

1. Что такое L1 регуляризация?

L1 регуляризация - это метод, позволяющий уменьшить сеть за счет исключения признаков, которые не влияют на конечный результат. Для реализации данного метода для каждого веса w к целевой функции прибавляется сумма весов в модуле $|w|$, таким образом, если веса на расстоянии V близки к 0, то они заменяются 0 и игнорируются. В итоге отбираются только важные признаки.

2. Что такое слой MaxPooling в сверточных нейронных сетях?

Слой MaxPooling нужен для уменьшения размерности изображения. Изображение делится на блоки и для каждого вызывается функция MaxPooling(функция максимума). Этот слой нужен для уменьшения изображения для того чтобы последующие свертки обрабатывали большую часть исходного изображения, а также для облегчения вычислений.

3. Что такое овражный эффект во время обучения ИНС?

Особенность применения метода градиентного спуска.

Если точка лежит на склоне оврага, то направление спуска из этой точки будет почти перпендикулярным к направлению дна оврага и в результате точки последовательности, полученные градиентным методом будут поочередно находиться то на одном то на другом склоне оврага. Если склоны оврага достаточно круты, то такие скачки со склона на склон точек могут сильно замедлить сходимость градиентного спуска.

4. Почему Вы посчитали, что на 10 эпохах началось переобучение, а не какая-то случайность в инициализации весов?

Запустила несколько раз, графики были похожи: с сильно колеблющейся ошибкой проверочных данных после этой эпохи(примерно), при этом тренировочная ошибка продолжает уменьшаться. То есть сеть очень хорошо распознает данные при обучении(ошибка стремится к 0), но при этом плохо распознает другие.

5. Сколько весов в Вашей модели?

```
model.add(Dense(64, activation='relu', input_shape=(train_data.shape[1],)))  
model.add(Dense(64, activation='relu'))  
model.add(Dense(1))
```

Данная модель содержит $\text{train_data.shape}[1] * 64$ весов между входным и первым слоем, $64 * 64$ весов между первым и скрытым слоем, 64 весов к выходному слою

Итого: $(\text{train_data.shape}[1] * 64 + 64 * 64 + 64)$

6. Каким образом Вы нормируете данные?

```
mean = train_data.mean(axis=0)  
std = train_data.std(axis=0)  
train_data -= mean  
train_data /= std  
test_data -= mean  
test_data /= std
```

Для каждого признака во входных данных (столбца в матрице входных данных) из каждого значения вычитается среднее по этому признаку(mean), и разность делится на стандартное отклонение(std), в результате признак центрируется по нулевому значению и имеет стандартное отклонение, равное единице

7. Почему в качестве метрики Вы выбрали mae ?

MAE - абсолютное отклонение значения от ожидаемого результата(средняя абсолютная ошибка), используем ее, так как в лабораторной работе решается задача регрессии, для которой нельзя использовать понятие точности.