**Практическая работа «Генерация текста»**

* [Полезная презентация с подсказками](https://marigostra.ru/materials/presentation-nlp-lab05.pdf)

Используя модель RuGPT от Сбера, необходимо реализовать возможность генерации текста по заданному промпту. Допускается использование как старой модели rugpt3large\_based\_on\_gpt2, так и новой ruGPT-3.5-13B.

Каждому студенту преподавателем будет дана пара слов, и требуется подобрать промпт таким образом, чтобы выданная пара слов встречалась бы в сгенерированном тексте с учётом порядка и с учётом словоформ (как в предыдущей работе). Допускается использовать обе модели, но пара слов преподавателем подбирается на rugpt3large\_based\_on\_gpt2. Ограничений на значения параметров нет (даже на длину генерируемого текста), но подобранный промпт не должен содержать искомые слова ни в одной из их словоформ.

Функция generate для практической работы «Генерация текста»:

def generate(

model, tok, text,

do\_sample=True, max\_length=100, repetition\_penalty=5.0,

top\_k=5, top\_p=0.95, temperature=1,

num\_beams=None,

no\_repeat\_ngram\_size=3

):

input\_ids = tok.encode(text, return\_tensors="pt")

print(model.generate.\_\_globals\_\_['\_\_file\_\_'])

out = model.generate(

input\_ids,

max\_length=max\_length,

repetition\_penalty=repetition\_penalty,

do\_sample=do\_sample,

top\_k=top\_k, top\_p=top\_p, temperature=temperature,

num\_beams=num\_beams, no\_repeat\_ngram\_size=no\_repeat\_ngram\_size

)

return list(map(tok.decode, out))