

### 1. Что такое Бэггинг?

Bagging в ансамблевом машинном обучении использует несколько слабых моделей, объединяя прогнозы, чтобы выбрать лучший прогноз для каждого участка пространства признаков, т.к. каждая из слабых моделей специализируется (прогнозирует точнее прочих) на отдельном участке пространства признаков.

2. В недостаток того, если бы сеть училась предсказывать не символы, а не слова?

В чем недостаток того, что сеть училась предсказывать символы, а не слова?

Недостатки предсказания слов вместо символов?

character-level архитектуры используют значительно меньше памяти, так как они не требуют большой embedding матрицы, как требовали бы слова. Кроме того, они могут использовать информацию подслов, например морфемы.

word-level модели все же на text prediction обычно работают лучше, но они должны быть действительно огромными, чтобы охватить все возможные слова. Сложнее заставить все знаки препинания появляться в сгенерированном/предсказанном тексте. Также усложняется ситуация для морфологически богатого языка, как русский или турецкий.

character-level модели мощны, они могут найти применение в условиях, где память ограничена, например, в мобильных устройствах. Кроме того, если дать модели больше параметров (например, добавить сверточные слои и больше сверточных фильтров), это может еще больше повысить ее результаты.

### 3. Для чего используется метод epoch\_predict?

Среди приложенных мной файлов либо TextGeneration.py лишний, либо надо было просто импортировать его метод в файл lr8.py. Он генерирует последовательность символов: берется та же модель, что и при непосредственном обучении, в нее загружаются предварительно записанные в

файл веса с последней эпохи, и на основе этих весов генерируется последовательность символов.