посвящена исследованию алгоритма DARTS поиска оптимальной структуры нейронной сети. Основной предмет исследования – это способы регуляризации гиперсети, как один из возможных вариантов преодоления проблемы нестабильности в обучении. Научная новизна работы состоит в описании зависимости метрики precision на валидационной выборке от параметров регуляризации алгоритма. В работе представлены информативные графики и таблицы, доказывающие теоретические рассуждения автора о контроле сложности сети.

Аннотация и введение:

В этой части с первых строк появляется большое множество новых терминов. Для удобства чтения их можно выделить курсивом, а самые важные (такие как «модель») вынести отдельно в абзац «определение», как в книгах по математике.

Абзац 1, последняя строка: неясно, как применить алгоритм к рекуррентным сетям (к свёрточным тоже не прокомментировано). Возможно, стоит это доказать ссылкой (для любознательного читателя) или убрать вовсе (далее это не пригождается).

Абзац 3, строка 7: неясно, какие альтернативные подходы применяются, не мешало бы их перечислить. Возможно, они находятся в теоретической части, но она читается позже введения. Если имеется ввиду «в других работах предлагаются», то это неоднозначная трактовка написанного предложения. Далее приводится комментарий к другому алгоритму, который может являться лишним (читатель, желающий понять вероятностный подход получает ссылку на оригинальную статью, а не желающему эта информация только мешает)

Постановка задачи:

Для визуализации хорошо могут помочь изображения ячеек и слоёв (в литературе нейронные сети принято рассказывать с рисунками)

Часть 2.2, первое предложение: «множество параметров, контролирующЕе»

Также, возможно, следует больше описать корректность перехода от дискретного пространства описаний модели к непрерывному (например, отсутствие потери качества). Этот пункт является хорошей идеей для рассмотрения автору и, если это объяснено в других работах, может быть проигнорирован.

Описание алгоритма:

Алгоритм описан достаточно сжато, но понятно и с комментариями к месту. Может быть полезно добавить рассуждения об оптимальности модели после перехода к дискретному описанию (класс моделей мог значительно измениться)

Вычислительный эксперимент:

Эксперимент реализован согласно заявленному описанию, а результаты обоснованы теоретически. Так как модель hurnet показала сравнимые, но всё же более низкие результаты, по сравнению с базовым алгоритмом DARTS, может быть полезно рассказать читателю, в чём её плюсы перед уже построенным алгоритмом и почему она может быть предпочтительнее

Вывод:

Статья однозначно рекомендуется к публикации, однако является трудновоспринимаемой для широкой публики. Это приводит к тому, что будет полезно сделать пометку о сложности статьи и требуемой подготовке к её восприятию, или добавлению подробностей в саму статью, если она планировалась для более широкой публики.

Статья будет полезна специалистам в области и интересующимся людям как хороший анализ одного из алгоритмов контроля сложности нейросети.

Рецензент:
Пырзу В.В.