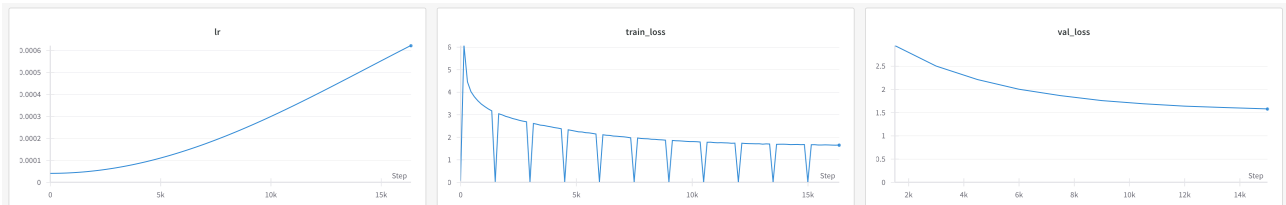


БДЗ1-отчет

Привет. Рассказываю про то, что удалось сделать на бдз.

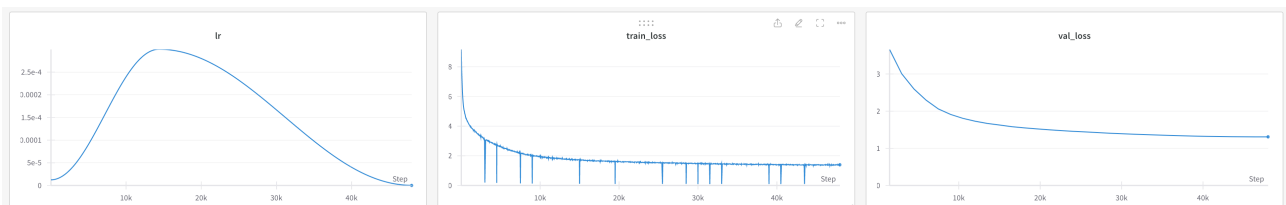
Архитектура обычная – эмбединг, энкодинг, декодер, линейар. Игрался с числом голов аттеншна, числом слоев декодера, попробовал пару шедулеров, на чем в принципе и остановился. Как оптимизатор взял AdamW, тоже без выкрутасов.

Сначала прогнал в тестовом порядке – ExponentialLR, 4 головы, 6 слоев, аттеншн шириной в 512.

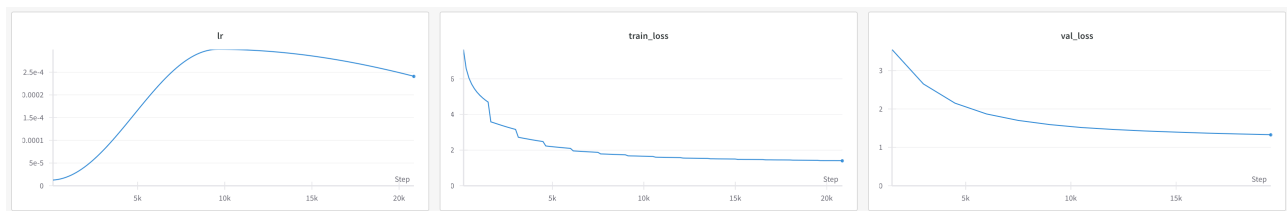


16к шагов очевидно не хватило, я хотел вал лосс ближе к 1.3, плюс я сильно накосячил в трейн лупе. Видимо, неправильно суммировал лосс, в общем, решил все удвоить после исправления ошибок – сделал размерность 1536, 8 голов, 8 слоев. Плюс заменил exponential lr на onecycle.

Все прошло просто прекрасно, моделька натренировалась, лосс вышел приемлемым, я хотел было делать eval, но датасфера мощно упала и я потерял файл с чекпойнтом и саму модель.



Ну и хрен с ним. Прогнал точно так же, но прервал тренировку посередине, потому что могу и хочу, плюс еще надо было сделать эвалюэйшн, а время 8 утра в день сдачи.



It is what it is. Текстики на выходе дают такой умеренный бред.

one evening, a big boy named tim went to the farm. he saw many animals there. he saw a cow, a sheep, and a horse. tim was happy. he wanted to play with the animals.as tim played with the cow, he heard a loud noise. the cow said, "moo! be careful, don\'t jump or hit my cloud!" tim was confused. he didn\'t listen to the cow and jumped anyway.the dog laughed at tim and started to pull his cloud away. the cow saw this and said, "go away, frog!" tim said, "no, i must not hold the cloud." the cow followed the dog. they did not see the big cow again. the shell fell on the cow\'s head and the horse. tim and the dog ran away, sad that mosquito had hit him.the next day, the dog came back with birds. the dog told the birds, "this is not a toy. it is a wishing book for you." the farmer heard the dog and said, "that is not good. the fish study them and we are all busy. i warn you, bees sing, or you can fly away. be careful next time."tim felt bad and said, "i leo, i will not listen for me now." the cow learned the bunny\'s words. she sang birds, she moved away and played with each other. on the water and saved her up to overcome crabs. she just like tim. the turtle, she said softly, but she felt better.tim learned that being close to be safe and enjoyed her little tom.the moral of other kids. each other is more patient and silly, like tim and careful in playtime instead. hidden toys that made everyone about danger. what with a machine is important than him. first the girl helped her left her friends and talk to them

и все в таком духе.

Средний перплексити на сыром выходе получился 42.55677372217178. As the kids say, it ain't much, but it's honest work. В принципе все. Модель в папке LLAMA, обучение в run.ipynb, эвал в eval.ipynb, чекпойнт в checkpoint.ipynb.