

1. Для чего может добавляться шум к выходным меткам?

Добавление шума к выходным меткам, то есть придание значениям меток некоторого небольшого случайного отклонения, может сделать модель более устойчивой к незначительным ошибкам, из-за которых она могла обучаться неверно. Особенно полезны подобные подходы при небольших размерах обучающих выборок (в этих случаях часто появляется проблема переобучения), так как эти подходы призваны также искусственно разнообразить выборки.

2. Какой геометрический смысл имеет регуляризация весов?

Регуляризация весов призвана исправить проблему переобучения путем наложения «штрафов» на веса со слишком большими значениями (принудительное занижение этих значений). На графиках это выглядит как установление некоторых границ значений, за которые не может выйти обучающаяся модель.

3. Для чего используется LearningRateScheduler?

Это один из Callback-ов в TF.Keras, который предназначен для предоставления возможности изменения значения learning rate на любой эпохе обучения. Так как в процессе обучения нас интересует «место» (конфигурация сети), где программа выдает наименьшие потери, то при высоком значении скорости обучения мы можем «проскочить» этот момент и это место. При низкой же скорости можно «застрять» в локальном минимуме. Также обучение в таком случае занимает больше времени. Компромиссный вариант – это, для примера, задать изначально более высокую скорость, а при снижении потерь – уменьшать ее, что позволит быстро и эффективно найти то самое «место» с наименьшими потерями. LearningRateScheduler позволяет реализовывать подобные подходы при обучении сетей.