

1. Вклад: Исследуют известный феномен deep double descent, проводят очень много экспериментов (статья размером 24 страницы), вводят понятие EMC и показывают его связь с deep double descent, вводят epoch-wise double descent.
2. Сильные стороны: Очень много экспериментов, интересные объяснения.
  - 2.1 Корректность - утверждений не делается, поэтому все корректно, все попытки объяснить результаты позиционированы как гипотезы
  - 2.2 Значимость - проводится много экспериментов, достаточно чтобы показать что феномен глобально существует в Deep Learning
  - 2.3 Актуальность - достаточно высокая, и даже не только в теории DL.
3. Слабые стороны:
  - 3.1 Deep Double Descent не новое понятие
  - 3.2 Свои объяснения феномену не доказывают и не дают достаточно обоснований, а оставляют как гипотезы
  - 3.3 Процедура обучения слишком проста, только SGD. Я бы попробовал лучшие гиперпараметры для текущей модели.
  - 3.4 EMC введена недостаточно формально чтобы быть термином, а не пояснением к гипотезе
4. Статья написана очень хорошо
5. Воспроизводима полностью
6. Хотел бы увидеть другие способы оптимизации, Adam, SWA, SAM.
7. 6-7
8. 3-4