

1. Каким образом можно представить датасет изображений в формате RGBA?
Данный датасет можно представить в виде тензора 4-го ранга со следующими осями: высота, ширина, каналы цвета и образцы.
2. Что такое скорость обучения?
Скоростью обучения называется шаг градиентного спуска.
3. В чём разница методов обучения второго порядка от методов обучения первого порядка?
В методах обучения первого порядка используется только вектор наибольшего роста функции потерь, в методах второго порядка также используется информация о кривизне функции потерь.
4. Чем можно обосновать колебания для контрольной выборки на графиках у Вас в отчете?
Колебания на графиках можно обосновать частыми попаданиями в локальные минимумы при оптимизации на основе градиентного спуска, движение из которых увеличивает потери.
5. На строчках 25-29 Вы создаете модель. Можете сказать, сколько в ней слоев и сколько нейронов на каждом слое?
В данной модели три слоя: входной – 60 нейронов, скрытый – 15 нейронов, выходной – 1 нейрон.
6. Какой тип данных имеет history.history?
Типом объекта является словарь
7. Что такое callback в функции fit?
Это функция представляющая из себя обратный вызов на определенных этапах обучения, в моём случае это ModelCheckpoint, который сохраняет модель после каждой эпохи обучения.