

ОТВЕТЫ НА ВОПРОСЫ К 5 ЛАБОРАТОРНОЙ.

Что такое проблема долговременной зависимости в рекуррентных сетях?

Это проблема утраты способности использовать накопленную информацию РНС-ю при увеличении «расстояния» между подаваемыми на вход данными. Например, искажение информации при потере удаленного контекста в тексте.

Что такое блок GRU?

GRU - управляемый рекуррентный нейрон. Используется в вариации сети с долгой краткосрочной памятью (LSTM, которая, кстати, создана для решения вышеописанной проблемы, используя фильтры и явно заданную клетку памяти). У них на один фильтр меньше (2), и связи реализованы иначе. Фильтр обновления определяет, сколько информации останется от прошлого состояния и сколько будет взято из предыдущего слоя. Фильтр сброса работает примерно как забывающий фильтр.

Почему при меньшем ядре точность выше? Как это можно обосновать?

Должно быть, исходные признаки не такие расплывчатые, как требуют ядра с бóльшим размером. Также как и на изображениях с плохим разрешением порой сложно распознать объект, так и здесь.

Что происходит, если часть свертки выходит за границу изображения?

Нулевое дополнение. Недостающие пиксели изображения полагаются равными нулю.