

1. Что такое кодер-декодер архитектура?

Архитектура, состоящая из двух моделей: одна для считывания входной последовательности и кодирования ее в вектор фиксированной длины, а вторая для декодирования вектора фиксированной длины и вывода предсказанной последовательности. Используется, например, для сегментации изображений, тогда Encoder преобразует входное изображение в многомерное представление признаков и выполняет функцию извлечения признаков, а Decoder создает сегментированное изображение на основе признаков, полученных от сверточной сети.

2. Какие задачи решаются обучением с учителем?

Решается следующая задача. Имеется множество объектов (ситуаций) и множество возможных ответов (откликов, реакций). Существует некоторая зависимость между ответами и объектами, но она неизвестна. Известна только конечная совокупность прецедентов – пар «объект, ответ», называемая обучающей выборкой. На основе этих данных требуется восстановить зависимость, то есть построить алгоритм, способный для любого объекта выдать достаточно точный ответ.

Обучение с учителем может использоваться для решения двух типов задач: задача регрессии и задача классификации.

3. Какие проблемы возникают при семантической сегментации изображений?

1. Уменьшение пространственного разрешения.
2. Существование объекта в различных масштабах.
3. Уменьшение локальной точности.

4. В отчете приведены тесты? Как по результатам понять, что классификация произошла правильно?

В отчете приведены тесты. По пользовательскому изображению, с помощью функции `model.predict` строится вектор, который по сути является вектором вероятностей быть каждой из цифр. Если вероятность для загаданной

цифры наибольшая, и это подтверждается на нескольких тестах, то можем считать, что классификация произошла правильно.

5. Для чего нужен слой Flatten?

Этот слой преобразует формат изображения из двумерного массива (28 на 28 пикселей) в одномерный (размерностью $28 * 28 = 784$ пикселя). Слой извлекает строки пикселей из изображения и выстраивает их в один ряд.

6. Почему в первом скрытом слое у Вас 1024 нейрона? Чем это обусловлено?

Я проверял различные архитектуры, и пришел к выводу, в данной работе лучше всего по точности себя показывает архитектура с единственным скрытым слоем. Увеличивая количество нейронов на нем, увеличивалась и точность. При увеличении количества нейронов более чем на ~ 1024 , рост точности был незначителен, и уже был достигнут результат в $\sim 98\%$ точности, поэтому я остановился на этом числе.

7. При каком методе обучения был достигнут лучший результат?

Не уверен, правильно ли понял вопрос, но в работе сравнивались различные оптимизаторы и их параметры, их влияние на процесс обучения. В итоге выяснилось, что оптимизаторы Adam, RMSprop, Adagrad и Nadam дают примерно одинаковый результат на архитектуре с одним скрытым слоем на 1024 нейрона, размером батча 128 и 5-ю эпохами. Беспристрастным путем был выбран базовый оптимизатор Adam и протестированы его параметры обучения. В итоге лучший результат достигается на базовых параметрах оптимизатора Adam и вышеописанной архитектуре сети.