

1) Можно ли сказать, что сеть не обучилась корректно, и поэтому не генерирует осмысленный текст?

Это зависит от того, что мы понимаем под некорректностью сети. Можно сказать, что сеть не самым лучшим образом генерирует текст, однако говорить о том, что она построена некорректно нельзя. Так как в данном случае мы видим, что сеть выполняет свою задачу, она генерирует символ по последовательности и смогла обучиться некоторым словам. Мы можем продолжить обучение или добавить новые слои, возможно тогда сеть станет выдавать результат лучше.

2) Что такое бутстреп?

Бутстреп (bootstrap) – это процедура генерации повторных случайных выборок из исходного набора данных.

Бутстреп-процедура состоит в том, чтобы случайным образом многократно извлекать повторные выборки из эмпирического распределения. Конкретно, если мы имеем исходную выборку из n членов $x_1, x_2, \dots, x_{n-1}, x_n$, то с помощью датчика случайных чисел, равномерно распределенных на интервале $[1, n]$, можем вытянуть из нее произвольный элемент x_k , который снова вернем в исходную выборку для возможного повторного извлечения. Такая процедура повторяется n раз и образуется бутстреп-выборка, в которой одни элементы могут повторяться два или более раз, тогда как другие элементы - отсутствовать. Бутстреп-выборки производятся равномерно и с возвращением, поэтому некоторые исходные примеры будут отсутствовать, а другие - дублироваться: в среднем одна такая выборка содержит около $2/3$ уникальных исходных наблюдений.

Вероятность того, что конкретное наблюдение не войдет в бутстреп-выборку размера n , равна $\left(1 - \frac{1}{n}\right)^n$ и стремится к $\frac{1}{e} = 0.368$ при $n \rightarrow \infty$.

3) За что отвечает параметр *min* в строке `"checkpoint = ModelCheckpoint(filepath, monitor='loss', verbose=1, save_best_only=True, mode='min')"`?

За то, чтобы в лучшей модели было самое маленькое значение `loss`. Так как параметр `"mode=min"` отвечает за минимизацию отслеживаемого значения `"monitor=loss"`.