

Рецензия на рукопись.  
Дистилляция моделей и данных.  
Баринов Н.

**Статья:**

- Introduction написан хорошо, мне нравится.
- В пункте 3.1 дистилляция данных мне непонятно,  $\theta_t^* := \theta_t^*$ , т.к. там написано, что инициализируем случайным значением параметры модели-студента, нужно поправить обозначения.
- В пункте 3.2 Дистилляция моделей совсем не понятна форма записи функции потерь, наверное, стоит пояснить её.
- В пункте 3.3 не написана новая функция потерь.
- В пункте 4.1 в таблице пропущено значение для ConvNet, хотя эксперименты для неё проведены.
- Вывода пока нет

**Графики:**

Стоит сохранить картинки, а не делать скриншоты, это сильно портит качество картинок.

Можно зарегистрироваться на kaggle, чтобы проделать более масштабные эксперименты с большим количеством эпох, 30 часов вычислительной мощности в неделю это позволяет. Я делал там, это удобно.

**Код:**

- Модуль networks не загружен на github.
- В статье идет речь об экспериментах только на CIFAR10, внутри есть ещё код для экспериментов на MNIST, может быть ещё не добавили или уже убрали.
- Код понятный, читаемый, эксперимент воспроизводится, нужно убрать импорт из networks, т.к. модели уже были добавлены в файл, а не импортируется из модуля.

Рецензент:  
Крейнин М.В.