

УДК 004.852

Использование многомерных пространств для определения минимального значения ошибки и оптимального количества нейронов в архитектуре нейросети

Черников К.М. (ГБНОУ СО Самарский региональный центр для одарённых детей)

Научный руководитель - Сухаренко Д.В

(ГБНОУ СО Самарский региональный центр для одарённых детей)

Введение. При решении практических задач с применением AI-алгоритмов возникает проблема определения количества нейронов в слоях при обучении ИНС, а также в поиске минимального значения ошибки. На данный момент наиболее популярный подход для определения минимума ошибки – метод градиентного спуска, суть которого заключается в высчитывании соотношения производной функции нейрона к функции ошибки сети[1]. Такой подход наиболее оптимален с точки зрения асимптотики и результативности. Существует также подход представления закономерности показателя ошибки от гиперпараметров в виде многомерного пространства, где значение ошибки является одной из осей, минимизация по которой необходима для подбора оптимальных параметров сети[2]. Особенности и улучшения этого подхода были рассмотрены в работе.

Основная часть. С помощью многомерной оптимизации решается множество вопросов определения оптимальных параметров ИНС[3]. Но ввиду ограничений значений весов эта оптимизация не может предсказывать релевантность использования того или иного количества нейронов. При добавлении в область допустимых значений нуля, можно будет с помощью многомерной оптимизации определить, без каких нейронов ИНС работает лучше. Таким образом можно определить оптимальное количество нейронов для конкретного слоя, при учёте что оптимизация происходит при достаточном количестве примеров метрики при различных весах. Можно получить это при интерполяции кривыми Безье для отдельных отрезков, концы которого – точки в многомерном пространстве, где одна из осей – ошибка, а остальные – веса. Тем самым при объединении этих кривых можно получить плоскость, благодаря которой сможем определить необходимое количество нейронов.

Выводы. Проведя сравнительный анализ методов стохастического градиентного спуска и многомерной оптимизации путём интерполяции было выявлено, что второй подход может помочь в автоматизации подбора не только весов, но и количества нейронов, ввиду включения нуля в возможные значения веса.

Список использованных источников:

1. Оптимизация в ML // Учебник по машинному обучению URL: <https://stage.bigenc.ru/geography/text/5684038> (дата обращения: 17.01.2024).
2. Е.М.Захарова, И.К.Минашина Обзор методов многомерной оптимизации // Информационные процессы Том 14, № 3. - М: Институт проблем передачи информации им. А.А. Харкевича РАН, 2014. - С. 256–274.
3. Методы многомерной оптимизации // Studfile URL: <https://studfile.net/preview/3653207/page:5/> (дата обращения: 28.01.2024).