



Deep Learning School

✓ Глубокое обучение. Часть 2

Домашнее задание по теме "Механизм внимания"

Это домашнее задание проходит в формате peer-review. Это означает, что его будут проверять ваши однокурсники. Поэтому пишите разборчивый код, добавляйте комментарии и пишите выводы после проделанной работы.

В этом задании вы будете решать задачу классификации математических задач по темам (многоклассовая классификация) с помощью Transformer.

В качестве датасета возьмем датасет математических задач по разным темам. Нам необходим следующий файл:

[Файл с классами](#)

Hint: не перезаписывайте модели, которые вы получите на каждом из этапов этого дз. Они ещё понадобятся.

✓ Задание 1 (2 балла)

Напишите кастомный класс для модели трансформера для задачи классификации, использующей в качестве backbone какую-то из моделей huggingface.

Т.е. конструктор класса должен принимать на вход название модели и подгружать её из huggingface, а затем использовать в качестве backbone (достаточно возможности использовать в качестве backbone те модели, которые упомянуты в последующих пунктах)

```
!pip install accelerate -U
```

```
➡ Requirement already satisfied: accelerate in /usr/local/lib/python3.11/dist
Collecting accelerate
  Downloading accelerate-1.6.0-py3-none-any.whl.metadata (19 kB)
Requirement already satisfied: numpy<3.0.0,>=1.17 in /usr/local/lib/python3.11/dist-packages (from accelerate)
Requirement already satisfied: packaging>=20.0 in /usr/local/lib/python3.11/dist-packages (from accelerate)
Requirement already satisfied: psutil in /usr/local/lib/python3.11/dist-packages (from accelerate)
Requirement already satisfied: pyyaml in /usr/local/lib/python3.11/dist-packages (from accelerate)
Requirement already satisfied: torch>=2.0.0 in /usr/local/lib/python3.11/dist-packages (from accelerate)
Requirement already satisfied: huggingface-hub>=0.21.0 in /usr/local/lib/python3.11/dist-packages (from accelerate)
Requirement already satisfied: safetensors>=0.4.3 in /usr/local/lib/python3.11/dist-packages (from accelerate)
Requirement already satisfied: filelock in /usr/local/lib/python3.11/dist-packages (from accelerate)
Requirement already satisfied: fsspec>=2023.5.0 in /usr/local/lib/python3.11/dist-packages (from accelerate)
Requirement already satisfied: requests in /usr/local/lib/python3.11/dist-packages (from accelerate)
Requirement already satisfied: tqdm>=4.42.1 in /usr/local/lib/python3.11/dist-packages (from accelerate)
Requirement already satisfied: typing-extensions>=3.7.4.3 in /usr/local/lib/python3.11/dist-packages (from accelerate)
Requirement already satisfied: networkx in /usr/local/lib/python3.11/dist-packages (from accelerate)
Requirement already satisfied: jinja2 in /usr/local/lib/python3.11/dist-packages (from accelerate)
Collecting nvidia-cuda-nvrtc-cu12==12.4.127 (from torch>=2.0.0->accelerate)
  Downloading nvidia_cuda_nvrtc_cu12-12.4.127-py3-none-manylinux2014_x86_64.whl (1.9 MB)
Collecting nvidia-cuda-runtime-cu12==12.4.127 (from torch>=2.0.0->accelerate)
  Downloading nvidia_cuda_runtime_cu12-12.4.127-py3-none-manylinux2014_x86_64.whl (59.5 MB)
Collecting nvidia-cuda-cupti-cu12==12.4.127 (from torch>=2.0.0->accelerate)
  Downloading nvidia_cuda_cupti_cu12-12.4.127-py3-none-manylinux2014_x86_64.whl (1.9 MB)
Collecting nvidia-cudnn-cu12==9.1.0.70 (from torch>=2.0.0->accelerate)
  Downloading nvidia_cudnn_cu12-9.1.0.70-py3-none-manylinux2014_x86_64.whl (669.0 MB)
Collecting nvidia-cublas-cu12==12.4.5.8 (from torch>=2.0.0->accelerate)
  Downloading nvidia_cublas_cu12-12.4.5.8-py3-none-manylinux2014_x86_64.whl (363.1 MB)
Collecting nvidia-cufft-cu12==11.2.1.3 (from torch>=2.0.0->accelerate)
  Downloading nvidia_cufft_cu12-11.2.1.3-py3-none-manylinux2014_x86_64.whl (193.1 MB)
Collecting nvidia-curand-cu12==10.3.5.147 (from torch>=2.0.0->accelerate)
  Downloading nvidia_curand_cu12-10.3.5.147-py3-none-manylinux2014_x86_64.whl (56.0 MB)
Collecting nvidia-cusolver-cu12==11.6.1.9 (from torch>=2.0.0->accelerate)
  Downloading nvidia_cusolver_cu12-11.6.1.9-py3-none-manylinux2014_x86_64.whl (163.0 MB)
Collecting nvidia-cusparse-cu12==12.3.1.170 (from torch>=2.0.0->accelerate)
  Downloading nvidia_cusparse_cu12-12.3.1.170-py3-none-manylinux2014_x86_64.whl (195.7 MB)
Requirement already satisfied: nvidia-cusparselt-cu12==0.6.2 in /usr/local/lib/python3.11/dist-packages (from torch>=2.0.0->accelerate)
Requirement already satisfied: nvidia-nccl-cu12==2.21.5 in /usr/local/lib/python3.11/dist-packages (from torch>=2.0.0->accelerate)
Requirement already satisfied: nvidia-nvtx-cu12==12.4.127 in /usr/local/lib/python3.11/dist-packages (from torch>=2.0.0->accelerate)
Collecting nvidia-nvjitlink-cu12==12.4.127 (from torch>=2.0.0->accelerate)
  Downloading nvidia_nvjitlink_cu12-12.4.127-py3-none-manylinux2014_x86_64.whl (14.0 MB)
Requirement already satisfied: triton==3.2.0 in /usr/local/lib/python3.11/dist-packages (from torch>=2.0.0->accelerate)
Requirement already satisfied: sympy==1.13.1 in /usr/local/lib/python3.11/dist-packages (from torch>=2.0.0->accelerate)
Requirement already satisfied: mpmath<1.4,>=1.1.0 in /usr/local/lib/python3.11/dist-packages (from sympy==1.13.1->torch>=2.0.0->accelerate)
Requirement already satisfied: MarkupSafe>=2.0 in /usr/local/lib/python3.11/dist-packages (from jinja2->torch>=2.0.0->accelerate)
Requirement already satisfied: charset-normalizer<4,>=2 in /usr/local/lib/python3.11/dist-packages (from requests->torch>=2.0.0->accelerate)
```

```

Requirement already satisfied: idna<4,>=2.5 in /usr/local/lib/python3.11/di
Requirement already satisfied: urllib3<3,>=1.21.1 in /usr/local/lib/python3
Requirement already satisfied: certifi>=2017.4.17 in /usr/local/lib/python3
Downloading accelerate-1.6.0-py3-none-any.whl (354 kB)
_____ 354.7/354.7 kB 11.2 MB/s eta 0:
Downloading nvidia_cublas_cu12-12.4.5.8-py3-none-manylinux2014_x86_64.whl (
_____ 363.4/363.4 MB 3.6 MB/s eta 0:
Downloading nvidia_cuda_cupti_cu12-12.4.127-py3-none-manylinux2014_x86_64.v
_____ 13.8/13.8 MB 60.2 MB/s eta 0:0
Downloading nvidia_cuda_nvrtc_cu12-12.4.127-py3-none-manylinux2014_x86_64.v
_____ 24.6/24.6 MB 41.0 MB/s eta 0:0
Downloading nvidia_cuda_runtime_cu12-12.4.127-py3-none-manylinux2014_x86_64

```

!pip install datasets



Collecting datasets

```

Downloading datasets-3.5.0-py3-none-any.whl.metadata (19 kB)
Requirement already satisfied: filelock in /usr/local/lib/python3.11/dist-pa
Requirement already satisfied: numpy>=1.17 in /usr/local/lib/python3.11/dist
Requirement already satisfied: pyarrow>=15.0.0 in /usr/local/lib/python3.11/
Collecting dill<0.3.9,>=0.3.0 (from datasets)
Downloading dill-0.3.8-py3-none-any.whl.metadata (10 kB)
Requirement already satisfied: pandas in /usr/local/lib/python3.11/dist-pack
Requirement already satisfied: requests>=2.32.2 in /usr/local/lib/python3.11
Requirement already satisfied: tqdm>=4.66.3 in /usr/local/lib/python3.11/dis
Collecting xxhash (from datasets)
Downloading xxhash-3.5.0-cp311-cp311-manylinux_2_17_x86_64.manylinux2014_x
Collecting multiprocess<0.70.17 (from datasets)
Downloading multiprocess-0.70.16-py311-none-any.whl.metadata (7.2 kB)
Collecting fsspec<=2024.12.0,>=2023.1.0 (from fsspec[http]<=2024.12.0,>=2023
Downloading fsspec-2024.12.0-py3-none-any.whl.metadata (11 kB)
Requirement already satisfied: aiohttp in /usr/local/lib/python3.11/dist-pac
Requirement already satisfied: huggingface-hub>=0.24.0 in /usr/local/lib/pyt
Requirement already satisfied: packaging in /usr/local/lib/python3.11/dist-p
Requirement already satisfied: pyyaml>=5.1 in /usr/local/lib/python3.11/dist
Requirement already satisfied: aiohappyeyeballs>=2.3.0 in /usr/local/lib/pyt
Requirement already satisfied: aiosignal>=1.1.2 in /usr/local/lib/python3.11
Requirement already satisfied: attrs>=17.3.0 in /usr/local/lib/python3.11/di
Requirement already satisfied: frozenlist>=1.1.1 in /usr/local/lib/python3.1
Requirement already satisfied: multidict<7.0,>=4.5 in /usr/local/lib/python3
Requirement already satisfied: propcache>=0.2.0 in /usr/local/lib/python3.11
Requirement already satisfied: yarll<2.0,>=1.17.0 in /usr/local/lib/python3.1
Requirement already satisfied: typing-extensions>=3.7.4.3 in /usr/local/lib/
Requirement already satisfied: charset-normalizer<4,>=2 in /usr/local/lib/py
Requirement already satisfied: idna<4,>=2.5 in /usr/local/lib/python3.11/dis
Requirement already satisfied: urllib3<3,>=1.21.1 in /usr/local/lib/python3.
Requirement already satisfied: certifi>=2017.4.17 in /usr/local/lib/python3.
Requirement already satisfied: python-dateutil>=2.8.2 in /usr/local/lib/pyth
Requirement already satisfied: pytz>=2020.1 in /usr/local/lib/python3.11/dis
Requirement already satisfied: tzdata>=2022.7 in /usr/local/lib/python3.11/d
Requirement already satisfied: six>=1.5 in /usr/local/lib/python3.11/dist-pa
Downloading datasets-3.5.0-py3-none-any.whl (491 kB)
_____ 491.2/491.2 kB 25.1 MB/s eta 0:0

```

```

Downloading dill-0.3.8-py3-none-any.whl (116 kB)
 116.3/116.3 kB 9.3 MB/s eta 0:00
Downloading fsspec-2024.12.0-py3-none-any.whl (183 kB)
 183.9/183.9 kB 14.4 MB/s eta 0:0
Downloading multiprocess-0.70.16-py311-none-any.whl (143 kB)
 143.5/143.5 kB 11.1 MB/s eta 0:0
Downloading xxhash-3.5.0-cp311-cp311-manylinux_2_17_x86_64.manylinux2014_x86_64.whl (194.8 kB)
 194.8/194.8 kB 14.1 MB/s eta 0:0
Installing collected packages: xxhash, fsspec, dill, multiprocess, datasets
Attempting uninstall: fsspec
  Found existing installation: fsspec 2025.3.2
  Uninstalling fsspec-2025.3.2:
    Successfully uninstalled fsspec-2025.3.2
ERROR: pip's dependency resolver does not currently take into account all the
dependencies of a package and you need to specify the target version: fsspec
gcsfs 2025.3.2 requires fsspec==2025.3.2, but you have fsspec 2024.12.0 which
Successfully installed datasets-3.5.0 dill-0.3.8 fsspec-2024.12.0 multiproc

```

!pip install transformers

```

⇒ Requirement already satisfied: transformers in /usr/local/lib/python3.11/dist-packages
Requirement already satisfied: filelock in /usr/local/lib/python3.11/dist-packages
Requirement already satisfied: huggingface-hub<1.0, >=0.26.0 in /usr/local/lib/python3.11/dist-packages
Requirement already satisfied: numpy>=1.17 in /usr/local/lib/python3.11/dist-packages
Requirement already satisfied: packaging>=20.0 in /usr/local/lib/python3.11/dist-packages
Requirement already satisfied: pyyaml>=5.1 in /usr/local/lib/python3.11/dist-packages
Requirement already satisfied: regex!=2019.12.17 in /usr/local/lib/python3.11/dist-packages
Requirement already satisfied: requests in /usr/local/lib/python3.11/dist-packages
Requirement already satisfied: tokenizers<0.22, >=0.21 in /usr/local/lib/python3.11/dist-packages
Requirement already satisfied: safetensors>=0.4.3 in /usr/local/lib/python3.11/dist-packages
Requirement already satisfied: tqdm>=4.27 in /usr/local/lib/python3.11/dist-packages
Requirement already satisfied: fsspec>=2023.5.0 in /usr/local/lib/python3.11/dist-packages
Requirement already satisfied: typing-extensions>=3.7.4.3 in /usr/local/lib/python3.11/dist-packages
Requirement already satisfied: charset-normalizer<4, >=2 in /usr/local/lib/python3.11/dist-packages
Requirement already satisfied: idna<4, >=2.5 in /usr/local/lib/python3.11/dist-packages
Requirement already satisfied: urllib3<3, >=1.21.1 in /usr/local/lib/python3.11/dist-packages
Requirement already satisfied: certifi>=2017.4.17 in /usr/local/lib/python3.11/dist-packages

```

!pip install evaluate

```

⇒ Collecting evaluate
  Downloading evaluate-0.4.3-py3-none-any.whl.metadata (9.2 kB)
Requirement already satisfied: datasets>=2.0.0 in /usr/local/lib/python3.11/dist-packages
Requirement already satisfied: numpy>=1.17 in /usr/local/lib/python3.11/dist-packages
Requirement already satisfied: dill in /usr/local/lib/python3.11/dist-packages
Requirement already satisfied: pandas in /usr/local/lib/python3.11/dist-packages
Requirement already satisfied: requests>=2.19.0 in /usr/local/lib/python3.11/dist-packages
Requirement already satisfied: tqdm>=4.62.1 in /usr/local/lib/python3.11/dist-packages
Requirement already satisfied: xxhash in /usr/local/lib/python3.11/dist-packages
Requirement already satisfied: multiprocess in /usr/local/lib/python3.11/dist-packages
Requirement already satisfied: fsspec>=2021.05.0 in /usr/local/lib/python3.11/dist-packages
Requirement already satisfied: huggingface-hub>=0.7.0 in /usr/local/lib/python3.11/dist-packages

```

Requirement already satisfied: packaging in /usr/local/lib/python3.11/dist-p
Requirement already satisfied: filelock in /usr/local/lib/python3.11/dist-pa
Requirement already satisfied: pyarrow>=15.0.0 in /usr/local/lib/python3.11/
Requirement already satisfied: aiohttp in /usr/local/lib/python3.11/dist-pac
Requirement already satisfied: pyyaml>=5.1 in /usr/local/lib/python3.11/dist
Requirement already satisfied: typing-extensions>=3.7.4.3 in /usr/local/lib/
Requirement already satisfied: charset-normalizer<4,>=2 in /usr/local/lib/py
Requirement already satisfied: idna<4,>=2.5 in /usr/local/lib/python3.11/dis
Requirement already satisfied: urllib3<3,>=1.21.1 in /usr/local/lib/python3.
Requirement already satisfied: certifi>=2017.4.17 in /usr/local/lib/python3.
Requirement already satisfied: python-dateutil>=2.8.2 in /usr/local/lib/pyth
Requirement already satisfied: pytz>=2020.1 in /usr/local/lib/python3.11/dis
Requirement already satisfied: tzdata>=2022.7 in /usr/local/lib/python3.11/d
Requirement already satisfied: aiohappyeyeballs>=2.3.0 in /usr/local/lib/pyt
Requirement already satisfied: aiosignal>=1.1.2 in /usr/local/lib/python3.11
Requirement already satisfied: attrs>=17.3.0 in /usr/local/lib/python3.11/di
Requirement already satisfied: frozenlist>=1.1.1 in /usr/local/lib/python3.1
Requirement already satisfied: multidict<7.0,>=4.5 in /usr/local/lib/python3
Requirement already satisfied: propcache>=0.2.0 in /usr/local/lib/python3.11
Requirement already satisfied: yarll<2.0,>=1.17.0 in /usr/local/lib/python3.1
Requirement already satisfied: six>=1.5 in /usr/local/lib/python3.11/dist-pa
Downloading evaluate-0.4.3-py3-none-any.whl (84 kB)

84.0/84.0 kB 5.6 MB/s eta 0:00:0

Installing collected packages: evaluate
Successfully installed evaluate-0.4.3

```
import transformers
from datasets import load_dataset
import evaluate
```

```
import pandas as pd
from tqdm.auto import tqdm
```

```
import torch
import torch.nn as nn
from torch.utils.data import DataLoader, Dataset
import numpy as np
```

```
import matplotlib.pyplot as plt
import seaborn as sns
```

```
sns.set(palette='summer')
```

```
device = 'cuda' if torch.cuda.is_available() else 'cpu'
device
```

```
→ 'cpu'
```

This is just an interface example. You may change it if you want.

```
class TransformerClassificationModel(nn.Module):
    def __init__(self, base_transformer_model_huggingface_name=None, base_transf
        super().__init__()
        if base_transformer_model_huggingface_name is not None:
            self.backbone = transformers.AutoModelForSequenceClassification.from_p
                base_transformer_model_huggingface_name,
                **kwargs,
                output_attentions=True
            )
        else:
            self.backbone = base_transformer_model

    def forward(self, attention_mask=None, labels=None, input_ids=None):
        outputs = self.backbone(attention_mask=attention_mask, labels=labels, in
        return outputs
```

✓ Задание 2 (1 балл)

Напишите функцию заморозки backbone у модели (если необходимо, возвращайте из функции модель)

```
def freeze_backbone_function(model: TransformerClassificationModel):
    for w in model.backbone.parameters():
        w._trainable = False
    for w in model.backbone.classifier.parameters():
        w._trainable = True
```

✓ Задание 3 (2 балла)

Напишите функцию, которая будет использована для тренировки (дообучения) трансформера (TransformerClassificationModel). Функция должна поддерживать обучение с замороженным и размороженным backbone.

```
dataset = pd.read_excel("/content/data_problems_translated.xlsx")
dataset
```



Unnamed: 0		problem_text	topic
0	0	To prove that the sum of the numbers of the ex...	number_theory
1	1	(b) Will the statement of the previous challe...	number_theory
2	2	The quadratic three-member graph with the coef...	polynoms
3	3	Can you draw on the surface of Rubik's cube a ...	combinatorics
4	4	Dima, who came from Vrunlandia, said that ther...	graphs
...
5268	5268	Can the number $n!$ end with the figures of 76.0...	number_theory
5269	5269	Find all natural m and n for which $m! + 12 = n^2$.	number_theory
5270	5270	The correct 5,000-column has a 2001 vertebrate...	dirichlet
5271	5271	Find the maximum number N for which there are ...	number_theory
5272	5272	(a) Of the 19 balls, 2 are radioactive. It is ...	combinatorics

5273 rows × 3 columns

```
dataset = dataset.to_csv("/content/data_problems_translated.csv", index=False)

dataset = pd.read_csv("/content/data_problems_translated.csv")
dataset
```




	Unnamed: 0	problem_text	topic
0	0	To prove that the sum of the numbers of the ex...	number_theory
1	1	(b) Will the statement of the previous challe...	number_theory
2	2	The quadratic three-member graph with the coef...	polynoms
3	3	Can you draw on the surface of Rubik's cube a ...	combinatorics
4	4	Dima, who came from Vrunlandia, said that ther...	graphs
...
5268	5268	Can the number $n!$ end with the figures of 76.0...	number_theory
5269	5269	Find all natural m and n for which $m! + 12 = n^2$.	number_theory
5270	5270	The correct 5,000-column has a 2001 vertebrate...	dirichlet
5271	5271	Find the maximum number N for which there are ...	number_theory
5272	5272	(a) Of the 19 balls, 2 are radioactive. It is ...	combinatorics

5273 rows × 3 columns

```
dataset = load_dataset("csv", data_files="/content/data_problems_translated.csv")
```



Generating train split: 5273/0 [00:00<00:00, 62665.59 examples/s]

```
dataset = dataset.remove_columns('Unnamed: 0')
dataset
```



```
DatasetDict({
  train: Dataset({
    features: ['problem_text', 'topic'],
    num_rows: 5273
  })
})
```

```
label_vocab = set(dataset['train']['topic'])
label2id = {char: i for i, char in enumerate(label_vocab)}
id2label = {i: char for i, char in enumerate(label_vocab)}
```

```
label2id
```



```
{'combinatorics': 0,
 'polynoms': 1,
 'graphs': 2,
 'number_theory': 3,
```



```

'geometry': 4,
'invariant': 5,
'dirichlet': 6}

```

```

dataset = dataset['train'].train_test_split(test_size=0.1)
dataset

```

```

DatasetDict({
  train: Dataset({
    features: ['problem_text', 'topic'],
    num_rows: 4745
  })
  test: Dataset({
    features: ['problem_text', 'topic'],
    num_rows: 528
  })
})

```

```

tokenizer = transformers.AutoTokenizer.from_pretrained('cointegrated/rubert-tiny')
data_collator = transformers.DataCollatorWithPadding(tokenizer=tokenizer)

```

```

/usr/local/lib/python3.11/dist-packages/huggingface_hub/utils/_auth.py:94: U
The secret `HF_TOKEN` does not exist in your Colab secrets.
To authenticate with the Hugging Face Hub, create a token in your settings t
You will be able to reuse this secret in all of your notebooks.
Please note that authentication is recommended but still optional to access
  warnings.warn(
tokenizer_config.json: 100% 401/401 [00:00<00:00, 18.8kB/s]
vocab.txt: 100% 1.08M/1.08M [00:00<00:00, 9.87MB/s]
tokenizer.json: 100% 1.74M/1.74M [00:00<00:00, 23.5MB/s]
special_tokens_map.json: 100% 112/112 [00:00<00:00, 4.97kB/s]

```

```

def preprocess_function(examples):
    examples['labels'] = [label2id.get(topic, -1) for topic in examples['topic']]
    examples['problem_text'] = [text if text else "[EMPTY]" for text in examples

    tokenized_inputs = tokenizer(
        examples["problem_text"],
        truncation=True,
        padding='max_length',
        max_length=128
    )

    tokenized_inputs['labels'] = examples['labels']

```

```
return tokenized_inputs
```

```
tokenized_dataset = dataset.map(preprocess_function, batched=True)
```



Map: 100%

4745/4745 [00:01<00:00, 3499.73 examples/s]

Map: 100%

528/528 [00:00<00:00, 2828.32 examples/s]

```
accuracy = evaluate.load("accuracy")
```

```
def compute_metrics(eval_pred):
    predictions, labels = eval_pred
    predictions, _ = predictions
    predictions = np.argmax(predictions, axis=1)
    return accuracy.compute(predictions=predictions, references=labels)
```



Downloading builder script: 100%

4.20k/4.20k [00:00<00:00, 161kB/s]

```
import copy
```

```
from transformers import TrainingArguments, Trainer
```

```
def train_transformer(transformer_model, freeze_backbone=True, batch_size=8, epo
```

```
    model = TransformerClassificationModel(
        transformer_model,
        None,
        num_labels=id2label.__len__(),
        id2label=id2label,
        label2id=label2id,
    )
```

```
    if freeze_backbone:
        freeze_backbone_function(model)
```

```
    training_args = TrainingArguments(
        output_dir='./results',
        learning_rate=2e-5,
        per_device_train_batch_size=batch_size,
        per_device_eval_batch_size=batch_size,
        num_train_epochs=epochs,
        weight_decay=0.01,
        eval_strategy="epoch",
        save_strategy="epoch",
        load_best_model_at_end=True,
```

```

report_to="none",
)

trainer = Trainer(
    model=model,
    args=training_args,
    train_dataset=tokenized_dataset["train"],
    eval_dataset=tokenized_dataset["test"],
    tokenizer=tokenizer,
    data_collator=data_collator,
    compute_metrics=compute_metrics,
)

trainer.train()

return model

```

✓ Задание 4 (1 балл)

Проверьте вашу функцию из предыдущего пункта, дообучив двумя способами *cointegrated/rubert-tiny2* из huggingface.

```

rubert_tiny_transformer_model = 'cointegrated/rubert-tiny2'
rubert_tiny_full_finetuned = train_transformer(rubert_tiny_transformer_model, fr

```



Some weights of BertForSequenceClassification were not initialized from the You should probably TRAIN this model on a down-stream task to be able to use
<ipython-input-122-aad168562657>:31: FutureWarning: `tokenizer` is deprecate
trainer = Trainer(

[5340/5340 02:54, Epoch 10/10]

Epoch	Training Loss	Validation Loss	Accuracy
1	1.559100	1.345125	0.513684
2	1.290600	1.193938	0.593684
3	1.145100	1.169543	0.585263
4	1.081300	1.118636	0.600000
5	1.030500	1.117225	0.593684
6	0.975400	1.124747	0.578947
7	0.927800	1.116880	0.581053
8	0.932400	1.122981	0.583158
9	0.880400	1.117669	0.583158
10	0.884100	1.121155	0.585263

```
rubert_tiny_transformer_model = 'cointegrated/rubert-tiny2'
```

```
rubert_tiny_finetuned_with_freezed_backbone = train_transformer(rubert_tiny_tran
```

➡ Some weights of BertForSequenceClassification were not initialized from the You should probably TRAIN this model on a down-stream task to be able to use
<ipython-input-134-aad168562657>:31: FutureWarning: `tokenizer` is deprecate
trainer = Trainer(

✓ Задание 5 (1 балл) [5340/5340 02:52, Epoch 10/10]

Epoch Training Loss Validation Loss Accuracy

Обучите *ths17/MathBERT* (с замороженным backbone и без заморозки),
проанализируйте результаты. Сравните скоры с первым заданием. Получилось
лучше или нет? Почему?

1	1.562200	1.358320	0.507368
2	1.304700	1.208773	0.557895
3	1.168500	1.186440	0.578947

YOUR CODE IS HERE (probably, similar on the previous step)

5	1.063300	1.127482	0.593684
---	----------	----------	----------

✓ Задание 6 (1 балл)

6	1.008400	1.141583	0.595789
7	0.948600	1.126490	0.587368