

### 1. Для чего нужен слой MaxPooling в сверточных нейронных сетях?

Пулинговые слои, в частности, слой с функцией максимума (max pooling) используется для уменьшения размерности карт признаков предыдущего слоя. Так как в изображениях соседние пиксели не сильно друг от друга отличаются, то при замене нескольких соседних пикселей на один агрегат (максимальный по признаку), значительной потери информации не произойдет. Следовательно это позволит ускорить работу нейронной сети, а также позволит увеличить инвариантность сети к небольшим преобразованиям входных данных.

### 2. В чем особенность методов обучения второго порядка?

Методы обучения второго порядка для аппроксимации функции в точке учитывают вторую производную (используют матрицу Гессе).

### 3. Что такое коэффициент момента в методах обучения?

Коэффициент момента – число в интервале  $(0, 1)$ , умножаемое на величину предыдущего изменения веса. Обычно равно 0,9. Используется для добавления к коррекции веса члена, пропорционального величине предыдущего изменения веса. Такой метод называется методом импульса, он увеличивает устойчивость алгоритма обучения и увеличивает скорость сходимости.

### 4. Почему в лаб. работе анализируется mae?

Так как в лабораторной ИНС решает задачу регрессии, для которой не применимо понятие точности (ИНС определяет не дискретное значение исследуемого параметра), для оценки качества сети в ходе корректировки ее параметров используется абсолютное значение разности между предсказанными и целевыми значениями (mae).

### 5. Как называется вид нормализации используемый в лаб. работе?

Такая нормализация называется нормализацией стандартным отклонением. В результате такой нормализации распределение признака по нулевому значению и имеет стандартное отклонение, равное единице.

#### 6. Почему количество блоков выбрано 8?

Так как мы имеем небольшой набор данных, набор для проверки ИНС получился бы слишком маленьким, следовательно, оценки сети имели бы большой разброс, т.е. сильно смещенные. Разбив исходные данные на 2, 4, 6 и 8 блоков, при применении перекрестной проверки я оценивал, как меняется средняя ошибка по эпохам. Наименьший разброс оценки у меня получился на графике для 6 блоков.

#### 7. Какое суммарное количество весов в Вашей модели?

Входных признаков 13, в первом (входном) слое 64 нейронов, во втором (скрытом) 64, в третьем (выходном) – 1. Слои полносвязные. Всего весов:  $13 \cdot 64 + 64 \cdot 64 + 64 \cdot 1 = 57\,408$