**Обучение без предобученных эмбеддингов:**

Config parameters:

batch\_size:16

d\_model:256

epochs:30

learning\_rate:0.00001

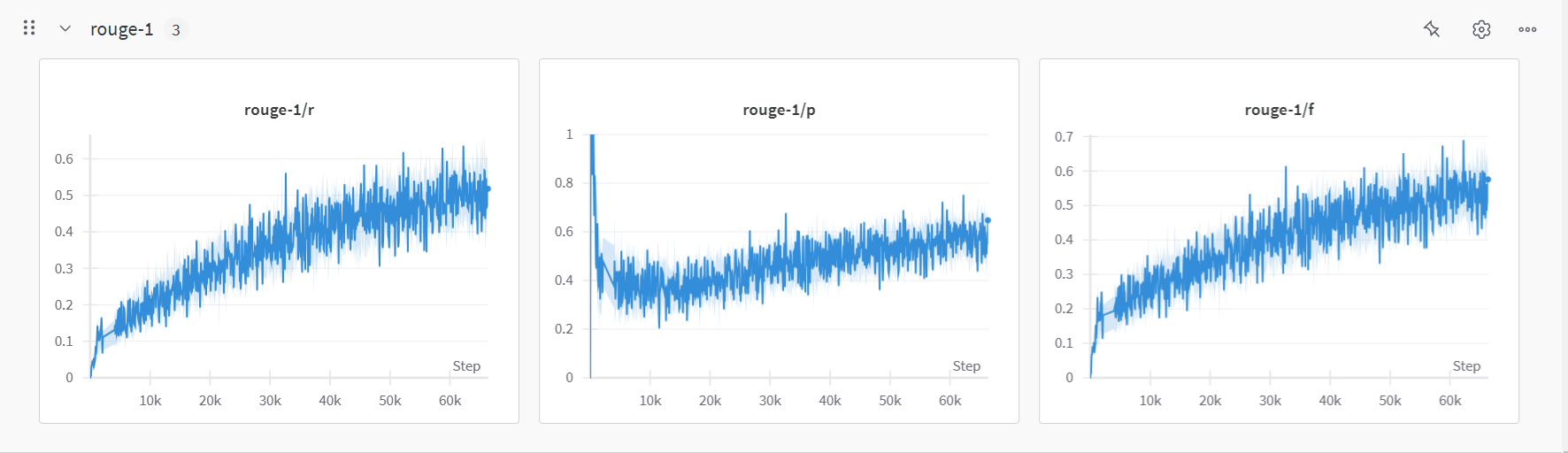
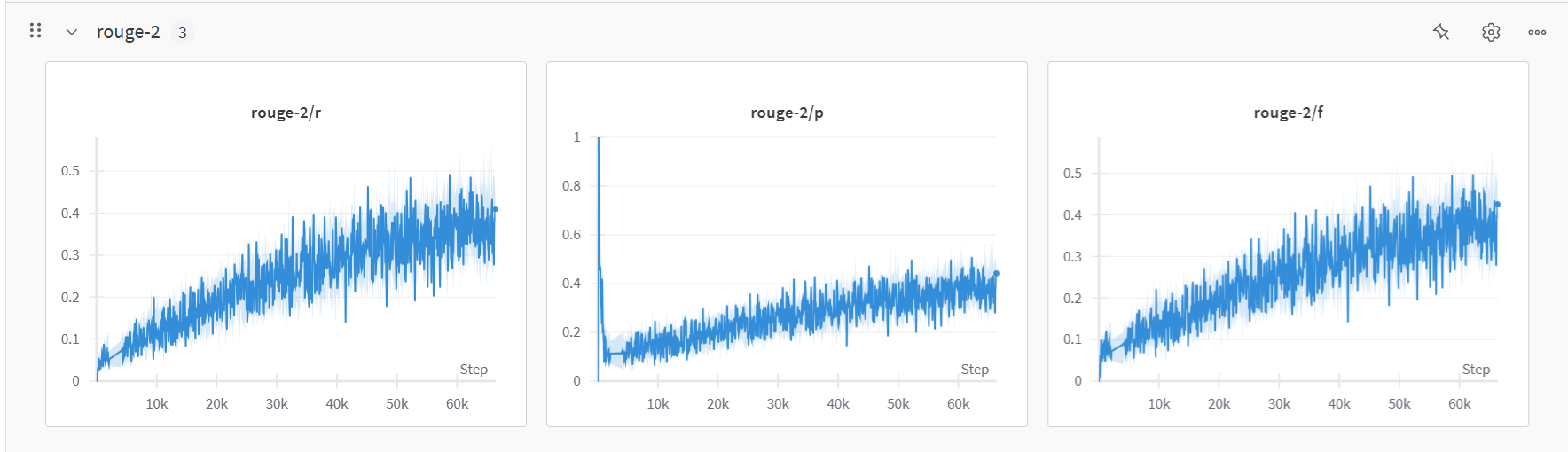
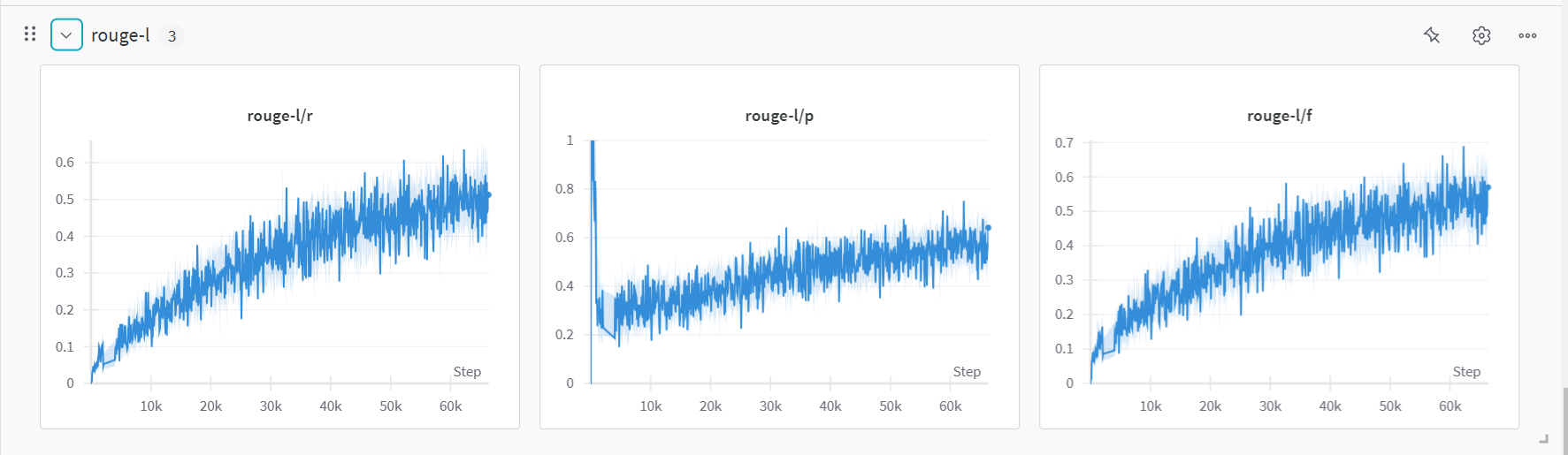
model:"EncoderDecoder"

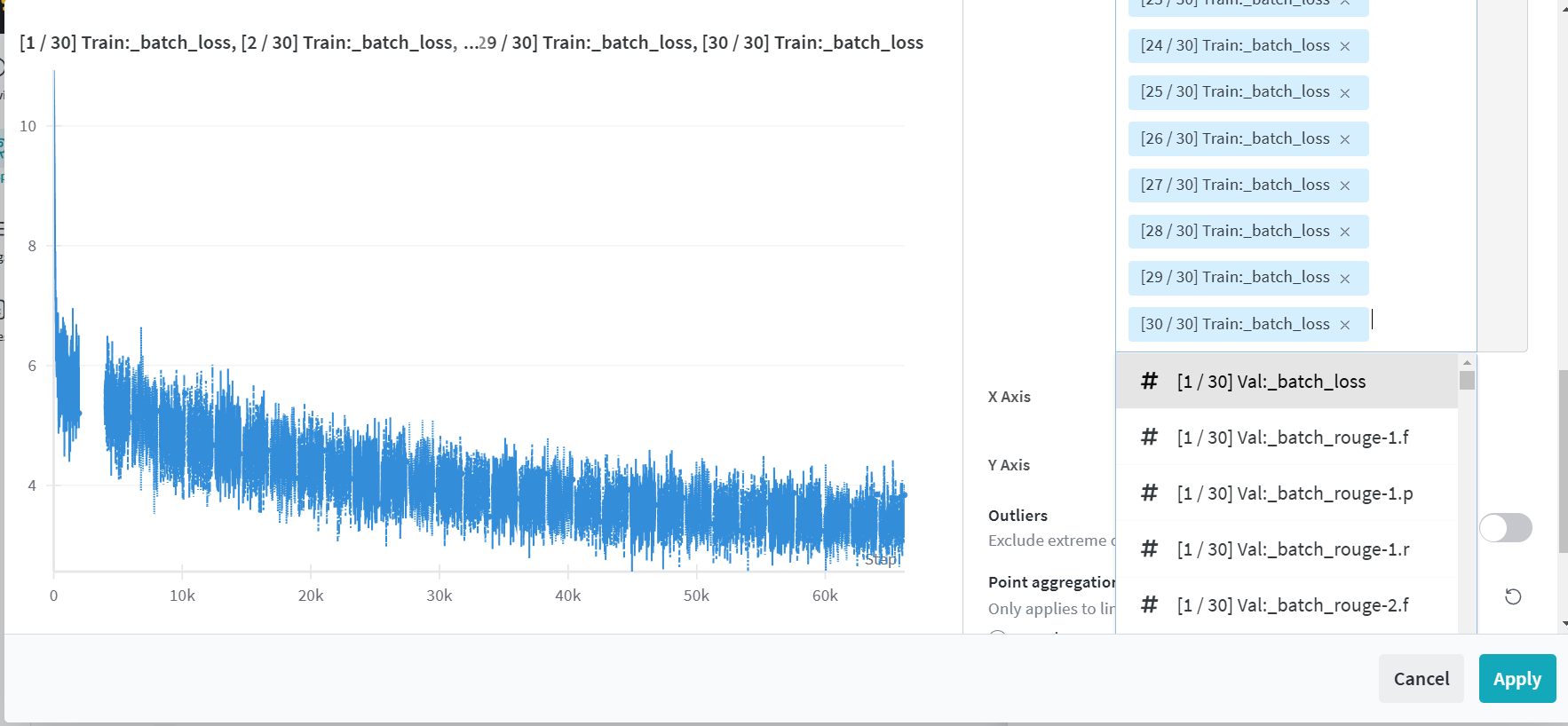
n\_heads:8

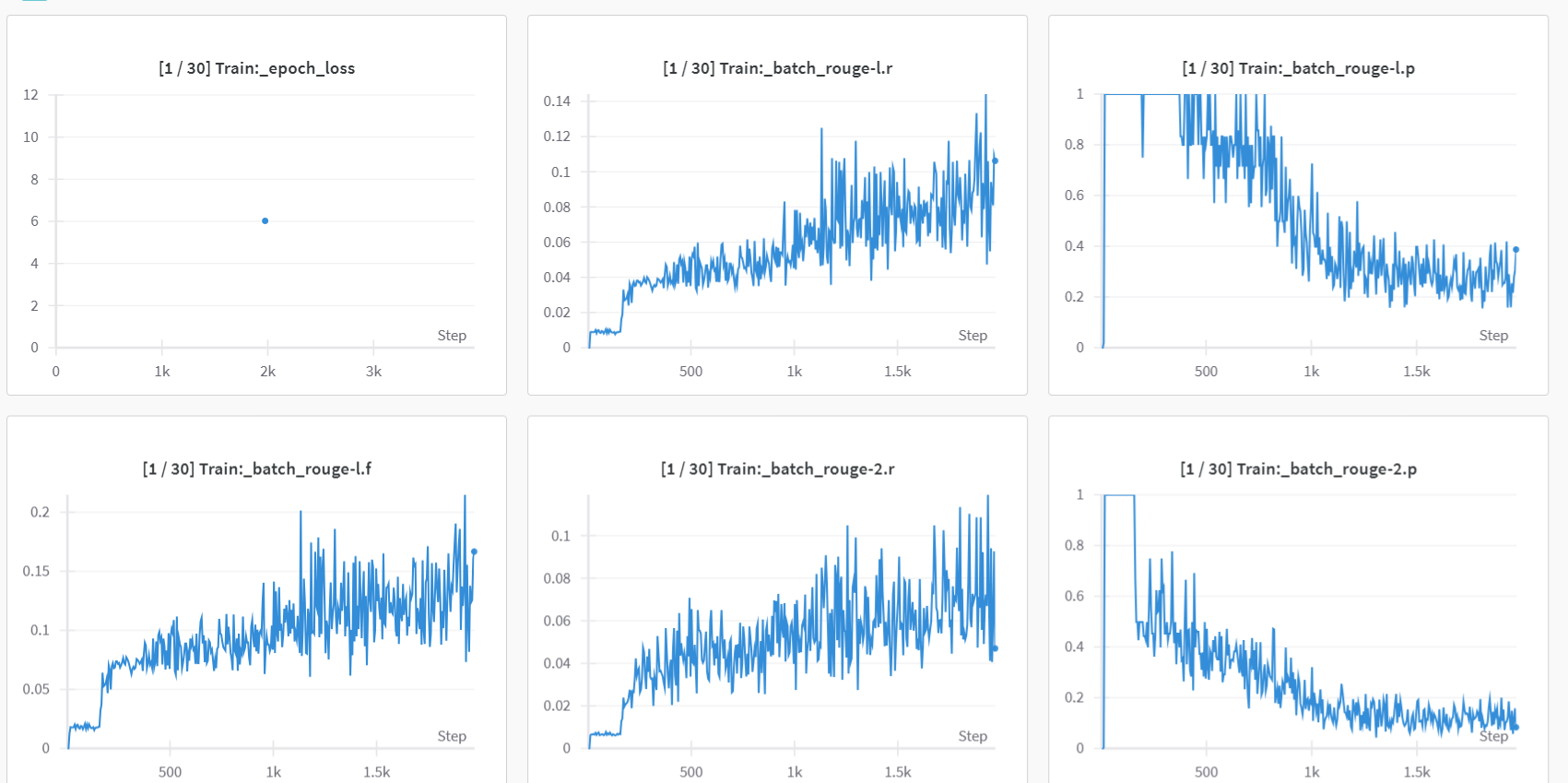


1`

Метрики по всем эпохам:



Batch\_loss по всем эпохам

1 эпоха, пример части графиков: 

**Обучение с предобученными эмбеддингами:**

Config parameters:

batch\_size:16

d\_model: 312, токенизация+эмбединги:

model\_name = "cointegrated/rubert-tiny2"

\_tokenizer = AutoTokenizer.from\_pretrained(model\_name)

\_model = AutoModel.from\_pretrained(model\_name).to(DEVICE)

epochs:10 + 10

learning\_rate:0.00001

model:"EncoderDecoder"

n\_heads:8

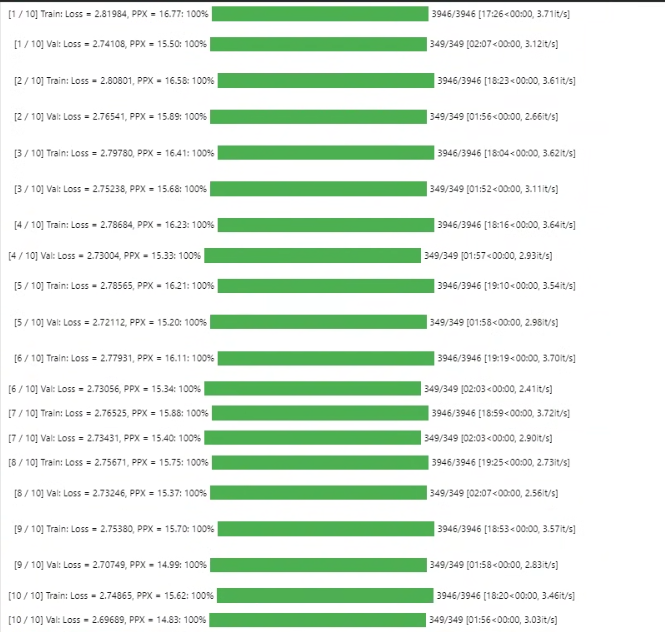
запустила обучаться 2 раза по 10 эпох (второй раз дообучала ту же модель). Заметна тенденция к уменьшению лосса, но не очень стремительная. Но каждую эпоху лосс немного, но падает. По итогу уже после 2 эпохи лосс оказался ниже, чем в случае обучения без предобученных эмбеддингов.

(Дообучение на еще 10-ти эпохах кажется не залогировалосьЮ ноу меня есть доказательства):

1 запуск 10 эпох:

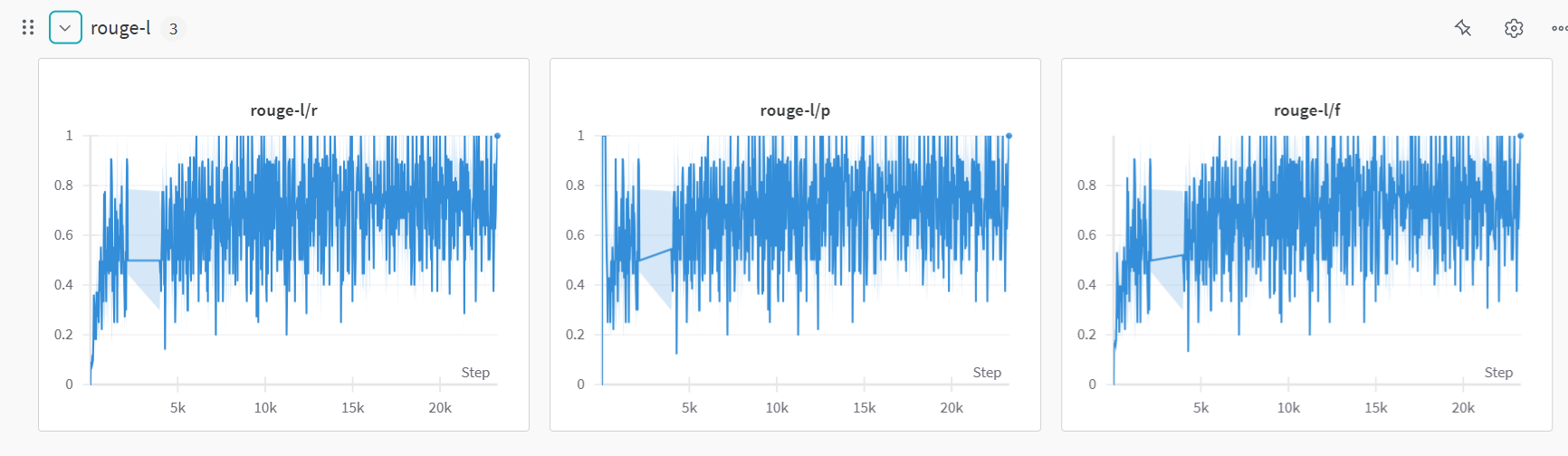


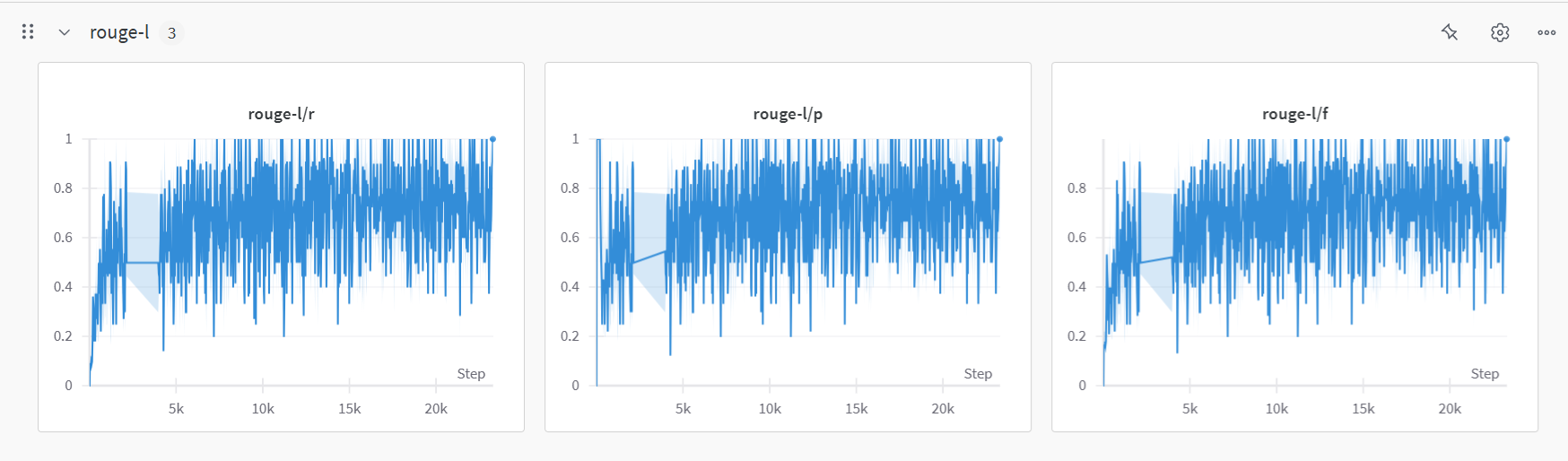
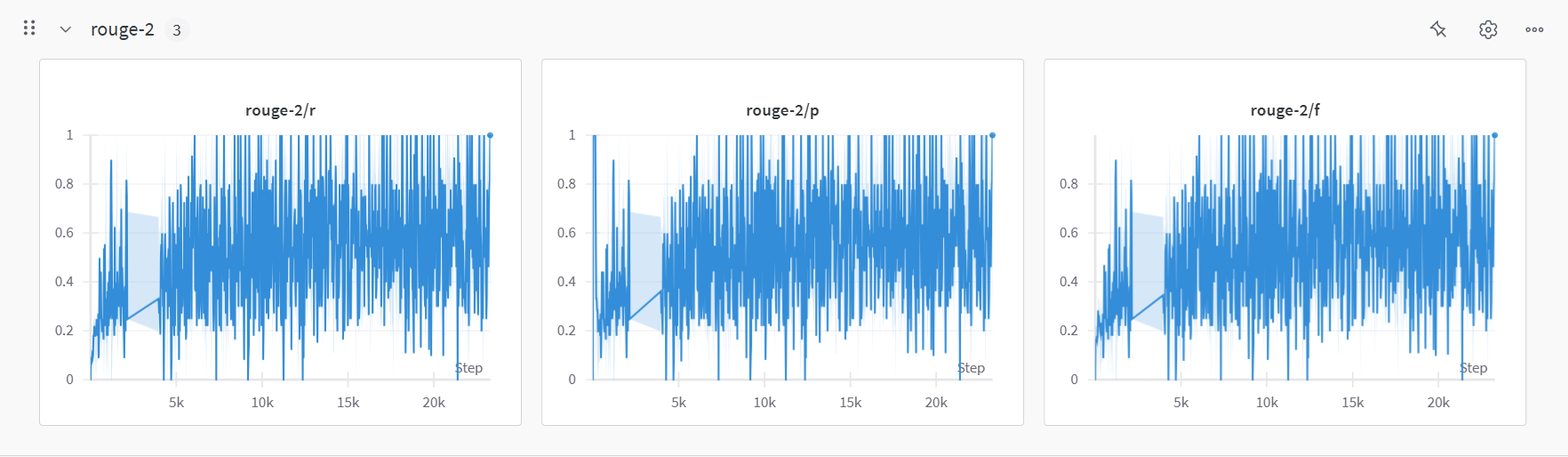
Второй запуск 10 эпох:



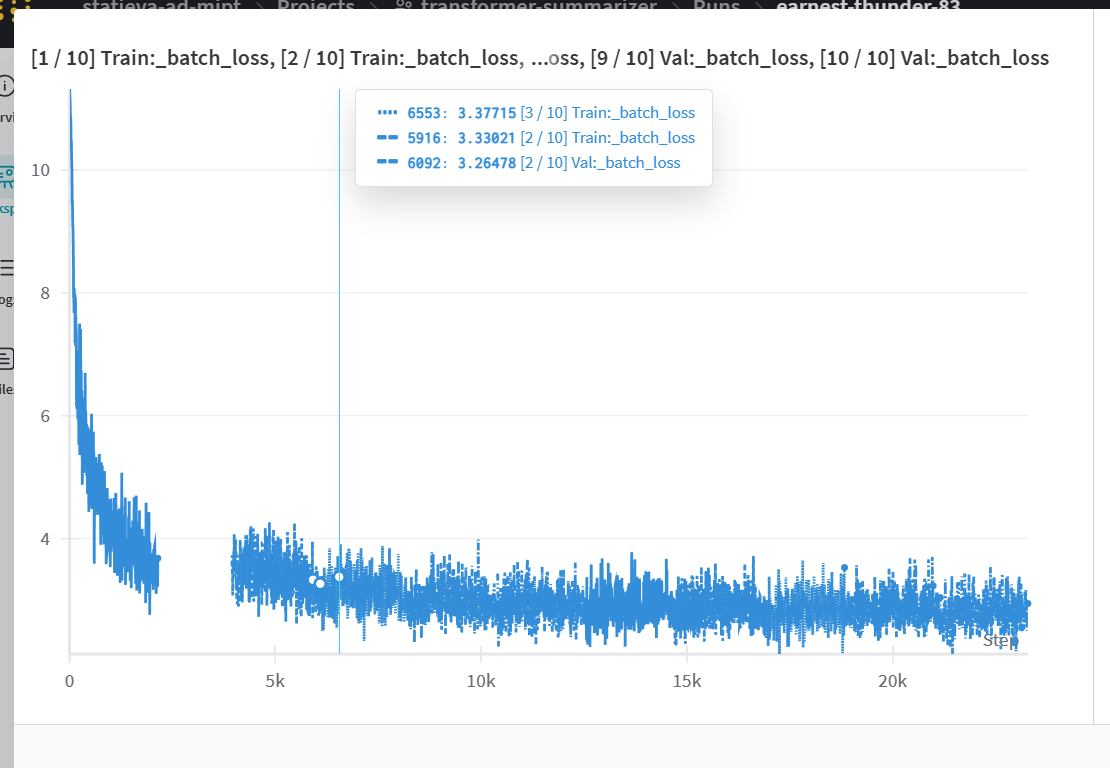
Графики для первых 10-ти эпох:

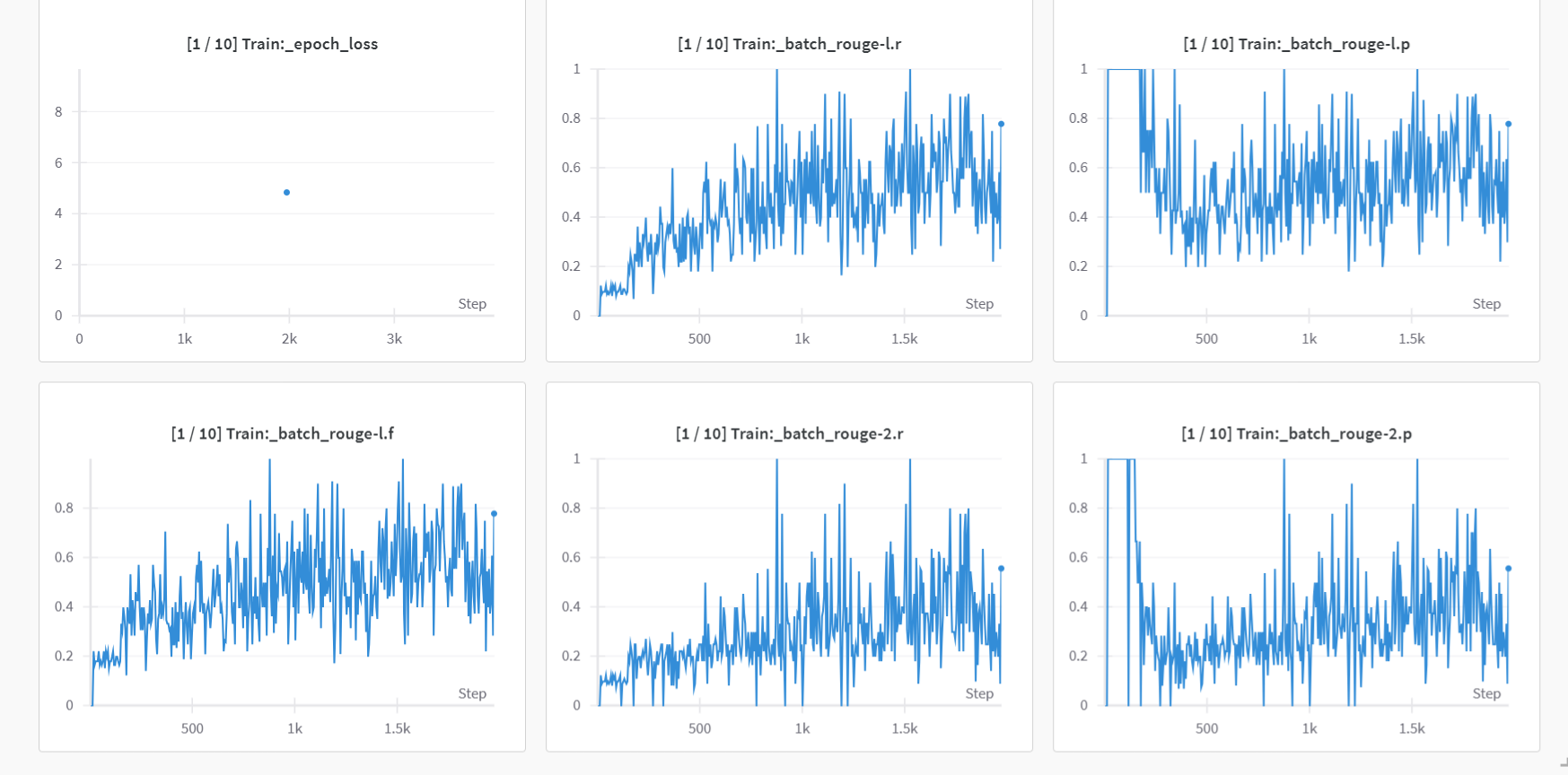
Метрики по всем эпохам:



Batch\_loss по всем эпохам(хз если честно что за дырка…):



1 эпоха, пример части графиков: 

Примеры предсказаний и визуализации механизма внимания в файле tests.ipynb

Я потратила очень много времени на эту домашку и я ей горжусь, какая бы она не была :D