

### Вопрос 1

Выполнен

Баллов: 1,00 из 1,00

Что такое teacher forcing?

Выберите один ответ:

- ☐ a. алгоритм изменения learning rate в процессе обучения сети для ускорения сходимости
- ☐ b. метод для обучения RNN, который всегда использует в качестве входа ground truth с предыдущего шага, а не предсказанное сетью значение
- ☒ c. метод для обучения RNN, который в некотором (заранее зафиксированном) проценте случаев использует в качестве входа ground truth с предыдущего шага, а не предсказанное сетью значение
- ☐ d. архитектура нейронной сети для машинного перевода

Ваш ответ верный.

[1] Yao Z. et al. [Stagc: A systematically mined question-code dataset from stack overflow](#) //Proceedings of the 2018 World Wide Web Conference. – 2018. – С. 1693-1703.

## Вопрос 2

Выполнен

Баллов: 1,00 из 1,00

Выберите правильные утверждения про загрузку данных через `torchtext.data.TabularDataset`.

Выберите один или несколько ответов:

- ☐ a. Обязательными параметрами являются `path`, `format`, `fields`, `skip_header`
- ☒ b. Неважно, заглавными или маленькими буквами прописывать формат входных данных в поле `format` (оба варианта `format='csv'` и `format='CSV'` сработают)
- ☒ c. В качестве формата входных данных можно использовать CSV, TSV или JSON
- ☐ d. При использовании `torchtext.data.TabularDataset.splits(...)` нужно обязательно указать пути до файлов с обучающей, валидационной и тестовой выборками (параметры `train`, `validation`, `test`). Поделить выборку две части (`train` и `test`) нельзя

Ваш ответ верный.

decode (декодирование), но, при этом, почему-то она здесь поставила слова "soup". Возможно, как-то это коррелирует с библиотекой [beautifulsoup](#) и, как-то, нашей сети показалось, что это здесь будет уместно. Если мы посмотрим на матрицу attention, то и здесь важно было только слово "UTF" — кажется, что если говорить про декодирование то слово "UTF" очень хорошо связано со словом "soup" (видимо, это отсылка к "beautifulsoup") и со словом "decode", с остальными скобками и "unknown token" немного — тоже. В предыдущем примере матрица attention выглядела примерно так же, то есть — слово "string", которое здесь было, видимо, ключевым, по мнению нашей сети было связано достаточно плотно, практически, со всеми остальными символами в нашем примере. Но учитывая, что это не слишком удачный пример, эта матрица не несёт достаточно информации.

### Вопрос 3

Выполнен

Баллов: 1,00 из 1,00

Формула перплексии:

$$PP(W) = P(w_1, w_2, \dots, w_N)^{-\frac{1}{N}} = \sqrt[N]{\frac{1}{P(w_1, w_2, \dots, w_N)}} = \sqrt[N]{\frac{1}{\prod_{i=1}^N P(w_i | w_1, \dots, w_{i-1})}}$$

Выберите верные утверждения про перплексию:

Выберите один или несколько ответов:

- ☒ a. Можно использовать перплексию для сравнения двух нейросетей с разными архитектурами на одной и той же задачи (обе нейросети обучены переводить текст с языка А на язык Б)
- ☐ b. Чтобы посчитать перплексию обязательно нужен ground truth (например, написанный человеком перевод текста). Другими словами, перплексию нельзя посчитать на данных без разметки.
- ☐ c. Можно использовать перплексию для сравнения нейросетей с одинаковой архитектурой, где первая сеть решает задачу перевода с языка А на язык Б, а вторая с языка В на язык Г
- ☒ d. Перплексия связного текста ниже, чем перплексия произвольного набора слов

Ваш ответ верный.