

Написать генератор текста и обучить его на баснях Эзопа.

## Данные

Будем использовать сборник басен Эзопа на английском языке, который можно найти [здесь](#).

Скрипт для загрузки (поместить в .sh файл):

```
FILE=$1
```

```
NAME=$2
```

```
URL=http://www.gutenberg.org/cache/epub/$FILE/pg$FILE.txt
```

```
TARGET_DIR=./data/$NAME/
```

```
mkdir $TARGET_DIR
```

```
TXT_FILE=./data/$NAME/data.txt
```

```
wget -N $URL -O $TXT_FILE
```

## Подготовка данных

1. перевести все в нижний регистр
2. разделить рассказы между собой группой символов: |||
3. применить токенайзер

## Модель

На pytorch имплементировать следующую архитектуру:

Layer (type)	Output Shape	Param #
input_1 (InputLayer)	(None, None)	0
embedding_1 (Embedding)	(None, None, 100)	416900
lstm_1 (LSTM)	(None, 256)	365568
dense_1 (Dense)	(None, 4169)	1071433
Total params: 1,853,901		
Trainable params: 1,853,901		
Non-trainable params: 0		

Обучить данную модель на наборе басен Эзопа

Написать скрипт инференса.

Пример задания на keras.

**Требования:**

- Репозиторий github/gitlab
- Мониторинг обучения: ClearML или Weights & Biases. Допустимо использование TensorBoard
- Поместите в README графики и ваши выводы по успешности обучения на их основе
- Проводить эксперименты вы можете в ноутбуках, но окончательный код обучения и инференса желательно оформить в .py файлах
- Неудавшиеся эксперименты так же будут засчитаны, главное указать выводы, почему эксперимент не удался, что и как нужно изменить