

Статья о предсказании параметров нейросетей при фиксированной задаче и наборе данных при помощи графовой архитектуры за один проход. Авторы статьи: Борис Князев, Michal Drozda, Graham W. Taylor, Adriana Romero-Soriano.

Среди сильных сторон работы можно отметить возможность быстрого fine-tuning (всего за 5 эпох) благодаря хорошей инициализации весов. Так же кроме весов нейронной сети модель способна предсказывать получаемое accuracy, inference speed и convergence speed.

Из слабых сторон можно отметить, что модель не показала SOTA результатов, проигрывает SGD и Adam. Так же существуют ограничения, что модель не может предсказывать параметры для новых задач и не может обобщаться на разные датасеты (на вход не подаются входные данные исходной нейронной сети на которой обучаемся).

Написано хорошо, но для полного и безболезненного понимания статьи нужны знания в графовых нейронных сетях.

Несмотря на отсутствие практика, качественное описание всех необходимых деталей для запуска дает уверенность в хорошей воспроизводимости. Есть хорошо проработанный авторский репозиторий с моделью.

По итогу имеем громкую статью с большим количеством интересных идей, с хорошим описанием всех деталей, проработанным авторским репозиторием и неплохими результатами.

Оценка по НИПС: 8 (уверенность 4)