

Федеративное обучение и сверхпараметризация в моделях

Saturday, 17 May 2025 13:45 (15 minutes)

В этом отчете будет представлено исследование влияния техники рестартов на оптимизацию обучения сверхпараметризованных моделей. Исследуется теоретическая связь между обобщенной (L_0 , L_1)-гладкостью функций потерь и явлением двойного спуска, характерным для сверхпараметризованных моделей. Основная гипотеза исследования состоит в том, что рестарты в градиентных методах могут выполнять роль механизма имплицитной регуляризации, способствующего обнаружению решений с повышенной обобщающей способностью.

Будут представлены результаты экспериментов на модифицированной архитектуре ResNet-18 с расширенными полносвязными слоями, демонстрирующие возможность обнаружения таких [с повышенной обобщающей способностью] решений.

Проведено больше экспериментов, уделяя особое внимание сравнению разных критериев для проведения рестартов, а также, ожидается, что будет сформулирована гипотеза о том, какой из критериев будет показывать лучшие результаты, в сравнении с другими.

Исследование использования техники рестартов в комбинации с методами понижения размерности и на других архитектурах.

Primary author: EIDLIN, Ivan

Co-author: Dr GASNIKOV, Alexander (Кафедра математических основ управления, Школа прикладной математики и информатики, Московский физико-технический институт (МФТИ))

Presenter: EIDLIN, Ivan

Session Classification: Оптимизация и машинное обучение

Track Classification: Машинное обучение и нейросети