

### **1. Что такое операции свертки в сверточных нейронных сетях?**

Свертка – прохождение частей изображения через слой одинаковых нейронов. Результат записывается в новое изображение, называемое картой признаков.

### **2. Можно ли для задачи регрессии в качестве метрики использовать бинарную кроссэнтропию?**

Нельзя, т.к. задача регрессии – предсказание какого-то вещественного значения на основе начальных данных, бинарная кроссэнтропия же применяется для задач бинарной классификации, где нужно предсказать, к какой из заранее заданных групп принадлежит каждый элемент.

### **3. Что такое коэффициент регуляризации?**

Регуляризация используется в нейронных сетях для борьбы со слишком большими весами сети и переобучением. Для этого вводится регуляризатор:

$\|w\|^2 = \sum_{i=1}^d w_i^2$  или  $\|w\| = \sum_{i=1}^d |w_i|$ . И тогда вместо задачи  $Q(w, X) \longrightarrow \min$  будем

решать задачу  $Q(w, X) + \alpha R(w) \longrightarrow \min$ , где  $R(w)$  – регуляризатор,  $\alpha$  – коэффициент регуляризации. При этом необходимо найти такое  $\alpha$ , чтобы модель не получилась слишком сложной (снова переобучение) или слишком простой (тогда от весов не будет смысла).

### **4. Гарантируется ли, что для данной задачи при одной и той же архитектуре сети, переобучение будет наблюдаться при 18 эпохах?**

Нет, т.к. данная архитектура может оказаться слишком сложной или слишком простой для других данных.

### **5. Какая функция позволяет применить к модели новые данные?**

Функция `fit` позволяет обучить нашу модель. При этом, если мы вызовем эту функцию на уже обученную модель, то обучение продолжится с той точки, на которой закончилось, таким образом можно добавлять новые данные в уже обученную сеть.

**6. Имеет ли смысл применять проверку по k блокам при большом размере датасета?**

Скорее не имеет, т.к. кросс-валидация и хороша тем, что может на небольшом размере датасета получать более лучшее качество обучения. Если же размер датасета будет большим, то качество результата не будет стоить затраченных на него вычислительных мощностей.

**7. Почему обучение идет на основании mse, но анализируется mae?**

Функция потерь mse используется Keras'ом для оптимизации модели во время обучения и это она минимизируется моделью. mae же здесь – метрика, которая используется для оценки производительности. mae удобней для анализа т.к. является по сути разницей между истинным и прогнозируемым значением; по mae легче судить о точности обученной модели.