

В итоговом тексте очень мало слов, большая часть — это просто набор символов. Получается, что модель обучена некорректно?

В данной работе сеть генерирует символ по последовательность символов, выбирая наиболее вероятный. Сеть свою работу выполнила, хоть и не сгенерировала много слов. Если обучать ее на большем количестве эпох, возможно, она сгенерирует последовательность слов. Или можно попробовать добавить случайную составляющую для выбора символа.

Что такое Бэггинг?

Для начала стоит пояснить, что такое Бутстрэп. Простыми словами мы из выборки равновероятно достаем n объектов (возможны повторения, т.е. можно представить, что мы достаем из мешка предмет и кладем его обратно) и повторяя m раз создаем m подвыборок.

С помощью бутстрепа создаем m выборок и на каждой обучим свой классификатор. Итоговый классификатор будет усреднять ответы всех этих алгоритмов в случае регрессии или вычисляется мода в случае классификации.

Бэггинг уменьшает средний квадрат ошибки и предотвращает переобучение.

Почему выбрано именно такое значение прореживания?

Как правило между скрытыми слоями используется прореживание в диапазоне 0.2 – 0.5. Такое прореживание было выбрано, чтобы избежать переобучения, но в то же время исключить не много нейронов. Данное прореживание с задачей справилось.

