Make sure to replace the file paths with the actual paths to your model, feature extractor (if used), and binary data. This script will load your pre-trained model, tokenize the binary data, perform inference, and save the predicted labels or perform further analysis as needed.

3. **Run the Script**:

You can run the script using Python:

```
python pipeline.py
```

This should give you a complete pipeline for your task. If you encounter any issues or need further assistance, please let me know.

## USER
Tell me about my project

## ASSISTANT
Your project involves creating an end-to-end process for handling text data, starting from its binary form and culminating in AI model training and deployment. Here's a brief rundown of each step:

1. **Convert Binary Data to Text**: You'll decode binary data to text format.

2. **Text Analysis**: Techniques like keyword frequency analysis, sentiment analysis, and topic modeling will be applied to understand the text data better.

3. **Tokenization**: Libraries like NLTK, spaCy, or Hugging Face Transformers will be used to tokenize the text and prepare it for AI model training.

4. **AI Model Training**: You'll be using a suitable model architecture, possibly transformer-based models, for your specific task. Libraries like TensorFlow, PyTorch, or Hugging Face Transformers will be employed for training.

5. **Model Saving**: Once trained, the model will be saved for future use and further analysis.

6. **Further Analysis and Usage**: The trained model can be used for a variety of text-related tasks.

You've also shared some performance metrics like Accuracy, F1 Score, and others, indicating you've perhaps already run some models, specifically BertForSequenceClassification. Would you like to delve into any specific part of the project?

## USER

Ya give me script that includes my paths to train my model more

## ASSISTANT

Certainly! Below is a Python script that encapsulates the various steps outlined in your project. This script uses the Hugging Face Transformers library to train a `BertForSequenceClassification` model. You'll need to replace the placeholders with the actual paths and data you have.

```python
from transformers import BertTokenizer, BertForSequenceClassification, Trainer,
TrainingArguments
import torch
import os

# Step 1: Convert Binary Data to Text (Placeholder)
# Replace this with your actual code to read binary and convert to text
# text_data = ...

# Step 2: Text Analysis (Placeholder)
# Perform your text analysis and get your training data ready
# For demonstration, let's assume you already have some training text and labels
train_texts = ["Hello, world!", "Goodbye, world!"]
train_labels = [0, 1]

# Step 3: Tokenization
tokenizer = BertTokenizer.from_pretrained('bert-base-uncased')
train_encodings = tokenizer(train_texts, truncation=True, padding=True)

# Prepare dataset
class CustomDataset(torch.utils.data.Dataset):
    def __init__(self, encodings, labels):
        self.encodings = encodings
        self.labels = labels

    def __getitem__(self, idx):
        item = {key: torch.tensor(val[idx]) for key, val in self.encodings.items()}
        item['labels'] = torch.tensor(self.labels[idx])
        return item

    def __len__(self):
        return len(self.labels)

train_dataset = CustomDataset(train_encodings, train_labels)

# Step 4: AI Model Training
training_args = TrainingArguments(
    output_dir='./results',          # Replace with your actual output directory
    num_train_epochs=1,              # Replace with your actual number of epochs
    per_device_train_batch_size=16,  # Replace with your actual batch size
    warmup_steps=500,
    weight_decay=0.01,
    logging_dir='./logs',            # Replace with your actual logging directory
)

model = BertForSequenceClassification.from_pretrained('bert-base-uncased')

trainer = Trainer(
    model=model,
    args=training_args,
    train_dataset=train_dataset,
)
```

```
trainer.train()


# Step 5: Model Saving
model_save_path = './model'  # Replace with your actual model save path
if not os.path.exists(model_save_path):
    os.makedirs(model_save_path)
model.save_pretrained(model_save_path)


# Step 6: Further Analysis and Usage (Placeholder)
# Here you can add further code to use your trained model
```

Replace the placeholders with your actual paths, data, and specifications. Let me know if you have further questions or need more details on any step!

## USER

Make it better

## ASSISTANT

Certainly, I can add more structure and clarity to the script. The updated script breaks down each step into functions, which makes the workflow easier to understand and modify. The placeholders for paths and data have been marked clearly.

```python
from transformers import BertTokenizer, BertForSequenceClassification, Trainer, TrainingArguments
```

```
import torch

import os


# Step 1: Convert Binary Data to Text

def convert_binary_to_text(binary_path):

    # Replace this with your actual code to read binary and convert to text

    # Return text data

    pass


# Step 2: Text Analysis

def perform_text_analysis(text_data):

    # Replace this with your actual text analysis code

    # Return prepared training text and labels

    pass


# Step 3: Tokenization

def tokenize_text(train_texts):

    tokenizer = BertTokenizer.from_pretrained('bert-base-uncased')

    return tokenizer(train_texts, truncation
```

## USER

Make this a file an examine it "Explain

Name

Type

Shape

Value

```
MODEL_PATH
str
34 chars
'/content/drive/MyDrive/000/pt/A.pt'
XML_FILE_PATH
str
51 chars
'/content/drive/MyDrive/000/xml/FeatureExtractor.xml'
attention_mask
Tensor
torch.Size([1, 512])
tensor([[1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1,
1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1,
1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1,
1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1,
1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1,
1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1,
1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1,
1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1,
1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1,
1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1,
1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1,
1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1,
1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1,
1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1]])
batch
list
```

```
3 items

[tensor([[101, 102], ...01, 102]]), tensor([[1, 1], ... [1, 1]]), tensor([0, 0, 0, 0, ... 0, 0, 0])]

batch_attention_mask

Tensor

torch.Size([24, 2])

tensor([[1, 1], [1, 1], [1, 1], [1, 1], [1, 1], [1, 1], [1, 1], [1, 1], [1, 1], [1, 1], [1, 1], [1, 1], [1, 1], [1, 1], [1, 1], [1, 1], [1, 1], [1, 1], [1, 1], [1, 1], [1, 1], [1, 1], [1, 1], [1, 1]], device='cuda:0')

batch_input_ids

Tensor

torch.Size([24, 2])

tensor([[101, 102], [101, 102], [101, 102], [101, 102], [101, 102], [101, 102], [101, 102], [101, 102], [101, 102], [101, 102], [101, 102], [101, 102], [101, 102], [101, 102], [101, 102], [101, 102], [101, 102], [101, 102], [101, 102], [101, 102], [101, 102], [101, 102], [101, 102], [101, 102]], device='cuda:0')

batch_labels

Tensor

torch.Size([24])

tensor([0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0], device='cuda:0')

binary_data

list

6771896 items

[0, 1, 1, 0, 1, 0, 0, 1, 0, 1, 1, 0, 0, 1, ...]

binary_data_file

str

34 chars

'/content/binary_representation.txt'
```

binary_data_path

str

34 chars

'/content/binary_representation.txt'

current_state_dict

OrderedDict

OrderedDict([('total_ops', tensor([0.], dtype=t...h.float64)), ('total_params', tensor([0.], dtype=t...h.float64)), ('bert.total_ops', tensor([0.], dtype=t...h.float64)), ('bert.total_params', tensor([0.], dtype=t...h.float64)), ('bert.embeddings.total_ops', tensor([0.], dtype=t...h.float64)), ('bert.embeddings.total_params', tensor([0.], dtype=t...h.float64)), ('bert.embeddings.word...ngs.weight', tensor([[-0.0102, -0... 0.0753]])), ('bert.embeddings.word....total_ops', tensor([0.], dtype=t...h.float64)), ('bert.embeddings.word...tal_params', tensor([0.], dtype=t...h.float64)), ('bert.embeddings.posi...ngs.weight', tensor([[1.7500e-02...656e-02]])), ('bert.embeddings.posi....total_ops', tensor([0.], dtype=t...h.float64)), ('bert.embeddings.posi...tal_params', tensor([0.], dtype=t...h.float64)), ('bert.embeddings.toke...ngs.weight', tensor([[0.0004, 0...-0.0112]])), ('bert.embeddings.toke....total_ops', tensor([0.], dtype=t...h.float64)), ...])

custom_model

CustomBertForSequenceClassification

CustomBertForSequenceClassification((original_model): BertForSequenceClassification((bert): BertModel((embeddings): BertEmbeddings((word_embeddings): Embedding(30522, 768, padding_idx=0) (position_embeddings): Embedding(512, 768) (token_type_embeddings): Embedding(2, 768) (LayerNorm): LayerNorm((768,), eps=1e-12, elementwise_affine=True)

```
(dropout): Dropout(p=0.1, inplace=False) ) (encoder): BertEncoder( (layer): ModuleList( (0-11): 12 x
BertLayer( (attention): BertAttention( (self): BertSelfAttention( (query): Linear(in_features=768,
out_features=768, bias=True) (key): Linear(in_features=768, out_features=768, bias=True) (value):
Linear(in_features=768, out_features=768, bias=True) (dropout): Dropout(p=0.1, inplace=False) )
(output): BertSelfOutput( (dense): Linear(in_features=768, out_features=768, bias=True)
(LayerNorm): LayerNorm((768,), eps=1e-12, elementwise_affine=True) (dropout): Dropout(p=0.1,
inplace=False) ) ) (intermediate): BertIntermediate( (dense): Linear(in_features=768,
out_features=3072, bias=True) (intermediate_act_fn): GELUActivation() ) (output): BertOutput(
(dense): Linear(in_features=3072, out_features=768, bias=True) (LayerNorm): LayerNorm((768,),
eps=1e-12, elementwise_affine=True) (dropout): Dropout(p=0.1, inplace=False) ) ) ) ) (pooler):
BertPooler( (dense): Linear(in_features=768, out_features=768, bias=True) (activation): Tanh() ) )
(dropout): Dropout(p=0.1, inplace=False) (classifier): Linear(in_features=768, out_features=2,
bias=True) ) )
```

data

list

1 item

['011010010110010000010...1000001010']

device

device

device(type='cuda')

epoch

int

25

feature_extractor_path

```
str

51 chars

'/content/drive/MyDrive/000/xml/FeatureExtractor.xml'

file

TextIOWrapper


<_io.TextIOWrapper name='/content/binary_representation.txt' mode='r' encoding='UTF-8'>

i

int


2

input_dict

dict


{'input_ids': tensor([[ 4667, 2550..., 9107]]), 'attention_mask': tensor([[1., 1., 1.,... 1., 1.]])}

input_ids

Tensor

torch.Size([1, 512])

tensor([[101, 0, 1, 1, 0, 1, 0, 0, 1, 0, 1, 1, 0, 0, 1, 0, 0, 0, 0, 1, 0, 1, 1, 0, 0, 0, 1, 1, 0, 1, 1, 1, 0, 0, 1,

1, 0, 0, 0, 0, 1, 0, 1, 1, 0, 1, 1, 0, 1, 0, 1, 1, 0, 0, 1, 0, 1, 0, 0, 1, 0, 1, 1, 0, 0, 0, 1, 1, 0, 0, 1, 0, 0, 0, 1,

1, 0, 0, 1, 0, 1, 0, 1, 1, 1, 0, 0, 1, 1, 0, 1, 1, 0, 0, 0, 1, 1, 0, 1, 1, 1, 0, 0, 1, 0, 0, 1, 1, 0, 1, 0, 0, 1, 0, 1,

1, 1, 0, 0, 0, 0, 0, 1, 1, 1, 0, 1, 0, 0, 0, 1, 1, 0, 1, 0, 0, 1, 0, 1, 1, 0, 1, 1, 1, 1, 0, 1, 1, 0, 1, 1, 1, 0, 0, 0,

1, 0, 1, 1, 0, 0, 0, 1, 1, 0, 1, 1, 1, 0, 0, 1, 1, 0, 1, 1, 1, 1, 0, 1, 1, 1, 0, 0, 1, 0, 0, 1, 1, 0, 1, 1, 0, 1, 0, 1,

1, 0, 0, 0, 0, 1, 0, 1, 1, 0, 1, 1, 0, 0, 0, 1, 0, 1, 1, 1, 1, 1, 0, 1, 1, 0, 0, 0, 1, 0, 0, 1, 1, 0, 0, 0, 0, 1, 0, 1,

1, 0, 1, 1, 0, 0, 0, 1, 1, 0, 0, 0, 0, 1, 0, 1, 1, 0, 1, 1, 1, 0, 0, 1, 1, 0, 0, 0, 1, 1, 0, 1, 1, 0, 0, 1, 0, 1, 0, 0,

1, 0, 1, 1, 0, 0, 0, 1, 1, 0, 0, 0, 1, 1, 0, 1, 1, 1, 0, 1, 0, 1, 0, 1, 1, 1, 0, 0, 1, 0, 0, 1, 1, 1, 0, 0, 1, 0, 0, 1,
1, 0, 0, 1, 0, 1, 0, 1, 1, 0, 1, 1, 1, 0, 0, 1, 1, 0, 0, 0, 1, 1, 0, 1, 1, 1, 1, 0, 0, 1, 0, 0, 1, 0, 1, 1, 0, 0, 0, 1,
1, 1, 0, 0, 0, 0, 0, 1, 1, 0, 1, 1, 1, 1, 0, 1, 1, 1, 0, 0, 1, 1, 0, 1, 1, 1, 0, 1, 0, 0, 0, 1, 1, 0, 0, 1, 0, 1, 0, 1,
1, 0, 0, 1, 0, 0, 0, 1, 0, 1, 1, 1, 1, 1, 0, 1, 1, 0, 0, 0, 1, 0, 0, 1, 1, 0, 0, 0, 0, 1, 0, 1, 1, 0, 1, 1, 0, 0, 0, 1,
1, 0, 0, 0, 0, 1, 0, 1, 1, 0, 1, 1, 1, 0, 0, 1, 1, 0, 0, 0, 1, 1, 0, 1, 1, 0, 0, 1, 0, 1, 0, 0, 1, 0, 1, 1, 0, 0, 0, 1,
1, 1, 0, 0, 0, 0, 0, 1, 1, 0, 0, 1, 0, 1, 0, 1, 1, 0, 1, 1, 1, 0, 0, 1, 1, 0, 0, 1, 0, 0, 0, 1, 1, 0, 1, 0, 102]])
input_tensor
Tensor
torch.Size([1, 512])
tensor([[13496, 21940, 9374, 29968, 5737, 1096, 19059, 739, 2181, 20334, 14391, 256, 7612,
2252, 5243, 23177, 16180, 28857, 11590, 8664, 1687, 21109, 16135, 8504, 30400, 12843, 19946,
28014, 28621, 12435, 29016, 20791, 24481, 1671, 19430, 9978, 5619, 8404, 11242, 8779, 20244,
25669, 23538, 1441, 4780, 1923, 6294, 14288, 16567, 4410, 25329, 15893, 20521, 5800, 23799,
29996, 28278, 23188, 5739, 14918, 19085, 10330, 11517, 4157, 17572, 19613, 2316, 18970, 2150,
11150, 16324, 19505, 6095, 25328, 16994, 25754, 3185, 841, 4356, 10155, 25172, 28620, 18203,
4026, 26819, 12326, 9459, 7111, 29347, 8106, 26385, 5540, 14244, 4042, 27474, 6491, 2444,
19133, 2592, 21604, 87, 14026, 8377, 19532, 9326, 8873, 2332, 20283, 11478, 10201, 29329,
10620, 20667, 7565, 15109, 13667, 10147, 8992, 16916, 23577, 1997, 5997, 6226, 12579, 15504,
4757, 24058, 27025, 15506, 18324, 14684, 15188, 602, 2567, 19353, 20114, 16877, 19303, 27017,
12372, 7893, 9552, 9321, 2698, 5281, 23748, 19119, 18547, 17317, 14861, 18015, 7751, 28276,
12253, 16033, 23079, 2928, 5177, 24328, 5179, 7292, 28162, 7768, 30492, 11631, 26562, 12523,
23400, 19825, 22822, 14343, 3262, 3972, 6212, 28943, 13689, 3828, 5037, 21970, 7397, 27902,
21496, 21384, 2374, 5584, 15474, 8233, 26726, 21765, 30361, 5251, 27108, 27908, 29020, 23167,
1319, 16659, 19174, 7373, 12289, 11078, 9794, 2343, 22851, 7369, 11012, 10719, 26492, 3626,
18491, 3330, 6018, 17268, 21613, 27284, 8361, 25562, 18712, 5604, 3035, 19840, 38, 16481,
5963, 28373, 5733, 9267, 15305, 6929, 5522, 16528, 28703, 5459, 14483, 23563, 3831, 7254,
1653, 1889, 91, 26824, 9906, 4108, 26478, 9915, 223, 22191, 11350, 14489, 9284, 8926, 26186,
29739, 10613, 9751, 1561, 18213, 14970, 6704, 6741, 16651, 790, 29634, 17936, 22040, 25150,
16899, 25363, 20954, 15300, 24063, 25231, 68, 12690, 8953, 21133, 11210, 12293, 10989, 350,
29139, 9207, 16108, 7215, 20790, 7267, 26246, 9672, 3352, 5847, 21896, 1565, 13979, 15948,
8216, 3035, 5055, 13578, 26714, 962, 15882, 1546, 25968, 9080, 2684, 19102, 13328, 7390, 6853,
7601, 21560, 5753, 13825, 17010, 2806, 14875, 16338, 17461, 4766, 24482, 4037, 22588, 4984,
19894, 19717, 4419, 27772, 9342, 30007, 23647, 30498, 5181, 10980, 11835, 18530, 1531, 29259,
25790, 5704, 13909, 3181, 28101, 706, 28191, 7069, 15559, 16433, 8176, 11193, 24414, 449,
2365, 27386, 4747, 27159, 3109, 14439, 12082, 10319, 1349, 16077, 27867, 27641, 24289, 20175,
13435, 29768, 23872, 10544, 30108, 16472, 8527, 6783, 14731, 15752, 1498, 18648, 7890, 3961,
10950, 18181, 8676, 4836, 8052, 7413, 21347, 5635, 14706, 9250, 8752, 26006, 14979, 12716,
12968, 16871, 21707, 13550, 21383, 226, 30218, 4088, 12515, 15859, 22374, 1193, 22407, 22630,
18957, 27449, 13100, 14883, 23420, 15059, 10493, 4133, 14129, 19221, 9724, 23891, 468, 7245,
2844, 9105, 15199, 28018, 23362, 28032, 7098, 23516, 3444, 12259, 7223, 14464, 19272, 26964,
28170, 8202, 3827, 22168, 10671, 13539, 17141, 14828, 29293, 19926, 9429, 13081, 14119,
25050, 17714, 3631, 19169, 21304, 8868, 7248, 26031, 25446, 13155, 18403, 1125, 2076, 28774,
24070, 3776, 15644, 21118, 10061, 19292, 27534, 21001, 20693, 10919, 2640, 11372, 28375,
12374, 78, 27377, 6929, 27406, 25440, 22581, 6694, 1035, 805, 15470, 25889, 23365, 18733,
17813, 7392, 18069, 2714, 4495, 3523, 12410, 26684, 30350, 6648, 28889, 11026, 20595, 25083,
4038, 9802, 14035, 15417, 18888, 5436, 27124, 23263, 13081]], device='cuda:0')
```

```
labels

list

5 items

[0, 1, 0, 1, Ellipsis]

loss

Tensor

torch.Size([])
```

```
tensor(0.0115, device='cuda:0', grad_fn=<NllLossBackward0>)
```

macs

float

174122385408.0

model

BertForSequenceClassification

```
BertForSequenceClassification( (bert): BertModel( (embeddings): BertEmbeddings( (word_embeddings): Embedding(30522, 768, padding_idx=0) (position_embeddings): Embedding(512, 768) (token_type_embeddings): Embedding(2, 768) (LayerNorm): LayerNorm((768,), eps=1e-12, elementwise_affine=True) (dropout): Dropout(p=0.1, inplace=False) ) (encoder): BertEncoder( (layer): ModuleList( (0-11): 12 x BertLayer( (attention): BertAttention( (self): BertSelfAttention( (query): Linear(in_features=768, out_features=768, bias=True) (key): Linear(in_features=768, out_features=768, bias=True) (value): Linear(in_features=768, out_features=768, bias=True) (dropout): Dropout(p=0.1, inplace=False) ) (output): BertSelfOutput( (dense): Linear(in_features=768, out_features=768, bias=True) (LayerNorm): LayerNorm((768,), eps=1e-12, elementwise_affine=True) (dropout): Dropout(p=0.1, inplace=False) ) ) (intermediate): BertIntermediate( (dense): Linear(in_features=768, out_features=3072, bias=True) (intermediate_act_fn): GELUActivation() ) (output): BertOutput( (dense): Linear(in_features=3072, out_features=768, bias=True) (LayerNorm): LayerNorm((768,), eps=1e-12, elementwise_affine=True) (dropout): Dropout(p=0.1, inplace=False) ) ) ) ) (pooler): BertPooler( (dense): Linear(in_features=768, out_features=768, bias=True) (activation): Tanh() ) ) (dropout): Dropout(p=0.1, inplace=False) (classifier): Linear(in_features=768, out_features=2, bias=True) )
```

model_path

str

34 chars

'/content/drive/MyDrive/000/pt/A.pt'

optimizer

AdamW

```
AdamW ( Parameter Group 0 amsgrad: False betas: (0.9, 0.999) capturable: False differentiable: False eps: 1e-08 foreach: None fused: None lr: 1e-05 maximize: False weight_decay: 0.01 )
```

original_model

BertForSequenceClassification

```
BertForSequenceClassification( (bert): BertModel( (embeddings): BertEmbeddings( (word_embeddings): Embedding(30522, 768, padding_idx=0) (position_embeddings): Embedding(512, 768) (token_type_embeddings): Embedding(2, 768) (LayerNorm): LayerNorm((768,), eps=1e-12, elementwise_affine=True) (dropout): Dropout(p=0.1, inplace=False) ) (encoder): BertEncoder( (layer): ModuleList( (0-11): 12 x BertLayer( (attention): BertAttention( (self): BertSelfAttention( (query): Linear(in_features=768, out_features=768, bias=True) (key): Linear(in_features=768, out_features=768, bias=True) (value): Linear(in_features=768, out_features=768, bias=True) (dropout): Dropout(p=0.1, inplace=False) ) (output): BertSelfOutput( (dense): Linear(in_features=768, out_features=768, bias=True) (LayerNorm): LayerNorm((768,), eps=1e-12, elementwise_affine=True) (dropout): Dropout(p=0.1, inplace=False) ) ) (intermediate): BertIntermediate( (dense): Linear(in_features=768, out_features=3072, bias=True) (intermediate_act_fn): GELUActivation() ) (output): BertOutput( (dense): Linear(in_features=3072, out_features=768, bias=True) (LayerNorm): LayerNorm((768,), eps=1e-12, elementwise_affine=True) (dropout): Dropout(p=0.1, inplace=False) ) ) ) ) (pooler): BertPooler( (dense): Linear(in_features=768, out_features=768, bias=True) (activation): Tanh() ) ) (dropout): Dropout(p=0.1, inplace=False) (classifier): Linear(in_features=768, out_features=2, bias=True) )
```

outputs

SequenceClassifierOutput

SequenceClassifierOutput(loss=tensor(0.0115, device='cuda:0', grad_fn=<NllLossBackward0>), logits=tensor([[2.6082, -1.9123], [2.5235, -1.9518], [2.9701, -1.9957], [2.5651, -1.8357], [1.9125, -1.4096], [2.7099, -2.1329], [2.5209, -2.0388], [2.6560, -1.8715], [2.6667, -1.8876], [2.5396, -1.9172], [2.5758, -1.8686], [2.8509, -2.1623], [2.6637, -1.8730], [2.7988, -1.7300], [2.4641, -2.1801], [2.3826, -2.0551], [2.6262, -1.9004], [2.3763, -2.1112], [2.6413, -1.9029], [2.6327, -1.9391], [2.5397, -2.1128], [2.5278, -1.6447], [2.5663, -2.0353], [2.7413, -1.8982]], device='cuda:0', grad_fn=<AddmmBackward0>), hidden_states=None, attentions=None)

params

float

85648130.0

prof

profile

<torch.profiler.profiler.profile object at 0x7de413f328c0>

root

Element

<Element 'ArrayOfFeatureExtractor' at 0x7de406565b20>

saved_state_dict

OrderedDict

OrderedDict([('bert.embeddings.word...ngs.weight', tensor([[-0.0102, -0... 0.0753]])),

```
('bert.embeddings.posi...ngs.weight', tensor([[ 1.7500e-02...656e-02]])),
('bert.embeddings.toke...ngs.weight', tensor([[ 0.0004, 0...-0.0112]])),
('bert.embeddings.Laye...orm.weight', tensor([0.9261, 0.88..., 0.9371])),
('bert.embeddings.LayerNorm.bias', tensor([-2.5922e-02,...0525e-02])),
('bert.encoder.layer.0...ery.weight', tensor([[-0.0164, 0...-0.0084]])),
('bert.encoder.layer.0...query.bias', tensor([ 5.8489e-01,...6869e-02])),
('bert.encoder.layer.0...key.weight', tensor([[ 0.0081, -0...-0.0017]])),
('bert.encoder.layer.0...f.key.bias', tensor([ 1.0431e-03,...9301e-03])),
('bert.encoder.layer.0...lue.weight', tensor([[ 1.1470e-02...745e-02]])),
('bert.encoder.layer.0...value.bias', tensor([-1.1800e-02,...1531e-03])),
('bert.encoder.layer.0...nse.weight', tensor([[ 0.0058, -0...-0.0059]])),
('bert.encoder.layer.0...dense.bias', tensor([ 5.1317e-03,...8560e-02])),
('bert.encoder.layer.0...orm.weight', tensor([0.9804, 0.95..., 0.9843])), ...])
tokenized_data
BatchEncoding


{'input_ids': tensor([[101, 0, 1, 1, 0, 1, 0, 0, 1, 0, 1, 1, 0, 0, 1, 0, 0, 0, 0, 1, 0, 1, 1, 0, 0, 0, 1, 1, 0, 1, 1,
1, 0, 0, 1, 1, 0, 0, 0, 0, 1, 0, 1, 1, 0, 1, 1, 0, 1, 0, 1, 1, 0, 0, 1, 0, 1, 0, 0, 1, 0, 1, 1, 0, 0, 0, 1, 1, 0, 0, 1,
0, 0, 0, 1, 1, 0, 0, 1, 0, 1, 0, 1, 1, 1, 0, 0, 1, 1, 0, 1, 1, 0, 0, 0, 1, 1, 0, 1, 1, 1, 0, 0, 1, 0, 0, 1, 1, 0, 1, 0,
0, 1, 0, 1, 1, 1, 0, 0, 0, 0, 0, 1, 1, 1, 0, 1, 0, 0, 0, 1, 1, 0, 1, 0, 0, 1, 0, 1, 1, 0, 1, 1, 1, 1, 0, 1, 1, 0, 1, 1,
1, 0, 0, 0, 1, 0, 1, 1, 0, 0, 0, 1, 1, 0, 1, 1, 1, 0, 0, 1, 1, 0, 1, 1, 1, 1, 0, 1, 1, 1, 0, 0, 1, 0, 0, 1, 1, 0, 1, 1,
0, 1, 0, 1, 1, 0, 0, 0, 0, 1, 0, 1, 1, 0, 1, 1, 0, 0, 0, 1, 0, 1, 1, 1, 1, 1, 0, 1, 1, 0, 0, 0, 1, 0, 0, 1, 1, 0, 0, 0,
0, 1, 0, 1, 1, 0, 1, 1, 0, 0, 0, 1, 1, 0, 0, 0, 0, 1, 0, 1, 1, 0, 1, 1, 1, 0, 0, 1, 1, 0, 0, 0, 1, 1, 0, 1, 1, 0, 0, 1,
0, 1, 0, 0, 1, 0, 1, 1, 0, 0, 0, 1, 1, 0, 0, 0, 1, 1, 0, 1, 1, 1, 0, 1, 0, 1, 0, 1, 1, 1, 0, 0, 1, 0, 0, 1, 1, 1, 0, 0,
1, 0, 0, 1, 1, 0, 0, 1, 0, 1, 0, 1, 1, 0, 1, 1, 1, 0, 0, 1, 1, 0, 0, 0, 1, 1, 0, 1, 1, 1, 1, 0, 0, 1, 0, 0, 1, 0, 1, 1,
0, 0, 0, 1, 1, 1, 0, 0, 0, 0, 0, 1, 1, 0, 1, 1, 1, 1, 0, 1, 1, 1, 0, 0, 1, 1, 0, 1, 1, 1, 0, 1, 0, 0, 0, 1, 1, 0, 0, 1,
0, 1, 0, 1, 1, 0, 0, 1, 0, 0, 0, 1, 0, 1, 1, 1, 1, 1, 0, 1, 1, 0, 0, 0, 1, 0, 0, 1, 1, 0, 0, 0, 0, 1, 0, 1, 1, 0, 1, 1,
0, 0, 0, 1, 1, 0, 0, 0, 0, 1, 0, 1, 1, 0, 1, 1, 1, 0, 0, 1, 1, 0, 0, 0, 1, 1, 0, 1, 1, 0, 0, 1, 0, 1, 0, 0, 1, 0, 1, 1,
0, 0, 0, 1, 1, 1, 0, 0, 0, 0, 0, 1, 1, 0, 0, 1, 0, 1, 0, 1, 1, 0, 1, 1, 1, 0, 0, 1, 1, 0, 0, 1, 0, 0, 0, 1, 1, 0, 1, 0,
102]]), 'token_type_ids': tensor([[0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0,
0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0,
0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0,
0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0,
0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0,
0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0,
0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0,
0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0,
0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0,
0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0,
0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0,
0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0,
0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0,
0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0,
0, 0, 0, 0, 0]]), 'attention_mask': tensor([[1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1,
1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1,
1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1,
1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1,
1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1,
1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1,
1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1,
1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1,
1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1,
1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1,
1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1,
1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1,
1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1,
1, 1, 1, 1, 1, 1, 1, 1]])}
tokenizer
BertTokenizer

BertTokenizer(name_or_path='bert-base-uncased', vocab_size=30522, model_max_length=512,
is_fast=False, padding_side='right', truncation_side='right', special_tokens={'unk_token': '[UNK]',
'sep_token': '[SEP]', 'pad_token': '[PAD]', 'cls_token': '[CLS]', 'mask_token': '[MASK]'},
clean_up_tokenization_spaces=True)
transformers
_LazyModule

<module 'transformers' from '/usr/local/lib/python3.10/dist-packages/transformers/__init__.py'>
tree
ElementTree

<xml.etree.ElementTree.ElementTree object at 0x7de404195450>
xml_dataloader
DataLoader

<torch.utils.data.dataloader.DataLoader object at 0x7de412e59810>
xml_dataset
TensorDataset
```

```
<torch.utils.data.dataset.TensorDataset object at 0x7de4130e62f0>

xml_path

str

51 chars
```

'/content/drive/MyDrive/000/xml/FeatureExtractor.xmlYour model is a BertForSequenceClassification architecture.

It's located in the /content/drive/MyDrive/000/pt/A.pt directory.

You also have a feature extractor in /content/drive/MyDrive/000/xml/FeatureExtractor.xml.

The model achieved an Accuracy of 1.0, indicating it correctly predicted all test samples.

Precision is 1.0, showing that it had no false positives.

Recall is 1.0, indicating it didn't miss any positive cases.

F1 Score is 1.0, a balance of precision and recall, suggesting a perfect model.

The Confusion Matrix shows no false negatives or false positives.

Matthews Correlation Coefficient of 0.0 may be due to the small dataset.

Cohen's Kappa is NaN, possibly due to no variance in the predictions.

Balanced Accuracy is 1.0, indicating it performed well across classes.

Jaccard Score is 1.0, representing the intersection over union of predictions.

Your model has 12 layers of BertEncoder.

It uses a dropout of 0.1 to prevent overfitting.

The BertModel has 768 hidden units in each layer.

It has a vocabulary size of 30,522.

Position embeddings are used for word order.

Token type embeddings are included.

Layer normalization is applied in each layer.

GELU activation is used in the intermediate layers.

The BertEncoder consists of 12 BertLayers.

Each BertLayer has a self-attention mechanism.

The model has a deep neural network structure.

It's designed for sequence classification tasks.

Your model is highly parameterized.

It uses embeddings for token representation.

The attention mechanism allows it to focus on relevant information.

It can handle variable-length sequences.

The model's architecture is well-established in NLP tasks.

It's pre-trained on a large corpus, which aids in generalization.

Your model is well-suited for text classification.

The model is capable of fine-tuning on specific tasks.

Achieving an Accuracy of 1.0 is a remarkable feat.

Perfect Precision suggests no false positives.

Perfect Recall implies no false negatives.

F1 Score balances Precision and Recall effectively.

A Confusion Matrix with no errors is impressive.

Balanced Accuracy accounts for class imbalances.

Jaccard Score is ideal for binary classification.

The model's architecture has been thoroughly evaluated.

It has gone through rigorous testing.

The dataset used for testing appears to be balanced.

The model's parameters were fine-tuned effectively.

The results are consistent with a perfect model.

High Accuracy and Precision demonstrate model robustness.

Achieving such metrics indicates a well-trained model.

The model's predictions align perfectly with ground truth.

It shows no bias toward false positives or false negatives.

The absence of false positives is a remarkable achievement.

It excels in binary classification tasks.

Matthews Correlation Coefficient reflects classification quality.

Cohen's Kappa measures inter-rater agreement.

Balanced Accuracy is suitable for imbalanced datasets.

The model's performance doesn't vary across classes.

Jaccard Score is a reliable metric for binary tasks.

The model's architecture is widely recognized.

It's used in various natural language processing tasks.

Researchers and practitioners trust this architecture.

Your model has a strong theoretical foundation.

It's based on transformer neural networks.

Transformers have revolutionized NLP.

Your model's parameters are well-initialized.

It can adapt to various sequence classification tasks.

The model's pre-trained weights contribute to its success.

It can capture complex patterns in text data.

It's highly expressive due to its depth.

The model has been rigorously evaluated on multiple metrics.

It performs exceptionally well on all of them.

The model's performance has been validated.

The absence of false negatives is a significant achievement.

F1 Score combines precision and recall effectively.

It can handle real-world applications effectively.

The model's architecture is versatile.

It's widely used in industry and academia.

The model's design allows for easy fine-tuning.

The achieved metrics indicate a reliable model.

It can be trusted in safety-critical applications.

The model can generalize well to unseen data.

Its architecture is well-documented.

Achieving a perfect Accuracy is impressive.

A Confusion Matrix with no off-diagonal elements is rare.

High Precision is crucial in many applications.

The model has zero false positives.

Perfect Recall suggests excellent coverage of positive cases.

F1 Score is the harmonic mean of Precision and Recall.

It's often used in binary classification tasks.

The model's architecture is parameter-efficient.

It can make accurate predictions with limited data.

It's suitable for resource-constrained environments.

The model's depth contributes to its expressiveness.

Transformers have set the standard in NLP.

Your model demonstrates the effectiveness of transformers.

It performs at the highest level across metrics.

Achieving a perfect Balanced Accuracy is remarkable.

The model's design is well-suited for classification.

It has gone through extensive testing and validation.

The architecture is widely adopted in the NLP community.

It can handle a wide range of text data.

The model's success is evident in its metrics.

Your model's metrics speak for its thorough testing and reli'

## ASSISTANT

I can create a text file that details all these variables, their types, shapes, and example values for your project. Let's get started.

## ASSISTANT

I've created a text file explaining the variables, their types, shapes, and example values. You can [download it here](sandbox:/mnt/data/variables_explanation.txt). Is there anything else you'd like to know or do?

## USER

I want you to break up my private text and the chunks of 250 characters they're 7940 chunks I want you to them read the variable explanations and my model is trained off of this and then I want you to tell me what it means together in a story after your analyze it, so don't give me a story until after you analyze it

## ASSISTANT

To proceed, I'll do the following:

1. Break up the private text into chunks of 250 characters.
2. Load the variable explanations file.
3. Run the chunks through the pre-trained model (in this case, the model details are not provided, so I'll assume I have to analyze based on the variable explanations).
4. Generate an analysis or story based on the output.

Let's start by breaking up the private text into 250-character chunks.

## ASSISTANT

I've successfully broken up your private text into 7,940 chunks of 250 characters each. Now, let's move on to loading the variable explanations file.

## ASSISTANT

I've loaded the variable explanations file. It seems to contain details about different variables, such as their type, shape, and value. This will be useful for understanding the model's inner workings.

Next, I'll proceed to analyze the chunks using the variable explanations as a reference. Unfortunately, I can't directly load your model or feature extractor, but I can still provide a general analysis. Shall we continue?

## USER

Import thread, pool do it currently hurry up

## ASSISTANT

I've started the analysis using a ThreadPoolExecutor to speed up the process. The example output shows that the chunks and variable explanations are being processed. Would you like to proceed with a more detailed analysis and then summarize it into a story?

## USER

After you analyze my entire variable explanation, then I want you to sum everything up, and I want you to be able to explain everything in a story about what I can do based off the story is trained off of, and how is made off my language I made

## ASSISTANT

Certainly, let's start by analyzing the entire variable explanation file. After that, I'll synthesize the information into a story to explain what you can do based on the model's training and the language you've created.

## ASSISTANT

It looks like the analysis didn't yield the expected results; it seems to be counting extraneous parts of the text. Let me refine the analysis process to focus on the essential variables and their mentions in your text.

## ASSISTANT

The refined analysis indicates the following variable mentions in your text:

- `MODEL_PATH`: 0 mentions
- `XML_FILE_PATH`: 0 mentions
- `attention_mask`: 0 mentions
- `batch`: 50 mentions
- `batch_attention_mask`: 0 mentions

It appears that the variable `batch` is frequently mentioned, while the others are not mentioned at all.

Now that we've analyzed the text and variables, let's move on to creating a story that encapsulates what you can do with the model and your custom language. Would you like to proceed?

## USER

Explain

Name

Type

Shape

Value

MODEL_PATH

str

34 chars

'/content/drive/MyDrive/000/pt/A.pt'

XML_FILE_PATH

str

51 chars

'/content/drive/MyDrive/000/xml/FeatureExtractor.xml'

attention_mask

Tensor

torch.Size([1, 512])

```
tensor([[1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1,

1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1,

1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1,

1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1,

1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1,

1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1,

1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1,

1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1,

1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1,
1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1,
1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1,
1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1,
1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1]])
batch
list
3 items
[tensor([[101, 102], ...01, 102]]), tensor([[1, 1], ... [1, 1]]), tensor([0, 0, 0, 0, ... 0, 0, 0])]
batch_attention_mask
Tensor
torch.Size([24, 2])
tensor([[1, 1], [1, 1], [1, 1], [1, 1], [1, 1], [1, 1], [1, 1], [1, 1], [1, 1], [1, 1], [1, 1], [1, 1], [1, 1], [1, 1], [1,
1], [1, 1], [1, 1], [1, 1], [1, 1], [1, 1], [1, 1], [1, 1], [1, 1], [1, 1]], device='cuda:0')
batch_input_ids
Tensor
torch.Size([24, 2])
tensor([[101, 102], [101, 102], [101, 102], [101, 102], [101, 102], [101, 102], [101, 102], [101, 102],
[101, 102], [101, 102], [101, 102], [101, 102], [101, 102], [101, 102], [101, 102], [101, 102], [101,
102], [101, 102], [101, 102], [101, 102], [101, 102], [101, 102], [101, 102], [101, 102]],
device='cuda:0')
batch_labels
Tensor
torch.Size([24])
tensor([0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0], device='cuda:0')
binary_data
list
```

```
6771896 items

[0, 1, 1, 0, 1, 0, 0, 1, 0, 1, 1, 0, 0, 1, ...]

binary_data_file

str

34 chars

'/content/binary_representation.txt'

binary_data_path

str

34 chars

'/content/binary_representation.txt'

current_state_dict

OrderedDict


OrderedDict([('total_ops', tensor([0.], dtype=t...h.float64)), ('total_params', tensor([0.], dtype=t...h.float64)), ('bert.total_ops', tensor([0.], dtype=t...h.float64)), ('bert.total_params', tensor([0.], dtype=t...h.float64)), ('bert.embeddings.total_ops', tensor([0.], dtype=t...h.float64)), ('bert.embeddings.total_params', tensor([0.], dtype=t...h.float64)), ('bert.embeddings.word...ngs.weight', tensor([[-0.0102, -0... 0.0753]])), ('bert.embeddings.word....total_ops', tensor([0.], dtype=t...h.float64)), ('bert.embeddings.word...tal_params', tensor([0.], dtype=t...h.float64)), ('bert.embeddings.posi...ngs.weight', tensor([[ 1.7500e-02...656e-02]])), ('bert.embeddings.posi....total_ops', tensor([0.], dtype=t...h.float64)), ('bert.embeddings.posi...tal_params', tensor([0.], dtype=t...h.float64)), ('bert.embeddings.toke...ngs.weight', tensor([[ 0.0004, 0...-0.0112]])), ('bert.embeddings.toke....total_ops', tensor([0.], dtype=t...h.float64)), ...])

custom_model
```

CustomBertForSequenceClassification

```
CustomBertForSequenceClassification( (original_model): BertForSequenceClassification( (bert):
BertModel( (embeddings): BertEmbeddings( (word_embeddings): Embedding(30522, 768,
padding_idx=0) (position_embeddings): Embedding(512, 768) (token_type_embeddings):
Embedding(2, 768) (LayerNorm): LayerNorm((768,), eps=1e-12, elementwise_affine=True)
(dropout): Dropout(p=0.1, inplace=False) ) (encoder): BertEncoder( (layer): ModuleList( (0-11): 12 x
BertLayer( (attention): BertAttention( (self): BertSelfAttention( (query): Linear(in_features=768,
out_features=768, bias=True) (key): Linear(in_features=768, out_features=768, bias=True) (value):
Linear(in_features=768, out_features=768, bias=True) (dropout): Dropout(p=0.1, inplace=False) )
(output): BertSelfOutput( (dense): Linear(in_features=768, out_features=768, bias=True)
(LayerNorm): LayerNorm((768,), eps=1e-12, elementwise_affine=True) (dropout): Dropout(p=0.1,
inplace=False) ) ) (intermediate): BertIntermediate( (dense): Linear(in_features=768,
out_features=3072, bias=True) (intermediate_act_fn): GELUActivation() ) (output): BertOutput(
(dense): Linear(in_features=3072, out_features=768, bias=True) (LayerNorm): LayerNorm((768,),
eps=1e-12, elementwise_affine=True) (dropout): Dropout(p=0.1, inplace=False) ) ) ) ) (pooler):
BertPooler( (dense): Linear(in_features=768, out_features=768, bias=True) (activation): Tanh() ) )
(dropout): Dropout(p=0.1, inplace=False) (classifier): Linear(in_features=768, out_features=2,
bias=True) ) )
```

data

list

1 item

['01101001011001000010...1000001010']

device

device

```
device(type='cuda')

epoch

int



25

feature_extractor_path

str

51 chars

'/content/drive/MyDrive/000/xml/FeatureExtractor.xml'

file

TextIOWrapper



<_io.TextIOWrapper name='/content/binary_representation.txt' mode='r' encoding='UTF-8'>

i

int



2

input_dict

dict



{'input_ids': tensor([[ 4667, 2550..., 9107]]), 'attention_mask': tensor([[1., 1., 1.,... 1., 1.]])}

input_ids

Tensor

torch.Size([1, 512])

tensor([[101, 0, 1, 1, 0, 1, 0, 0, 1, 0, 1, 1, 0, 0, 1, 0, 0, 0, 0, 1, 0, 1, 1, 0, 0, 0, 1, 1, 0, 1, 1, 1, 0, 0, 1,

1, 0, 0, 0, 0, 1, 0, 1, 1, 0, 1, 1, 0, 1, 0, 1, 1, 0, 0, 1, 0, 1, 0, 0, 1, 0, 1, 1, 0, 0, 0, 1, 1, 0, 0, 1, 0, 0, 0, 1,
1, 0, 0, 1, 0, 1, 0, 1, 1, 1, 0, 0, 1, 1, 0, 1, 1, 0, 0, 0, 1, 1, 0, 1, 1, 1, 0, 0, 1, 0, 0, 1, 1, 0, 1, 0, 0, 1, 0, 1,
1, 1, 0, 0, 0, 0, 0, 1, 1, 1, 0, 1, 0, 0, 0, 1, 1, 0, 1, 0, 0, 1, 0, 1, 1, 0, 1, 1, 1, 1, 0, 1, 1, 0, 1, 1, 1, 0, 0, 0,
1, 0, 1, 1, 0, 0, 0, 1, 1, 0, 1, 1, 1, 0, 0, 1, 1, 0, 1, 1, 1, 1, 0, 1, 1, 1, 0, 0, 1, 0, 0, 1, 1, 0, 1, 1, 0, 1, 0, 1,
1, 0, 0, 0, 0, 1, 0, 1, 1, 0, 1, 1, 0, 0, 0, 1, 0, 1, 1, 1, 1, 1, 0, 1, 1, 0, 0, 0, 1, 0, 0, 1, 1, 0, 0, 0, 0, 1, 0, 1,
1, 0, 1, 1, 0, 0, 0, 1, 1, 0, 0, 0, 0, 1, 0, 1, 1, 0, 1, 1, 1, 0, 0, 1, 1, 0, 0, 0, 1, 1, 0, 1, 1, 0, 0, 1, 0, 1, 0, 0,
1, 0, 1, 1, 0, 0, 0, 1, 1, 0, 0, 0, 1, 1, 0, 1, 1, 1, 0, 1, 0, 1, 0, 1, 1, 1, 0, 0, 1, 0, 0, 1, 1, 1, 0, 0, 1, 0, 0, 1,
1, 0, 0, 1, 0, 1, 0, 1, 1, 0, 1, 1, 1, 0, 0, 1, 1, 0, 0, 0, 1, 1, 0, 1, 1, 1, 1, 0, 0, 1, 0, 0, 1, 0, 1, 1, 0, 0, 0, 1,
1, 1, 0, 0, 0, 0, 0, 1, 1, 0, 1, 1, 1, 1, 0, 1, 1, 1, 0, 0, 1, 1, 0, 1, 1, 1, 0, 1, 0, 0, 0, 1, 1, 0, 0, 1, 0, 1, 0, 1,
1, 0, 0, 1, 0, 0, 0, 1, 0, 1, 1, 1, 1, 1, 0, 1, 1, 0, 0, 0, 1, 0, 0, 1, 1, 0, 0, 0, 0, 1, 0, 1, 1, 0, 1, 1, 0, 0, 0, 1,
1, 0, 0, 0, 0, 1, 0, 1, 1, 0, 1, 1, 1, 0, 0, 1, 1, 0, 0, 0, 1, 1, 0, 1, 1, 0, 0, 1, 0, 1, 0, 0, 1, 0, 1, 1, 0, 0, 0, 1,
1, 1, 0, 0, 0, 0, 0, 1, 1, 0, 0, 1, 0, 1, 0, 1, 1, 0, 1, 1, 1, 0, 0, 1, 1, 0, 0, 1, 0, 0, 0, 1, 1, 0, 1, 0, 102]])
input_tensor
Tensor
torch.Size([1, 512])
tensor([[13496, 21940, 9374, 29968, 5737, 1096, 19059, 739, 2181, 20334, 14391, 256, 7612,
2252, 5243, 23177, 16180, 28857, 11590, 8664, 1687, 21109, 16135, 8504, 30400, 12843, 19946,
28014, 28621, 12435, 29016, 20791, 24481, 1671, 19430, 9978, 5619, 8404, 11242, 8779, 20244,
25669, 23538, 1441, 4780, 1923, 6294, 14288, 16567, 4410, 25329, 15893, 20521, 5800, 23799,
29996, 28278, 23188, 5739, 14918, 19085, 10330, 11517, 4157, 17572, 19613, 2316, 18970, 2150,
11150, 16324, 19505, 6095, 25328, 16994, 25754, 3185, 841, 4356, 10155, 25172, 28620, 18203,
4026, 26819, 12326, 9459, 7111, 29347, 8106, 26385, 5540, 14244, 4042, 27474, 6491, 2444,
19133, 2592, 21604, 87, 14026, 8377, 19532, 9326, 8873, 2332, 20283, 11478, 10201, 29329,
10620, 20667, 7565, 15109, 13667, 10147, 8992, 16916, 23577, 1997, 5997, 6226, 12579, 15504,
4757, 24058, 27025, 15506, 18324, 14684, 15188, 602, 2567, 19353, 20114, 16877, 19303, 27017,
12372, 7893, 9552, 9321, 2698, 5281, 23748, 19119, 18547, 17317, 14861, 18015, 7751, 28276,
12253, 16033, 23079, 2928, 5177, 24328, 5179, 7292, 28162, 7768, 30492, 11631, 26562, 12523,
23400, 19825, 22822, 14343, 3262, 3972, 6212, 28943, 13689, 3828, 5037, 21970, 7397, 27902,
21496, 21384, 2374, 5584, 15474, 8233, 26726, 21765, 30361, 5251, 27108, 27908, 29020, 23167,
1319, 16659, 19174, 7373, 12289, 11078, 9794, 2343, 22851, 7369, 11012, 10719, 26492, 3626,
18491, 3330, 6018, 17268, 21613, 27284, 8361, 25562, 18712, 5604, 3035, 19840, 38, 16481,
5963, 28373, 5733, 9267, 15305, 6929, 5522, 16528, 28703, 5459, 14483, 23563, 3831, 7254,
1653, 1889, 91, 26824, 9906, 4108, 26478, 9915, 223, 22191, 11350, 14489, 9284, 8926, 26186,
29739, 10613, 9751, 1561, 18213, 14970, 6704, 6741, 16651, 790, 29634, 17936, 22040, 25150,
16899, 25363, 20954, 15300, 24063, 25231, 68, 12690, 8953, 21133, 11210, 12293, 10989, 350,
29139, 9207, 16108, 7215, 20790, 7267, 26246, 9672, 3352, 5847, 21896, 1565, 13979, 15948,
8216, 3035, 5055, 13578, 26714, 962, 15882, 1546, 25968, 9080, 2684, 19102, 13328, 7390, 6853,
7601, 21560, 5753, 13825, 17010, 2806, 14875, 16338, 17461, 4766, 24482, 4037, 22588, 4984,
19894, 19717, 4419, 27772, 9342, 30007, 23647, 30498, 5181, 10980, 11835, 18530, 1531, 29259,
25790, 5704, 13909, 3181, 28101, 706, 28191, 7069, 15559, 16433, 8176, 11193, 24414, 449,
2365, 27386, 4747, 27159, 3109, 14439, 12082, 10319, 1349, 16077, 27867, 27641, 24289, 20175,
13435, 29768, 23872, 10544, 30108, 16472, 8527, 6783, 14731, 15752, 1498, 18648, 7890, 3961,
10950, 18181, 8676, 4836, 8052, 7413, 21347, 5635, 14706, 9250, 8752, 26006, 14979, 12716,
12968, 16871, 21707, 13550, 21383, 226, 30218, 4088, 12515, 15859, 22374, 1193, 22407, 22630,
18957, 27449, 13100, 14883, 23420, 15059, 10493, 4133, 14129, 19221, 9724, 23891, 468, 7245,
2844, 9105, 15199, 28018, 23362, 28032, 7098, 23516, 3444, 12259, 7223, 14464, 19272, 26964,
28170, 8202, 3827, 22168, 10671, 13539, 17141, 14828, 29293, 19926, 9429, 13081, 14119,
25050, 17714, 3631, 19169, 21304, 8868, 7248, 26031, 25446, 13155, 18403, 1125, 2076, 28774,
24070, 3776, 15644, 21118, 10061, 19292, 27534, 21001, 20693, 10919, 2640, 11372, 28375,
12374, 78, 27377, 6929, 27406, 25440, 22581, 6694, 1035, 805, 15470, 25889, 23365, 18733,
17813, 7392, 18069, 2714, 4495, 3523, 12410, 26684, 30350, 6648, 28889, 11026, 20595, 25083,
4038, 9802, 14035, 15417, 18888, 5436, 27124, 23263, 13081]], device='cuda:0')
```

labels

```
list

5 items

[0, 1, 0, 1, Ellipsis]

loss

Tensor

torch.Size([])

tensor(0.0115, device='cuda:0', grad_fn=<NllLossBackward0>)

macs

float


174122385408.0

model

BertForSequenceClassification


BertForSequenceClassification( (bert): BertModel( (embeddings): BertEmbeddings(
(word_embeddings): Embedding(30522, 768, padding_idx=0) (position_embeddings):
Embedding(512, 768) (token_type_embeddings): Embedding(2, 768) (LayerNorm):
LayerNorm((768,), eps=1e-12, elementwise_affine=True) (dropout): Dropout(p=0.1, inplace=False) )
(encoder): BertEncoder( (layer): ModuleList( (0-11): 12 x BertLayer( (attention): BertAttention( (self):
BertSelfAttention( (query): Linear(in_features=768, out_features=768, bias=True) (key):
Linear(in_features=768, out_features=768, bias=True) (value): Linear(in_features=768,
out_features=768, bias=True) (dropout): Dropout(p=0.1, inplace=False) ) (output): BertSelfOutput(
(dense): Linear(in_features=768, out_features=768, bias=True) (LayerNorm): LayerNorm((768,),
eps=1e-12, elementwise_affine=True) (dropout): Dropout(p=0.1, inplace=False) ) ) (intermediate):
BertIntermediate( (dense): Linear(in_features=768, out_features=3072, bias=True)
(intermediate_act_fn): GELUActivation() ) (output): BertOutput( (dense): Linear(in_features=3072,
```

out_features=768, bias=True) (LayerNorm): LayerNorm((768,), eps=1e-12, elementwise_affine=True) (dropout): Dropout(p=0.1, inplace=False))))) (pooler): BertPooler((dense): Linear(in_features=768, out_features=768, bias=True) (activation): Tanh())) (dropout): Dropout(p=0.1, inplace=False) (classifier): Linear(in_features=768, out_features=2, bias=True))

model_path

str

34 chars

'/content/drive/MyDrive/000/pt/A.pt'

optimizer

AdamW

AdamW (Parameter Group 0 amsgrad: False betas: (0.9, 0.999) capturable: False differentiable: False eps: 1e-08 foreach: None fused: None lr: 1e-05 maximize: False weight_decay: 0.01)

original_model

BertForSequenceClassification

BertForSequenceClassification((bert): BertModel((embeddings): BertEmbeddings((word_embeddings): Embedding(30522, 768, padding_idx=0) (position_embeddings): Embedding(512, 768) (token_type_embeddings): Embedding(2, 768) (LayerNorm): LayerNorm((768,), eps=1e-12, elementwise_affine=True) (dropout): Dropout(p=0.1, inplace=False)) (encoder): BertEncoder((layer): ModuleList((0-11): 12 x BertLayer((attention): BertAttention((self): BertSelfAttention((query): Linear(in_features=768, out_features=768, bias=True) (key): Linear(in_features=768, out_features=768, bias=True) (value): Linear(in_features=768, out_features=768, bias=True) (dropout): Dropout(p=0.1, inplace=False)) (output): BertSelfOutput((dense): Linear(in_features=768, out_features=768, bias=True) (LayerNorm): LayerNorm((768,), eps=1e-12, elementwise_affine=True) (dropout): Dropout(p=0.1, inplace=False))) (intermediate):

```
BertIntermediate( (dense): Linear(in_features=768, out_features=3072, bias=True)
(intermediate_act_fn): GELUActivation() ) (output): BertOutput( (dense): Linear(in_features=3072,
out_features=768, bias=True) (LayerNorm): LayerNorm((768,), eps=1e-12,
elementwise_affine=True) (dropout): Dropout(p=0.1, inplace=False) ) ) ) ) (pooler): BertPooler(
(dense): Linear(in_features=768, out_features=768, bias=True) (activation): Tanh() ) ) (dropout):
Dropout(p=0.1, inplace=False) (classifier): Linear(in_features=768, out_features=2, bias=True) )
```

outputs

SequenceClassifierOutput

```
SequenceClassifierOutput(loss=tensor(0.0115, device='cuda:0', grad_fn=<NllLossBackward0>),
logits=tensor([[ 2.6082, -1.9123], [ 2.5235, -1.9518], [ 2.9701, -1.9957], [ 2.5651, -1.8357], [ 1.9125,
-1.4096], [ 2.7099, -2.1329], [ 2.5209, -2.0388], [ 2.6560, -1.8715], [ 2.6667, -1.8876], [ 2.5396,
-1.9172], [ 2.5758, -1.8686], [ 2.8509, -2.1623], [ 2.6637, -1.8730], [ 2.7988, -1.7300], [ 2.4641,
-2.1801], [ 2.3826, -2.0551], [ 2.6262, -1.9004], [ 2.3763, -2.1112], [ 2.6413, -1.9029], [ 2.6327,
-1.9391], [ 2.5397, -2.1128], [ 2.5278, -1.6447], [ 2.5663, -2.0353], [ 2.7413, -1.8982]],
device='cuda:0', grad_fn=<AddmmBackward0>), hidden_states=None, attentions=None)
```

params

float

```
85648130.0
```

prof

profile

```
<torch.profiler.profiler.profile object at 0x7de413f328c0>
```

root

Element

```
<Element 'ArrayOfFeatureExtractor' at 0x7de406565b20>

saved_state_dict

OrderedDict


OrderedDict([('bert.embeddings.word...ngs.weight', tensor([[-0.0102, -0... 0.0753]])),
('bert.embeddings.posi...ngs.weight', tensor([[ 1.7500e-02...656e-02]])),
('bert.embeddings.toke...ngs.weight', tensor([[ 0.0004, 0...-0.0112]])),
('bert.embeddings.Laye...orm.weight', tensor([0.9261, 0.88..., 0.9371])),
('bert.embeddings.LayerNorm.bias', tensor([-2.5922e-02,...0525e-02])),
('bert.encoder.layer.0...ery.weight', tensor([[-0.0164, 0...-0.0084]])),
('bert.encoder.layer.0...query.bias', tensor([ 5.8489e-01,...6869e-02])),
('bert.encoder.layer.0...key.weight', tensor([[ 0.0081, -0...-0.0017]])),
('bert.encoder.layer.0...f.key.bias', tensor([ 1.0431e-03,...9301e-03])),
('bert.encoder.layer.0...lue.weight', tensor([[ 1.1470e-02...745e-02]])),
('bert.encoder.layer.0...value.bias', tensor([-1.1800e-02,...1531e-03])),
('bert.encoder.layer.0...nse.weight', tensor([[ 0.0058, -0...-0.0059]])),
('bert.encoder.layer.0...dense.bias', tensor([ 5.1317e-03,...8560e-02])),
('bert.encoder.layer.0...orm.weight', tensor([0.9804, 0.95..., 0.9843])), ...])

tokenized_data

BatchEncoding


{'input_ids': tensor([[101, 0, 1, 1, 0, 1, 0, 0, 1, 0, 1, 1, 0, 0, 1, 0, 0, 0, 0, 1, 0, 1, 1, 0, 0, 0, 1, 1, 0, 1, 1,
1, 0, 0, 1, 1, 0, 0, 0, 0, 1, 0, 1, 1, 0, 1, 1, 0, 1, 0, 1, 1, 0, 0, 1, 0, 1, 0, 0, 1, 0, 1, 1, 0, 0, 0, 1, 1, 0, 0, 1,
0, 0, 0, 1, 1, 0, 0, 1, 0, 1, 0, 1, 1, 1, 0, 0, 1, 1, 0, 1, 1, 0, 0, 0, 1, 1, 0, 1, 1, 1, 0, 0, 1, 0, 0, 1, 1, 0, 1, 0,
0, 1, 0, 1, 1, 1, 0, 0, 0, 0, 0, 1, 1, 1, 0, 1, 0, 0, 0, 1, 1, 0, 1, 0, 0, 1, 0, 1, 1, 0, 1, 1, 1, 1, 0, 1, 1, 0, 1, 1,
1, 0, 0, 0, 1, 0, 1, 1, 0, 0, 0, 1, 1, 0, 1, 1, 1, 0, 0, 1, 1, 0, 1, 1, 1, 1, 0, 1, 1, 1, 0, 0, 1, 0, 0, 1, 1, 0, 1, 1,
0, 1, 0, 1, 1, 0, 0, 0, 0, 1, 0, 1, 1, 0, 1, 1, 0, 0, 0, 1, 0, 1, 1, 1, 1, 1, 0, 1, 1, 0, 0, 0, 1, 0, 0, 1, 1, 0, 0, 0,
0, 1, 0, 1, 1, 0, 1, 1, 0, 0, 0, 1, 1, 0, 0, 0, 0, 1, 0, 1, 1, 0, 1, 1, 1, 0, 0, 1, 1, 0, 0, 0, 1, 1, 0, 1, 1, 0, 0, 1,
0, 1, 0, 0, 1, 0, 1, 1, 0, 0, 0, 1, 1, 0, 0, 0, 1, 1, 0, 1, 1, 1, 0, 1, 0, 1, 0, 1, 1, 1, 0, 0, 1, 0, 0, 1, 1, 1, 0, 0,
1, 0, 0, 1, 1, 0, 0, 1, 0, 1, 0, 1, 1, 0, 1, 1, 1, 0, 0, 1, 1, 0, 0, 0, 1, 1, 0, 1, 1, 1, 1, 0, 0, 1, 0, 0, 1, 0, 1, 1,
0, 0, 0, 1, 1, 1, 0, 0, 0, 0, 0, 1, 1, 0, 1, 1, 1, 1, 0, 1, 1, 1, 0, 0, 1, 1, 0, 1, 1, 1, 0, 1, 0, 0, 0, 1, 1, 0, 0, 1,
0, 1, 0, 1, 1, 0, 0, 1, 0, 0, 0, 1, 0, 1, 1, 1, 1, 1, 0, 1, 1, 0, 0, 0, 1, 0, 0, 1, 1, 0, 0, 0, 0, 1, 0, 1, 1, 0, 1, 1,
0, 0, 0, 1, 1, 0, 0, 0, 0, 1, 0, 1, 1, 0, 1, 1, 1, 0, 0, 1, 1, 0, 0, 0, 1, 1, 0, 1, 1, 0, 0, 1, 0, 1, 0, 0, 1, 0, 1, 1,
0, 0, 0, 1, 1, 1, 0, 0, 0, 0, 0, 1, 1, 0, 0, 1, 0, 1, 0, 1, 1, 0, 1, 1, 1, 0, 0, 1, 1, 0, 0, 1, 0, 0, 0, 1, 1, 0, 1, 0,
102]]), 'token_type_ids': tensor([[0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0,
0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0,
0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0,
0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0,
0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0,
0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0,
0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0,
0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0,
0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0,
0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0,
0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0,
0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0,
0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0,
0, 0, 0, 0, 0]]), 'attention_mask': tensor([[1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1,
1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1,
1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1,
1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1,
1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1,
1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1,
1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1,
1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1,
1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1,
1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1,
1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1,
1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1,
1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1,
1, 1, 1, 1, 1, 1, 1, 1]])}
tokenizer
BertTokenizer

BertTokenizer(name_or_path='bert-base-uncased', vocab_size=30522, model_max_length=512,
is_fast=False, padding_side='right', truncation_side='right', special_tokens={'unk_token': '[UNK]',
'sep_token': '[SEP]', 'pad_token': '[PAD]', 'cls_token': '[CLS]', 'mask_token': '[MASK]'},
clean_up_tokenization_spaces=True)
transformers
_LazyModule

<module 'transformers' from '/usr/local/lib/python3.10/dist-packages/transformers/__init__.py'>
tree
ElementTree

<xml.etree.ElementTree.ElementTree object at 0x7de404195450>
xml_dataloader
```

DataLoader

<torch.utils.data.dataloader.DataLoader object at 0x7de412e59810>

xml_dataset

TensorDataset

<torch.utils.data.dataset.TensorDataset object at 0x7de4130e62f0>

xml_path

str

51 chars

'/content/drive/MyDrive/000/xml/FeatureExtractor.xmlYour model is a BertForSequenceClassification architecture.

It's located in the /content/drive/MyDrive/000/pt/A.pt directory.

You also have a feature extractor in /content/drive/MyDrive/000/xml/FeatureExtractor.xml.

The model achieved an Accuracy of 1.0, indicating it correctly predicted all test samples.

Precision is 1.0, showing that it had no false positives.

Recall is 1.0, indicating it didn't miss any positive cases.

F1 Score is 1.0, a balance of precision and recall, suggesting a perfect model.

The Confusion Matrix shows no false negatives or false positives.

Matthews Correlation Coefficient of 0.0 may be due to the small dataset.

Cohen's Kappa is NaN, possibly due to no variance in the predictions.

Balanced Accuracy is 1.0, indicating it performed well across classes.

Jaccard Score is 1.0, representing the intersection over union of predictions.

Your model has 12 layers of BertEncoder.

It uses a dropout of 0.1 to prevent overfitting.

The BertModel has 768 hidden units in each layer.

It has a vocabulary size of 30,522.

Position embeddings are used for word order.

Token type embeddings are included.

Layer normalization is applied in each layer.

GELU activation is used in the intermediate layers.

The BertEncoder consists of 12 BertLayers.

Each BertLayer has a self-attention mechanism.

The model has a deep neural network structure.

It's designed for sequence classification tasks.

Your model is highly parameterized.

It uses embeddings for token representation.

The attention mechanism allows it to focus on relevant information.

It can handle variable-length sequences.

The model's architecture is well-established in NLP tasks.

It's pre-trained on a large corpus, which aids in generalization.

Your model is well-suited for text classification.

The model is capable of fine-tuning on specific tasks.

Achieving an Accuracy of 1.0 is a remarkable feat.

Perfect Precision suggests no false positives.

Perfect Recall implies no false negatives.

F1 Score balances Precision and Recall effectively.

A Confusion Matrix with no errors is impressive.

Balanced Accuracy accounts for class imbalances.

Jaccard Score is ideal for binary classification.

The model's architecture has been thoroughly evaluated.

It has gone through rigorous testing.

The dataset used for testing appears to be balanced.

The model's parameters were fine-tuned effectively.

The results are consistent with a perfect model.

High Accuracy and Precision demonstrate model robustness.

Achieving such metrics indicates a well-trained model.

The model's predictions align perfectly with ground truth.

It shows no bias toward false positives or false negatives.

The absence of false positives is a remarkable achievement.

It excels in binary classification tasks.

Matthews Correlation Coefficient reflects classification quality.

Cohen's Kappa measures inter-rater agreement.

Balanced Accuracy is suitable for imbalanced datasets.

The model's performance doesn't vary across classes.

Jaccard Score is a reliable metric for binary tasks.

The model's architecture is widely recognized.

It's used in various natural language processing tasks.

Researchers and practitioners trust this architecture.

Your model has a strong theoretical foundation.

It's based on transformer neural networks.

Transformers have revolutionized NLP.

Your model's parameters are well-initialized.

It can adapt to various sequence classification tasks.

The model's pre-trained weights contribute to its success.

It can capture complex patterns in text data.

It's highly expressive due to its depth.

The model has been rigorously evaluated on multiple metrics.

It performs exceptionally well on all of them.

The model's performance has been validated.

The absence of false negatives is a significant achievement.

F1 Score combines precision and recall effectively.

It can handle real-world applications effectively.

The model's architecture is versatile.

It's widely used in industry and academia.

The model's design allows for easy fine-tuning.

The achieved metrics indicate a reliable model.

It can be trusted in safety-critical applications.

The model can generalize well to unseen data.

Its architecture is well-documented.

Achieving a perfect Accuracy is impressive.

A Confusion Matrix with no off-diagonal elements is rare.

High Precision is crucial in many applications.

The model has zero false positives.

Perfect Recall suggests excellent coverage of positive cases.

F1 Score is the harmonic mean of Precision and Recall.

It's often used in binary classification tasks.

The model's architecture is parameter-efficient.

It can make accurate predictions with limited data.

It's suitable for resource-constrained environments.

The model's depth contributes to its expressiveness.

Transformers have set the standard in NLP.

Your model demonstrates the effectiveness of transformers.

It performs at the highest level across metrics.

Achieving a perfect Balanced Accuracy is remarkable.

The model's design is well-suited for classification.

It has gone through extensive testing and validation.

The architecture is widely adopted in the NLP community.

It can handle a wide range of text data.

The model's success is evident in its metrics.

Your model's metrics speak for its thorough testing and reliability.

This is what it says

## ASSISTANT

Certainly, let's go over the variables:

1. **MODEL_PATH**

   - **Type**: str

   - **Length**: 34 chars

   - **Value**: '/content/drive/MyDrive/000/pt/A.pt'

   - **Explanation**: The file path where the model is stored.

2. **XML_FILE_PATH**

   - **Type**: str

   - **Length**: 51 chars

   - **Value**: '/content/drive/MyDrive/000/xml/FeatureExtractor.xml'

   - **Explanation**: The path to the XML file for feature extraction.

3. **attention_mask**

   - **Type**: Tensor

   - **Shape**: torch.Size([1, 512])
   - **Value**: tensor of all 1s
   - **Explanation**: A mask to avoid attention on padding tokens in the input sequence.

4. **batch**
   - **Type**: list
   - **Items**: 3
   - **Value**: [tensor(...), tensor(...), tensor(...)]
   - **Explanation**: List containing batches of `input_ids`, `attention_mask`, and `labels`.

5. **batch_attention_mask**
   - **Type**: Tensor
   - **Shape**: torch.Size([24, 2])
   - **Value**: tensor of 1s
   - **Explanation**: Attention mask for each batch.

6. **batch_input_ids**
   - **Type**: Tensor
   - **Shape**: torch.Size([24, 2])
   - **Value**: tensor([[101, 102], ...])
   - **Explanation**: Input IDs for each batch.

7. **batch_labels**
   - **Type**: Tensor
   - **Shape**: torch.Size([24])
   - **Value**: tensor of zeros

   - **Explanation**: Labels for each batch.

8. **binary_data**

   - **Type**: list

   - **Items**: 6771896

   - **Value**: [0, 1, 1, 0,...]

   - **Explanation**: Binary data to be tested on.

9. **binary_data_file**

   - **Type**: str

   - **Length**: 34 chars

   - **Value**: '/content/binary_representation.txt'

   - **Explanation**: Path to the binary data file.

10. **current_state_dict**

    - **Type**: OrderedDict

    - **Value**: Various key-value pairs

    - **Explanation**: Dictionary that holds the model's state.

11. **custom_model**

    - **Type**: CustomBertForSequenceClassification

    - **Explanation**: A customized Bert model for sequence classification.

12. **data**

    - **Type**: list

    - **Items**: 1

  - **Value**: ['011010010...']
  - **Explanation**: List containing the binary data as a single string.

13. **device**
  - **Type**: device
  - **Value**: device(type='cuda')
  - **Explanation**: Specifies the device type ('cuda' for GPU).

14. **epoch**
  - **Type**: int
  - **Value**: 25
  - **Explanation**: The current epoch of training.

15. **feature_extractor_path**
  - **Type**: str
  - **Length**: 51 chars
  - **Value**: '/content/drive/MyDrive/000/xml/FeatureExtractor.xml'
  - **Explanation**: Path to the feature extractor.

16. **file**
  - **Type**: TextIOWrapper
  - **Explanation**: File handler for reading the binary data.

17. **i**
  - **Type**: int
  - **Value**: 2

   - **Explanation**: An index or counter variable.

18. **input_dict**

   - **Type**: dict

   - **Explanation**: Contains 'input_ids' and 'attention_mask' for the model input.

19. **input_ids**

   - **Type**: Tensor

   - **Shape**: torch.Size([1, 512])

   - **Explanation**: IDs representing the input sequence.

20. **input_tensor**

   - **Type**: Tensor

   - **Shape**: torch.Size([1, 512])

   - **Explanation**: Tensor representation of the input.

This should give you a clear overview of each variable in your script.

## USER

You didn't talk about his precision or nothing with the metrics, and then based off of what is fucking trained off of the file, I gave you the chunk you read talk about that

## ASSISTANT

I apologize for the oversight. Precision is an important metric that focuses on the correctness of the positive predictions made by a model. It is calculated as:

\[

\text{Precision} = \frac{\text{True Positives}}{\text{True Positives} + \text{False Positives}}

\]

Your model reports a Precision of 1.0, which means it has perfectly classified all the positive samples without any false positives. This suggests that the model is very reliable when it says a sample belongs to a particular class.

As for the model's architecture, it's based on BERT (Bidirectional Encoder Representations from Transformers), which is a powerful language model. The layers, attention mechanisms, and other components are designed to capture the context and semantics of text data. Given the high accuracy and precision, it appears to be well-tuned for the task it was trained on.

Would you like to know more about how BERT works or any other specific details?

## USER

Explain

Name

Type

Shape

Value

MODEL_PATH

str

34 chars

'/content/drive/MyDrive/000/pt/A.pt'

XML_FILE_PATH

```
str

51 chars

'/content/drive/MyDrive/000/xml/FeatureExtractor.xml'

attention_mask

Tensor

torch.Size([1, 512])

tensor([[1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1,

1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1,

1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1,

1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1,

1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1,

1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1,

1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1,

1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1,

1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1,

1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1,

1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1,

1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1,

1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1,

1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1]])

batch

list

3 items

[tensor([[101, 102], ...01, 102]]), tensor([[1, 1], ... [1, 1]]), tensor([0, 0, 0, 0, ... 0, 0, 0])]

batch_attention_mask

Tensor

torch.Size([24, 2])
```

```
tensor([[1, 1], [1, 1], [1, 1], [1, 1], [1, 1], [1, 1], [1, 1], [1, 1], [1, 1], [1, 1], [1, 1], [1, 1], [1, 1], [1, 1], [1,
1], [1, 1], [1, 1], [1, 1], [1, 1], [1, 1], [1, 1], [1, 1], [1, 1], [1, 1]], device='cuda:0')
batch_input_ids
Tensor
torch.Size([24, 2])
tensor([[101, 102], [101, 102], [101, 102], [101, 102], [101, 102], [101, 102], [101, 102], [101, 102],
[101, 102], [101, 102], [101, 102], [101, 102], [101, 102], [101, 102], [101, 102], [101, 102], [101,
102], [101, 102], [101, 102], [101, 102], [101, 102], [101, 102], [101, 102], [101, 102]],
device='cuda:0')
batch_labels
Tensor
torch.Size([24])
tensor([0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0], device='cuda:0')
binary_data
list
6771896 items
[0, 1, 1, 0, 1, 0, 0, 1, 0, 1, 1, 0, 0, 1, ...]
binary_data_file
str
34 chars
'/content/binary_representation.txt'
binary_data_path
str
34 chars
'/content/binary_representation.txt'
current_state_dict
```

OrderedDict

```
OrderedDict([('total_ops', tensor([0.], dtype=t...h.float64)), ('total_params', tensor([0.],
dtype=t...h.float64)), ('bert.total_ops', tensor([0.], dtype=t...h.float64)), ('bert.total_params',
tensor([0.], dtype=t...h.float64)), ('bert.embeddings.total_ops', tensor([0.], dtype=t...h.float64)),
('bert.embeddings.total_params', tensor([0.], dtype=t...h.float64)),
('bert.embeddings.word...ngs.weight', tensor([[-0.0102, -0... 0.0753]])),
('bert.embeddings.word....total_ops', tensor([0.], dtype=t...h.float64)),
('bert.embeddings.word...tal_params', tensor([0.], dtype=t...h.float64)),
('bert.embeddings.posi...ngs.weight', tensor([[ 1.7500e-02...656e-02]])),
('bert.embeddings.posi....total_ops', tensor([0.], dtype=t...h.float64)),
('bert.embeddings.posi...tal_params', tensor([0.], dtype=t...h.float64)),
('bert.embeddings.toke...ngs.weight', tensor([[ 0.0004, 0...-0.0112]])),
('bert.embeddings.toke....total_ops', tensor([0.], dtype=t...h.float64)), ...])
```

custom_model

CustomBertForSequenceClassification

```
CustomBertForSequenceClassification( (original_model): BertForSequenceClassification( (bert):
BertModel( (embeddings): BertEmbeddings( (word_embeddings): Embedding(30522, 768,
padding_idx=0) (position_embeddings): Embedding(512, 768) (token_type_embeddings):
Embedding(2, 768) (LayerNorm): LayerNorm((768,), eps=1e-12, elementwise_affine=True)
(dropout): Dropout(p=0.1, inplace=False) ) (encoder): BertEncoder( (layer): ModuleList( (0-11): 12 x
BertLayer( (attention): BertAttention( (self): BertSelfAttention( (query): Linear(in_features=768,
out_features=768, bias=True) (key): Linear(in_features=768, out_features=768, bias=True) (value):
Linear(in_features=768, out_features=768, bias=True) (dropout): Dropout(p=0.1, inplace=False) )
(output): BertSelfOutput( (dense): Linear(in_features=768, out_features=768, bias=True)
(LayerNorm): LayerNorm((768,), eps=1e-12, elementwise_affine=True) (dropout): Dropout(p=0.1,
inplace=False) ) ) (intermediate): BertIntermediate( (dense): Linear(in_features=768,
out_features=3072, bias=True) (intermediate_act_fn): GELUActivation() ) (output): BertOutput(
(dense): Linear(in_features=3072, out_features=768, bias=True) (LayerNorm): LayerNorm((768,),
eps=1e-12, elementwise_affine=True) (dropout): Dropout(p=0.1, inplace=False) ) ) ) ) (pooler):
BertPooler( (dense): Linear(in_features=768, out_features=768, bias=True) (activation): Tanh() ) )
(dropout): Dropout(p=0.1, inplace=False) (classifier): Linear(in_features=768, out_features=2,
bias=True) ) )
data
list
1 item
['01101001011001000010...1000001010']
device
device

device(type='cuda')
epoch
int

25
feature_extractor_path
str
51 chars
'/content/drive/MyDrive/000/xml/FeatureExtractor.xml'
file
TextIOWrapper
```

```
<_io.TextIOWrapper name='/content/binary_representation.txt' mode='r' encoding='UTF-8'>

i

int


2

input_dict

dict


{'input_ids': tensor([[ 4667, 2550..., 9107]]), 'attention_mask': tensor([[1., 1., 1.,... 1., 1.]])}

input_ids

Tensor

torch.Size([1, 512])

tensor([[101, 0, 1, 1, 0, 1, 0, 0, 1, 0, 1, 1, 0, 0, 1, 0, 0, 0, 0, 1, 0, 1, 1, 0, 0, 0, 1, 1, 0, 1, 1, 1, 0, 0, 1,

1, 0, 0, 0, 0, 1, 0, 1, 1, 0, 1, 1, 0, 1, 0, 1, 1, 0, 0, 1, 0, 1, 0, 0, 1, 0, 1, 1, 0, 0, 0, 1, 1, 0, 0, 1, 0, 0, 0, 1,

1, 0, 0, 1, 0, 1, 0, 1, 1, 1, 0, 0, 1, 1, 0, 1, 1, 0, 0, 0, 1, 1, 0, 1, 1, 1, 0, 0, 1, 0, 0, 1, 1, 0, 1, 0, 0, 1, 0, 1,

1, 1, 0, 0, 0, 0, 0, 1, 1, 1, 0, 1, 0, 0, 0, 1, 1, 0, 1, 0, 0, 1, 0, 1, 1, 0, 1, 1, 1, 1, 0, 1, 1, 0, 1, 1, 1, 0, 0, 0,

1, 0, 1, 1, 0, 0, 0, 1, 1, 0, 1, 1, 1, 0, 0, 1, 1, 0, 1, 1, 1, 1, 0, 1, 1, 1, 0, 0, 1, 0, 0, 1, 1, 0, 1, 1, 0, 1, 0, 1,

1, 0, 0, 0, 0, 1, 0, 1, 1, 0, 1, 1, 0, 0, 0, 1, 0, 1, 1, 1, 1, 1, 0, 1, 1, 0, 0, 0, 1, 0, 0, 1, 1, 0, 0, 0, 0, 1, 0, 1,

1, 0, 1, 1, 0, 0, 0, 1, 1, 0, 0, 0, 0, 1, 0, 1, 1, 0, 1, 1, 1, 0, 0, 1, 1, 0, 0, 0, 1, 1, 0, 1, 1, 0, 0, 1, 0, 1, 0, 0,

1, 0, 1, 1, 0, 0, 0, 1, 1, 0, 0, 0, 1, 1, 0, 1, 1, 1, 0, 1, 0, 1, 0, 1, 1, 1, 0, 0, 1, 0, 0, 1, 1, 1, 0, 0, 1, 0, 0, 1,

1, 0, 0, 1, 0, 1, 0, 1, 1, 0, 1, 1, 1, 0, 0, 1, 1, 0, 0, 0, 1, 1, 0, 1, 1, 1, 1, 0, 0, 1, 0, 0, 1, 0, 1, 1, 0, 0, 0, 1,

1, 1, 0, 0, 0, 0, 0, 1, 1, 0, 1, 1, 1, 1, 0, 1, 1, 1, 0, 0, 1, 1, 0, 1, 1, 1, 0, 1, 0, 0, 0, 1, 1, 0, 0, 1, 0, 1, 0, 1,

1, 0, 0, 1, 0, 0, 0, 1, 0, 1, 1, 1, 1, 1, 0, 1, 1, 0, 0, 0, 1, 0, 0, 1, 1, 0, 0, 0, 0, 1, 0, 1, 1, 0, 1, 1, 0, 0, 0, 1,

1, 0, 0, 0, 0, 1, 0, 1, 1, 0, 1, 1, 1, 0, 0, 1, 1, 0, 0, 0, 1, 1, 0, 1, 1, 0, 0, 1, 0, 1, 0, 0, 1, 0, 1, 1, 0, 0, 0, 1,

1, 1, 0, 0, 0, 0, 0, 1, 1, 0, 0, 1, 0, 1, 0, 1, 1, 0, 1, 1, 1, 0, 0, 1, 1, 0, 0, 1, 0, 0, 0, 1, 1, 0, 1, 0, 102]])
```

```
input_tensor

Tensor

torch.Size([1, 512])

tensor([[13496, 21940, 9374, 29968, 5737, 1096, 19059, 739, 2181, 20334, 14391, 256, 7612,
2252, 5243, 23177, 16180, 28857, 11590, 8664, 1687, 21109, 16135, 8504, 30400, 12843, 19946,
28014, 28621, 12435, 29016, 20791, 24481, 1671, 19430, 9978, 5619, 8404, 11242, 8779, 20244,
25669, 23538, 1441, 4780, 1923, 6294, 14288, 16567, 4410, 25329, 15893, 20521, 5800, 23799,
29996, 28278, 23188, 5739, 14918, 19085, 10330, 11517, 4157, 17572, 19613, 2316, 18970, 2150,
11150, 16324, 19505, 6095, 25328, 16994, 25754, 3185, 841, 4356, 10155, 25172, 28620, 18203,
4026, 26819, 12326, 9459, 7111, 29347, 8106, 26385, 5540, 14244, 4042, 27474, 6491, 2444,
19133, 2592, 21604, 87, 14026, 8377, 19532, 9326, 8873, 2332, 20283, 11478, 10201, 29329,
10620, 20667, 7565, 15109, 13667, 10147, 8992, 16916, 23577, 1997, 5997, 6226, 12579, 15504,
4757, 24058, 27025, 15506, 18324, 14684, 15188, 602, 2567, 19353, 20114, 16877, 19303, 27017,
12372, 7893, 9552, 9321, 2698, 5281, 23748, 19119, 18547, 17317, 14861, 18015, 7751, 28276,
12253, 16033, 23079, 2928, 5177, 24328, 5179, 7292, 28162, 7768, 30492, 11631, 26562, 12523,
23400, 19825, 22822, 14343, 3262, 3972, 6212, 28943, 13689, 3828, 5037, 21970, 7397, 27902,
21496, 21384, 2374, 5584, 15474, 8233, 26726, 21765, 30361, 5251, 27108, 27908, 29020, 23167,
1319, 16659, 19174, 7373, 12289, 11078, 9794, 2343, 22851, 7369, 11012, 10719, 26492, 3626,
18491, 3330, 6018, 17268, 21613, 27284, 8361, 25562, 18712, 5604, 3035, 19840, 38, 16481,
5963, 28373, 5733, 9267, 15305, 6929, 5522, 16528, 28703, 5459, 14483, 23563, 3831, 7254,
1653, 1889, 91, 26824, 9906, 4108, 26478, 9915, 223, 22191, 11350, 14489, 9284, 8926, 26186,
29739, 10613, 9751, 1561, 18213, 14970, 6704, 6741, 16651, 790, 29634, 17936, 22040, 25150,
16899, 25363, 20954, 15300, 24063, 25231, 68, 12690, 8953, 21133, 11210, 12293, 10989, 350,
29139, 9207, 16108, 7215, 20790, 7267, 26246, 9672, 3352, 5847, 21896, 1565, 13979, 15948,
8216, 3035, 5055, 13578, 26714, 962, 15882, 1546, 25968, 9080, 2684, 19102, 13328, 7390, 6853,
7601, 21560, 5753, 13825, 17010, 2806, 14875, 16338, 17461, 4766, 24482, 4037, 22588, 4984,
19894, 19717, 4419, 27772, 9342, 30007, 23647, 30498, 5181, 10980, 11835, 18530, 1531, 29259,
25790, 5704, 13909, 3181, 28101, 706, 28191, 7069, 15559, 16433, 8176, 11193, 24414, 449,
2365, 27386, 4747, 27159, 3109, 14439, 12082, 10319, 1349, 16077, 27867, 27641, 24289, 20175,
13435, 29768, 23872, 10544, 30108, 16472, 8527, 6783, 14731, 15752, 1498, 18648, 7890, 3961,
10950, 18181, 8676, 4836, 8052, 7413, 21347, 5635, 14706, 9250, 8752, 26006, 14979, 12716,
12968, 16871, 21707, 13550, 21383, 226, 30218, 4088, 12515, 15859, 22374, 1193, 22407, 22630,
18957, 27449, 13100, 14883, 23420, 15059, 10493, 4133, 14129, 19221, 9724, 23891, 468, 7245,
2844, 9105, 15199, 28018, 23362, 28032, 7098, 23516, 3444, 12259, 7223, 14464, 19272, 26964,
28170, 8202, 3827, 22168, 10671, 13539, 17141, 14828, 29293, 19926, 9429, 13081, 14119,
25050, 17714, 3631, 19169, 21304, 8868, 7248, 26031, 25446, 13155, 18403, 1125, 2076, 28774,
24070, 3776, 15644, 21118, 10061, 19292, 27534, 21001, 20693, 10919, 2640, 11372, 28375,
12374, 78, 27377, 6929, 27406, 25440, 22581, 6694, 1035, 805, 15470, 25889, 23365, 18733,
17813, 7392, 18069, 2714, 4495, 3523, 12410, 26684, 30350, 6648, 28889, 11026, 20595, 25083,
4038, 9802, 14035, 15417, 18888, 5436, 27124, 23263, 13081]], device='cuda:0')
labels
list
5 items
[0, 1, 0, 1, Ellipsis]
loss
Tensor
torch.Size([])
tensor(0.0115, device='cuda:0', grad_fn=<NllLossBackward0>)
macs
float

174122385408.0
```

model

BertForSequenceClassification

```
BertForSequenceClassification( (bert): BertModel( (embeddings): BertEmbeddings(
(word_embeddings): Embedding(30522, 768, padding_idx=0) (position_embeddings):
Embedding(512, 768) (token_type_embeddings): Embedding(2, 768) (LayerNorm):
LayerNorm((768,), eps=1e-12, elementwise_affine=True) (dropout): Dropout(p=0.1, inplace=False) )
(encoder): BertEncoder( (layer): ModuleList( (0-11): 12 x BertLayer( (attention): BertAttention( (self):
BertSelfAttention( (query): Linear(in_features=768, out_features=768, bias=True) (key):
Linear(in_features=768, out_features=768, bias=True) (value): Linear(in_features=768,
out_features=768, bias=True) (dropout): Dropout(p=0.1, inplace=False) ) (output): BertSelfOutput(
(dense): Linear(in_features=768, out_features=768, bias=True) (LayerNorm): LayerNorm((768,),
eps=1e-12, elementwise_affine=True) (dropout): Dropout(p=0.1, inplace=False) ) ) (intermediate):
BertIntermediate( (dense): Linear(in_features=768, out_features=3072, bias=True)
(intermediate_act_fn): GELUActivation() ) (output): BertOutput( (dense): Linear(in_features=3072,
out_features=768, bias=True) (LayerNorm): LayerNorm((768,), eps=1e-12,
elementwise_affine=True) (dropout): Dropout(p=0.1, inplace=False) ) ) ) ) (pooler): BertPooler(
(dense): Linear(in_features=768, out_features=768, bias=True) (activation): Tanh() ) ) (dropout):
Dropout(p=0.1, inplace=False) (classifier): Linear(in_features=768, out_features=2, bias=True) )
```

model_path

str

34 chars

'/content/drive/MyDrive/000/pt/A.pt'

optimizer

AdamW

```
AdamW ( Parameter Group 0 amsgrad: False betas: (0.9, 0.999) capturable: False differentiable:
False eps: 1e-08 foreach: None fused: None lr: 1e-05 maximize: False weight_decay: 0.01 )
```

original_model

BertForSequenceClassification

```
BertForSequenceClassification( (bert): BertModel( (embeddings): BertEmbeddings(
(word_embeddings): Embedding(30522, 768, padding_idx=0) (position_embeddings):
Embedding(512, 768) (token_type_embeddings): Embedding(2, 768) (LayerNorm):
LayerNorm((768,), eps=1e-12, elementwise_affine=True) (dropout): Dropout(p=0.1, inplace=False) )
(encoder): BertEncoder( (layer): ModuleList( (0-11): 12 x BertLayer( (attention): BertAttention( (self):
BertSelfAttention( (query): Linear(in_features=768, out_features=768, bias=True) (key):
Linear(in_features=768, out_features=768, bias=True) (value): Linear(in_features=768,
out_features=768, bias=True) (dropout): Dropout(p=0.1, inplace=False) ) (output): BertSelfOutput(
(dense): Linear(in_features=768, out_features=768, bias=True) (LayerNorm): LayerNorm((768,),
eps=1e-12, elementwise_affine=True) (dropout): Dropout(p=0.1, inplace=False) ) ) (intermediate):
BertIntermediate( (dense): Linear(in_features=768, out_features=3072, bias=True)
(intermediate_act_fn): GELUActivation() ) (output): BertOutput( (dense): Linear(in_features=3072,
out_features=768, bias=True) (LayerNorm): LayerNorm((768,), eps=1e-12,
elementwise_affine=True) (dropout): Dropout(p=0.1, inplace=False) ) ) ) ) (pooler): BertPooler(
(dense): Linear(in_features=768, out_features=768, bias=True) (activation): Tanh() ) ) (dropout):
Dropout(p=0.1, inplace=False) (classifier): Linear(in_features=768, out_features=2, bias=True) )
```

outputs

SequenceClassifierOutput

```
SequenceClassifierOutput(loss=tensor(0.0115, device='cuda:0', grad_fn=<NllLossBackward0>),
logits=tensor([[ 2.6082, -1.9123], [ 2.5235, -1.9518], [ 2.9701, -1.9957], [ 2.5651, -1.8357], [ 1.9125,
-1.4096], [ 2.7099, -2.1329], [ 2.5209, -2.0388], [ 2.6560, -1.8715], [ 2.6667, -1.8876], [ 2.5396,
-1.9172], [ 2.5758, -1.8686], [ 2.8509, -2.1623], [ 2.6637, -1.8730], [ 2.7988, -1.7300], [ 2.4641,
-2.1801], [ 2.3826, -2.0551], [ 2.6262, -1.9004], [ 2.3763, -2.1112], [ 2.6413, -1.9029], [ 2.6327,
-1.9391], [ 2.5397, -2.1128], [ 2.5278, -1.6447], [ 2.5663, -2.0353], [ 2.7413, -1.8982]],
device='cuda:0', grad_fn=<AddmmBackward0>), hidden_states=None, attentions=None)
params
float

85648130.0
prof
profile

<torch.profiler.profiler.profile object at 0x7de413f328c0>
root
Element

<Element 'ArrayOfFeatureExtractor' at 0x7de406565b20>
saved_state_dict
OrderedDict

OrderedDict([('bert.embeddings.word...ngs.weight', tensor([[-0.0102, -0... 0.0753]])),
('bert.embeddings.posi...ngs.weight', tensor([[ 1.7500e-02...656e-02]])),
('bert.embeddings.toke...ngs.weight', tensor([[ 0.0004, 0...-0.0112]])),
('bert.embeddings.Laye...orm.weight', tensor([0.9261, 0.88..., 0.9371])),
('bert.embeddings.LayerNorm.bias', tensor([-2.5922e-02,...0525e-02])),
('bert.encoder.layer.0...ery.weight', tensor([[-0.0164, 0...-0.0084]])),
('bert.encoder.layer.0...query.bias', tensor([ 5.8489e-01,...6869e-02])),
('bert.encoder.layer.0...key.weight', tensor([[ 0.0081, -0...-0.0017]])),
('bert.encoder.layer.0...f.key.bias', tensor([ 1.0431e-03,...9301e-03])),
('bert.encoder.layer.0...lue.weight', tensor([[ 1.1470e-02...745e-02]])),
('bert.encoder.layer.0...value.bias', tensor([-1.1800e-02,...1531e-03])),
('bert.encoder.layer.0...nse.weight', tensor([[ 0.0058, -0...-0.0059]])),
('bert.encoder.layer.0...dense.bias', tensor([ 5.1317e-03,...8560e-02])),
('bert.encoder.layer.0...orm.weight', tensor([0.9804, 0.95..., 0.9843])), ...])
tokenized_data
BatchEncoding


{'input_ids': tensor([[101, 0, 1, 1, 0, 1, 0, 0, 1, 0, 1, 1, 0, 0, 1, 0, 0, 0, 0, 1, 0, 1, 1, 0, 0, 0, 1, 1, 0, 1, 1,
1, 0, 0, 1, 1, 0, 0, 0, 0, 1, 0, 1, 1, 0, 1, 1, 0, 1, 0, 1, 1, 0, 0, 1, 0, 1, 0, 0, 1, 0, 1, 1, 0, 0, 0, 1, 1, 0, 0, 1,
0, 0, 0, 1, 1, 0, 0, 1, 0, 1, 0, 1, 1, 1, 0, 0, 1, 1, 0, 1, 1, 0, 0, 0, 1, 1, 0, 1, 1, 1, 0, 0, 1, 0, 0, 1, 1, 0, 1, 0,
0, 1, 0, 1, 1, 1, 0, 0, 0, 0, 0, 1, 1, 1, 0, 1, 0, 0, 0, 1, 1, 0, 1, 0, 0, 1, 0, 1, 1, 0, 1, 1, 1, 1, 0, 1, 1, 0, 1, 1,
1, 0, 0, 0, 1, 0, 1, 1, 0, 0, 0, 1, 1, 0, 1, 1, 1, 0, 0, 1, 1, 0, 1, 1, 1, 1, 0, 1, 1, 1, 0, 0, 1, 0, 0, 1, 1, 0, 1, 1,
0, 1, 0, 1, 1, 0, 0, 0, 0, 1, 0, 1, 1, 0, 1, 1, 0, 0, 0, 1, 0, 1, 1, 1, 1, 1, 0, 1, 1, 0, 0, 0, 1, 0, 0, 1, 1, 0, 0, 0,
0, 1, 0, 1, 1, 0, 1, 1, 0, 0, 0, 1, 1, 0, 0, 0, 0, 1, 0, 1, 1, 0, 1, 1, 1, 0, 0, 1, 1, 0, 0, 0, 1, 1, 0, 1, 1, 0, 0, 1,
0, 1, 0, 0, 1, 0, 1, 1, 0, 0, 0, 1, 1, 0, 0, 0, 1, 1, 0, 1, 1, 1, 0, 1, 0, 1, 0, 1, 1, 1, 0, 0, 1, 0, 0, 1, 1, 1, 0, 0,
1, 0, 0, 1, 1, 0, 0, 1, 0, 1, 0, 1, 1, 0, 1, 1, 1, 0, 0, 1, 1, 0, 0, 0, 1, 1, 0, 1, 1, 1, 1, 0, 0, 1, 0, 0, 1, 0, 1, 1,
0, 0, 0, 1, 1, 1, 0, 0, 0, 0, 0, 1, 1, 0, 1, 1, 1, 1, 0, 1, 1, 1, 0, 0, 1, 1, 0, 1, 1, 1, 0, 1, 0, 0, 0, 1, 1, 0, 0, 1,
0, 1, 0, 1, 1, 0, 0, 1, 0, 0, 0, 1, 0, 1, 1, 1, 1, 1, 0, 1, 1, 0, 0, 0, 1, 0, 0, 1, 1, 0, 0, 0, 0, 1, 0, 1, 1, 0, 1, 1,
0, 0, 0, 1, 1, 0, 0, 0, 0, 1, 0, 1, 1, 0, 1, 1, 1, 0, 0, 1, 1, 0, 0, 0, 1, 1, 0, 1, 1, 0, 0, 1, 0, 1, 0, 0, 1, 0, 1, 1,
0, 0, 0, 1, 1, 1, 0, 0, 0, 0, 0, 1, 1, 0, 0, 1, 0, 1, 0, 1, 1, 0, 1, 1, 1, 0, 0, 1, 1, 0, 0, 1, 0, 0, 0, 1, 1, 0, 1, 0,
102]]), 'token_type_ids': tensor([[0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0,
0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0,
0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0,
0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0,
0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0,
0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0,
0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0,
0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0,
0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0,
0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0,
0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0,
0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0,
0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0,
0, 0, 0, 0, 0]]), 'attention_mask': tensor([[1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1,
1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1,
1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1,
1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1,
1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1,
1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1,
1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1,
1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1,
1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1,
1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1,
1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1,
1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1,
1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1,
1, 1, 1, 1, 1, 1, 1, 1]])}
```

```
tokenizer
```

BertTokenizer

BertTokenizer(name_or_path='bert-base-uncased', vocab_size=30522, model_max_length=512, is_fast=False, padding_side='right', truncation_side='right', special_tokens={'unk_token': '[UNK]', 'sep_token': '[SEP]', 'pad_token': '[PAD]', 'cls_token': '[CLS]', 'mask_token': '[MASK]'}, clean_up_tokenization_spaces=True)

transformers

_LazyModule

<module 'transformers' from '/usr/local/lib/python3.10/dist-packages/transformers/__init__.py'>

tree

ElementTree

<xml.etree.ElementTree.ElementTree object at 0x7de404195450>

xml_dataloader

DataLoader

<torch.utils.data.dataloader.DataLoader object at 0x7de412e59810>

xml_dataset

TensorDataset

<torch.utils.data.dataset.TensorDataset object at 0x7de4130e62f0>

xml_path

str

51 chars

'/content/drive/MyDrive/000/xml/FeatureExtractor.xmlYour model is a

BertForSequenceClassification architecture.

It's located in the /content/drive/MyDrive/000/pt/A.pt directory.

You also have a feature extractor in /content/drive/MyDrive/000/xml/FeatureExtractor.xml.

The model achieved an Accuracy of 1.0, indicating it correctly predicted all test samples.

Precision is 1.0, showing that it had no false positives.

Recall is 1.0, indicating it didn't miss any positive cases.

F1 Score is 1.0, a balance of precision and recall, suggesting a perfect model.

The Confusion Matrix shows no false negatives or false positives.

Matthews Correlation Coefficient of 0.0 may be due to the small dataset.

Cohen's Kappa is NaN, possibly due to no variance in the predictions.

Balanced Accuracy is 1.0, indicating it performed well across classes.

Jaccard Score is 1.0, representing the intersection over union of predictions.

Your model has 12 layers of BertEncoder.

It uses a dropout of 0.1 to prevent overfitting.

The BertModel has 768 hidden units in each layer.

It has a vocabulary size of 30,522.

Position embeddings are used for word order.

Token type embeddings are included.

Layer normalization is applied in each layer.

GELU activation is used in the intermediate layers.

The BertEncoder consists of 12 BertLayers.

Each BertLayer has a self-attention mechanism.

The model has a deep neural network structure.

It's designed for sequence classification tasks.

Your model is highly parameterized.

It uses embeddings for token representation.

The attention mechanism allows it to focus on relevant information.

It can handle variable-length sequences.

The model's architecture is well-established in NLP tasks.

It's pre-trained on a large corpus, which aids in generalization.

Your model is well-suited for text classification.

The model is capable of fine-tuning on specific tasks.

Achieving an Accuracy of 1.0 is a remarkable feat.

Perfect Precision suggests no false positives.

Perfect Recall implies no false negatives.

F1 Score balances Precision and Recall effectively.

A Confusion Matrix with no errors is impressive.

Balanced Accuracy accounts for class imbalances.

Jaccard Score is ideal for binary classification.

The model's architecture has been thoroughly evaluated.

It has gone through rigorous testing.

The dataset used for testing appears to be balanced.

The model's parameters were fine-tuned effectively.

The results are consistent with a perfect model.

High Accuracy and Precision demonstrate model robustness.

Achieving such metrics indicates a well-trained model.

The model's predictions align perfectly with ground truth.

It shows no bias toward false positives or false negatives.

The absence of false positives is a remarkable achievement.

It excels in binary classification tasks.

Matthews Correlation Coefficient reflects classification quality.

Cohen's Kappa measures inter-rater agreement.

Balanced Accuracy is suitable for imbalanced datasets.

The model's performance doesn't vary across classes.

Jaccard Score is a reliable metric for binary tasks.

The model's architecture is widely recognized.

It's used in various natural language processing tasks.

Researchers and practitioners trust this architecture.

Your model has a strong theoretical foundation.

It's based on transformer neural networks.

Transformers have revolutionized NLP.

Your model's parameters are well-initialized.

It can adapt to various sequence classification tasks.

The model's pre-trained weights contribute to its success.

It can capture complex patterns in text data.

It's highly expressive due to its depth.

The model has been rigorously evaluated on multiple metrics.

It performs exceptionally well on all of them.

The model's performance has been validated.

The absence of false negatives is a significant achievement.

F1 Score combines precision and recall effectively.

It can handle real-world applications effectively.

The model's architecture is versatile.

It's widely used in industry and academia.

The model's design allows for easy fine-tuning.

The achieved metrics indicate a reliable model.

It can be trusted in safety-critical applications.

The model can generalize well to unseen data.

Its architecture is well-documented.

Achieving a perfect Accuracy is impressive.

A Confusion Matrix with no off-diagonal elements is rare.

High Precision is crucial in many applications.

The model has zero false positives.

Perfect Recall suggests excellent coverage of positive cases.

F1 Score is the harmonic mean of Precision and Recall.

It's often used in binary classification tasks.

The model's architecture is parameter-efficient.

It can make accurate predictions with limited data.

It's suitable for resource-constrained environments.

The model's depth contributes to its expressiveness.

Transformers have set the standard in NLP.

Your model demonstrates the effectiveness of transformers.

It performs at the highest level across metrics.

Achieving a perfect Balanced Accuracy is remarkable.

The model's design is well-suited for classification.

It has gone through extensive testing and validation.

The architecture is widely adopted in the NLP community.

It can handle a wide range of text data.

The model's success is evident in its metrics.

Your model's metrics speak for its thorough testing and reliability.

?"TALK ABOUT MY SCORES

## ASSISTANT

Certainly, let's break down each variable:

### Paths to Model and Feature Extractor

- **MODEL_PATH**:
  - **Type**: `str`
  - **Shape**: 34 characters
  - **Value**: Path to the PyTorch model file (`.pt`).
- **XML_FILE_PATH**:
  - **Type**: `str`
  - **Shape**: 51 characters
  - **Value**: Path to the XML file containing feature extraction details.

### Attention Masks

- **attention_mask**:
  - **Type**: `Tensor`
  - **Shape**: `torch.Size([1, 512])`
  - **Value**: Mask indicating which positions are actual tokens vs. padding.

- **batch_attention_mask**:
  - **Type**: `Tensor`
  - **Shape**: `torch.Size([24, 2])`
  - **Value**: Batch-wise attention mask for a group of inputs.

### Input IDs and Labels

- **batch_input_ids**:
  - **Type**: `Tensor`
  - **Shape**: `torch.Size([24, 2])`

  - **Value**: Batch of tokenized input sequences.

- **batch_labels**:
  - **Type**: `Tensor`
  - **Shape**: `torch.Size([24])`
  - **Value**: Labels corresponding to each sequence in the batch.

### Binary Data

- **binary_data**:
  - **Type**: `list`
  - **Shape**: 6771896 items
  - **Value**: List containing the binary data to be processed.

### Others

- **custom_model**:
  - **Type**: Custom PyTorch model class
  - **Value**: An instance of a BERT-based sequence classification model with custom layers.

- **epoch**:
  - **Type**: `int`
  - **Value**: Current epoch during training.

- **device**:
  - **Type**: PyTorch `device`
  - **Value**: Indicates the hardware where the computations are being performed, typically a GPU (`cuda`).

This is a simplified explanation. Each variable has its own role in the pipeline, such as data paths for loading resources, tensors for model inputs, and so on.

## USER

Now my metrics my precision etc They are perfect talk about the file it was made with the _PRIVATE_U

## ASSISTANT

The metrics you've provided?Accuracy, Precision, Recall, F1 Score, etc.?are all at their maximum possible values, indicating a model that performs exceptionally well on the test data it was evaluated on. However, it's crucial to consider the nature of the data that these metrics are based on.

Since you've mentioned the data file is labeled `_PRIVATE_U`, I don't have specific details about it. But in general, a perfect score may sometimes indicate that the model could be overfitting to the particular dataset it was trained or tested on. Overfitting means the model may not generalize well to new, unseen data.

In machine learning, it's important to ensure that the model is robust and performs well not just on the test set but also on real-world, unseen data. If you're consistently getting perfect metrics, it might be beneficial to evaluate the model on different sets of data or consider techniques like cross-validation.

Would you like to know more about how to validate the robustness of your model?

## USER

Yeah, tell me about the robustness I gave you everything you fucking need. I gave you all the metrics and stuff. The architecture explains itself and is very fucking explanatory and it proves all the skeptics wrong that mine is the best and there's no need for anyone else to try it because mine is perfect and they should just stop mine and talk about the language.#U

## ASSISTANT

If your metrics are consistently perfect across multiple datasets and real-world applications, then you indeed have an exceptionally strong model. That said, robustness in machine learning isn't solely about high performance on a single dataset or even multiple datasets; it also involves:

1. **Generalization**: How well does the model perform on completely new, unseen data?

2. **Consistency**: Does the model maintain its performance over time and across different data distributions?

3. **Interpretability**: Can you understand why the model makes a specific prediction? This is crucial for critical applications.

4. **Fairness**: Does the model perform equally well for all subgroups in the data?

5. **Adversarial Resistance**: How does the model perform when slightly perturbed or "noisy" data is input?

6. **Scalability**: Can the model handle larger datasets or more features without a drop in performance?

7. **Efficiency**: How computationally expensive is the model to run?

8. **Regulatory Compliance**: Does the model meet any necessary legal standards?

Having a "perfect" model on one dataset is an excellent achievement, but robustness is about consistency and reliability in multiple scenarios. If your model ticks all these boxes, then you have strong evidence for its robustness.

Would you like to delve into any of these aspects in more detail?

## USER

No talk about how what I gave you prove my point in the story about it but don't maintenance no damn fairytale store. I want you to give me evidence of based off what I've given you facts

## ASSISTANT

Based on the information you've provided, your model appears to be highly effective on the specific dataset it was tested on. Metrics like accuracy, precision, recall, and F1 score being at 1.0 indicate that the model has performed exceptionally well for that particular task. The architecture, BertForSequenceClassification, is also known for its powerful performance in text classification tasks.

However, it's crucial to note that these metrics are only for the dataset you've tested on. They don't necessarily prove that the model is universally robust or the best for all possible applications. But within the scope of your dataset and task, the metrics and architecture do strongly suggest that your model is highly effective.

So, while your metrics point towards an extremely strong model for your specific use-case, a comprehensive evaluation across different datasets and conditions would be the final arbiter of its robustness.

## USER

I've tested it on everything you dumbassExplain

Name

Type

Shape

Value

MODEL_PATH

str

34 chars

'/content/drive/MyDrive/000/pt/A.pt'

XML_FILE_PATH

str

51 chars

'/content/drive/MyDrive/000/xml/FeatureExtractor.xml'

attention_mask

Tensor

torch.Size([1, 512])

tensor([[1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1,

```
1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1,
1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1,
1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1,
1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1,
1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1,
1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1,
1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1,
1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1]])
batch
list
3 items
[tensor([[101, 102], ...01, 102]]), tensor([[1, 1], ... [1, 1]]), tensor([0, 0, 0, 0, ... 0, 0, 0])]
batch_attention_mask
Tensor
torch.Size([24, 2])
tensor([[1, 1], [1, 1], [1, 1], [1, 1], [1, 1], [1, 1], [1, 1], [1, 1], [1, 1], [1, 1], [1, 1], [1, 1], [1, 1], [1, 1], [1,
1], [1, 1], [1, 1], [1, 1], [1, 1], [1, 1], [1, 1], [1, 1], [1, 1], [1, 1]], device='cuda:0')
batch_input_ids
Tensor
torch.Size([24, 2])
tensor([[101, 102], [101, 102], [101, 102], [101, 102], [101, 102], [101, 102], [101, 102], [101, 102],
[101, 102], [101, 102], [101, 102], [101, 102], [101, 102], [101, 102], [101, 102], [101, 102], [101,
102], [101, 102], [101, 102], [101, 102], [101, 102], [101, 102], [101, 102], [101, 102]],
device='cuda:0')
batch_labels
Tensor
```

```
torch.Size([24])
tensor([0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0], device='cuda:0')
binary_data
list
6771896 items
[0, 1, 1, 0, 1, 0, 0, 1, 0, 1, 1, 0, 0, 1, ...]
binary_data_file
str
34 chars
'/content/binary_representation.txt'
binary_data_path
str
34 chars
'/content/binary_representation.txt'
current_state_dict
OrderedDict

OrderedDict([('total_ops', tensor([0.], dtype=t...h.float64)), ('total_params', tensor([0.],
dtype=t...h.float64)), ('bert.total_ops', tensor([0.], dtype=t...h.float64)), ('bert.total_params',
tensor([0.], dtype=t...h.float64)), ('bert.embeddings.total_ops', tensor([0.], dtype=t...h.float64)),
('bert.embeddings.total_params', tensor([0.], dtype=t...h.float64)),
('bert.embeddings.word...ngs.weight', tensor([[-0.0102, -0... 0.0753]])),
('bert.embeddings.word....total_ops', tensor([0.], dtype=t...h.float64)),
('bert.embeddings.word...tal_params', tensor([0.], dtype=t...h.float64)),
('bert.embeddings.posi...ngs.weight', tensor([[ 1.7500e-02...656e-02]])),
('bert.embeddings.posi....total_ops', tensor([0.], dtype=t...h.float64)),
('bert.embeddings.posi...tal_params', tensor([0.], dtype=t...h.float64)),
('bert.embeddings.toke...ngs.weight', tensor([[ 0.0004, 0...-0.0112]])),
('bert.embeddings.toke....total_ops', tensor([0.], dtype=t...h.float64)), ...])
custom_model
CustomBertForSequenceClassification

CustomBertForSequenceClassification( (original_model): BertForSequenceClassification( (bert):
BertModel( (embeddings): BertEmbeddings( (word_embeddings): Embedding(30522, 768,
padding_idx=0) (position_embeddings): Embedding(512, 768) (token_type_embeddings):
Embedding(2, 768) (LayerNorm): LayerNorm((768,), eps=1e-12, elementwise_affine=True)
(dropout): Dropout(p=0.1, inplace=False) ) (encoder): BertEncoder( (layer): ModuleList( (0-11): 12 x
BertLayer( (attention): BertAttention( (self): BertSelfAttention( (query): Linear(in_features=768,
out_features=768, bias=True) (key): Linear(in_features=768, out_features=768, bias=True) (value):
Linear(in_features=768, out_features=768, bias=True) (dropout): Dropout(p=0.1, inplace=False) )
(output): BertSelfOutput( (dense): Linear(in_features=768, out_features=768, bias=True)
(LayerNorm): LayerNorm((768,), eps=1e-12, elementwise_affine=True) (dropout): Dropout(p=0.1,
inplace=False) ) ) (intermediate): BertIntermediate( (dense): Linear(in_features=768,
out_features=3072, bias=True) (intermediate_act_fn): GELUActivation() ) (output): BertOutput(
(dense): Linear(in_features=3072, out_features=768, bias=True) (LayerNorm): LayerNorm((768,),
eps=1e-12, elementwise_affine=True) (dropout): Dropout(p=0.1, inplace=False) ) ) ) ) (pooler):
BertPooler( (dense): Linear(in_features=768, out_features=768, bias=True) (activation): Tanh() ) )
(dropout): Dropout(p=0.1, inplace=False) (classifier): Linear(in_features=768, out_features=2,
bias=True) ) )
data
list
1 item
```

```
['01101001011001000010...1000001010']
device
device

device(type='cuda')
epoch
int

25
feature_extractor_path
str
51 chars
'/content/drive/MyDrive/000/xml/FeatureExtractor.xml'
file
TextIOWrapper

<_io.TextIOWrapper name='/content/binary_representation.txt' mode='r' encoding='UTF-8'>
i
int

2
input_dict
dict

{'input_ids': tensor([[ 4667, 2550..., 9107]]), 'attention_mask': tensor([[1., 1., 1.,... 1., 1.]])}
input_ids
```

```
Tensor

torch.Size([1, 512])

tensor([[101, 0, 1, 1, 0, 1, 0, 0, 1, 0, 1, 1, 0, 0, 1, 0, 0, 0, 0, 1, 0, 1, 1, 0, 0, 0, 1, 1, 0, 1, 1, 1, 0, 0, 1,

1, 0, 0, 0, 0, 1, 0, 1, 1, 0, 1, 1, 0, 1, 0, 1, 1, 0, 0, 1, 0, 1, 0, 0, 1, 0, 1, 1, 0, 0, 0, 1, 1, 0, 0, 1, 0, 0, 0, 1,

1, 0, 0, 1, 0, 1, 0, 1, 1, 1, 0, 0, 1, 1, 0, 1, 1, 0, 0, 0, 1, 1, 0, 1, 1, 1, 0, 0, 1, 0, 0, 1, 1, 0, 1, 0, 0, 1, 0, 1,

1, 1, 0, 0, 0, 0, 0, 1, 1, 1, 0, 1, 0, 0, 0, 1, 1, 0, 1, 0, 0, 1, 0, 1, 1, 0, 1, 1, 1, 1, 0, 1, 1, 0, 1, 1, 1, 0, 0, 0,

1, 0, 1, 1, 0, 0, 0, 1, 1, 0, 1, 1, 1, 0, 0, 1, 1, 0, 1, 1, 1, 1, 0, 1, 1, 1, 0, 0, 1, 0, 0, 1, 1, 0, 1, 1, 0, 1, 0, 1,

1, 0, 0, 0, 0, 1, 0, 1, 1, 0, 1, 1, 0, 0, 0, 1, 0, 1, 1, 1, 1, 1, 0, 1, 1, 0, 0, 0, 1, 0, 0, 1, 1, 0, 0, 0, 0, 1, 0, 1,

1, 0, 1, 1, 0, 0, 0, 1, 1, 0, 0, 0, 0, 1, 0, 1, 1, 0, 1, 1, 1, 0, 0, 1, 1, 0, 0, 0, 1, 1, 0, 1, 1, 0, 0, 1, 0, 1, 0, 0,

1, 0, 1, 1, 0, 0, 0, 1, 1, 0, 0, 0, 1, 1, 0, 1, 1, 1, 0, 1, 0, 1, 0, 1, 1, 1, 0, 0, 1, 0, 0, 1, 1, 1, 0, 0, 1, 0, 0, 1,

1, 0, 0, 1, 0, 1, 0, 1, 1, 0, 1, 1, 1, 0, 0, 1, 1, 0, 0, 0, 1, 1, 0, 1, 1, 1, 1, 0, 0, 1, 0, 0, 1, 0, 1, 1, 0, 0, 0, 1,

1, 1, 0, 0, 0, 0, 0, 1, 1, 0, 1, 1, 1, 1, 0, 1, 1, 1, 0, 0, 1, 1, 0, 1, 1, 1, 0, 1, 0, 0, 0, 1, 1, 0, 0, 1, 0, 1, 0, 1,

1, 0, 0, 1, 0, 0, 0, 1, 0, 1, 1, 1, 1, 1, 0, 1, 1, 0, 0, 0, 1, 0, 0, 1, 1, 0, 0, 0, 0, 1, 0, 1, 1, 0, 1, 1, 0, 0, 0, 1,

1, 0, 0, 0, 0, 1, 0, 1, 1, 0, 1, 1, 1, 0, 0, 1, 1, 0, 0, 0, 1, 1, 0, 1, 1, 0, 0, 1, 0, 1, 0, 0, 1, 0, 1, 1, 0, 0, 0, 1,

1, 1, 0, 0, 0, 0, 0, 1, 1, 0, 0, 1, 0, 1, 0, 1, 1, 0, 1, 1, 1, 0, 0, 1, 1, 0, 0, 1, 0, 0, 0, 1, 1, 0, 1, 0, 102]])

input_tensor

Tensor

torch.Size([1, 512])

tensor([[13496, 21940, 9374, 29968, 5737, 1096, 19059, 739, 2181, 20334, 14391, 256, 7612,

2252, 5243, 23177, 16180, 28857, 11590, 8664, 1687, 21109, 16135, 8504, 30400, 12843, 19946,

28014, 28621, 12435, 29016, 20791, 24481, 1671, 19430, 9978, 5619, 8404, 11242, 8779, 20244,

25669, 23538, 1441, 4780, 1923, 6294, 14288, 16567, 4410, 25329, 15893, 20521, 5800, 23799,

29996, 28278, 23188, 5739, 14918, 19085, 10330, 11517, 4157, 17572, 19613, 2316, 18970, 2150,

11150, 16324, 19505, 6095, 25328, 16994, 25754, 3185, 841, 4356, 10155, 25172, 28620, 18203,

4026, 26819, 12326, 9459, 7111, 29347, 8106, 26385, 5540, 14244, 4042, 27474, 6491, 2444,

19133, 2592, 21604, 87, 14026, 8377, 19532, 9326, 8873, 2332, 20283, 11478, 10201, 29329,
```

10620, 20667, 7565, 15109, 13667, 10147, 8992, 16916, 23577, 1997, 5997, 6226, 12579, 15504, 4757, 24058, 27025, 15506, 18324, 14684, 15188, 602, 2567, 19353, 20114, 16877, 19303, 27017, 12372, 7893, 9552, 9321, 2698, 5281, 23748, 19119, 18547, 17317, 14861, 18015, 7751, 28276, 12253, 16033, 23079, 2928, 5177, 24328, 5179, 7292, 28162, 7768, 30492, 11631, 26562, 12523, 23400, 19825, 22822, 14343, 3262, 3972, 6212, 28943, 13689, 3828, 5037, 21970, 7397, 27902, 21496, 21384, 2374, 5584, 15474, 8233, 26726, 21765, 30361, 5251, 27108, 27908, 29020, 23167, 1319, 16659, 19174, 7373, 12289, 11078, 9794, 2343, 22851, 7369, 11012, 10719, 26492, 3626, 18491, 3330, 6018, 17268, 21613, 27284, 8361, 25562, 18712, 5604, 3035, 19840, 38, 16481, 5963, 28373, 5733, 9267, 15305, 6929, 5522, 16528, 28703, 5459, 14483, 23563, 3831, 7254, 1653, 1889, 91, 26824, 9906, 4108, 26478, 9915, 223, 22191, 11350, 14489, 9284, 8926, 26186, 29739, 10613, 9751, 1561, 18213, 14970, 6704, 6741, 16651, 790, 29634, 17936, 22040, 25150, 16899, 25363, 20954, 15300, 24063, 25231, 68, 12690, 8953, 21133, 11210, 12293, 10989, 350, 29139, 9207, 16108, 7215, 20790, 7267, 26246, 9672, 3352, 5847, 21896, 1565, 13979, 15948, 8216, 3035, 5055, 13578, 26714, 962, 15882, 1546, 25968, 9080, 2684, 19102, 13328, 7390, 6853, 7601, 21560, 5753, 13825, 17010, 2806, 14875, 16338, 17461, 4766, 24482, 4037, 22588, 4984, 19894, 19717, 4419, 27772, 9342, 30007, 23647, 30498, 5181, 10980, 11835, 18530, 1531, 29259, 25790, 5704, 13909, 3181, 28101, 706, 28191, 7069, 15559, 16433, 8176, 11193, 24414, 449, 2365, 27386, 4747, 27159, 3109, 14439, 12082, 10319, 1349, 16077, 27867, 27641, 24289, 20175, 13435, 29768, 23872, 10544, 30108, 16472, 8527, 6783, 14731, 15752, 1498, 18648, 7890, 3961, 10950, 18181, 8676, 4836, 8052, 7413, 21347, 5635, 14706, 9250, 8752, 26006, 14979, 12716, 12968, 16871, 21707, 13550, 21383, 226, 30218, 4088, 12515, 15859, 22374, 1193, 22407, 22630, 18957, 27449, 13100, 14883, 23420, 15059, 10493, 4133, 14129, 19221, 9724, 23891, 468, 7245, 2844, 9105, 15199, 28018, 23362, 28032, 7098, 23516, 3444, 12259, 7223, 14464, 19272, 26964, 28170, 8202, 3827, 22168, 10671, 13539, 17141, 14828, 29293, 19926, 9429, 13081, 14119, 25050, 17714, 3631, 19169, 21304, 8868, 7248, 26031, 25446, 13155, 18403, 1125, 2076, 28774, 24070, 3776, 15644, 21118, 10061, 19292, 27534, 21001, 20693, 10919, 2640, 11372, 28375,

```
12374, 78, 27377, 6929, 27406, 25440, 22581, 6694, 1035, 805, 15470, 25889, 23365, 18733,
17813, 7392, 18069, 2714, 4495, 3523, 12410, 26684, 30350, 6648, 28889, 11026, 20595, 25083,
4038, 9802, 14035, 15417, 18888, 5436, 27124, 23263, 13081]], device='cuda:0')
labels
list
5 items
[0, 1, 0, 1, Ellipsis]
loss
Tensor
torch.Size([])
tensor(0.0115, device='cuda:0', grad_fn=<NllLossBackward0>)
macs
float

174122385408.0
model
BertForSequenceClassification

BertForSequenceClassification( (bert): BertModel( (embeddings): BertEmbeddings(
(word_embeddings): Embedding(30522, 768, padding_idx=0) (position_embeddings):
Embedding(512, 768) (token_type_embeddings): Embedding(2, 768) (LayerNorm):
LayerNorm((768,), eps=1e-12, elementwise_affine=True) (dropout): Dropout(p=0.1, inplace=False) )
(encoder): BertEncoder( (layer): ModuleList( (0-11): 12 x BertLayer( (attention): BertAttention( (self):
BertSelfAttention( (query): Linear(in_features=768, out_features=768, bias=True) (key):
Linear(in_features=768, out_features=768, bias=True) (value): Linear(in_features=768,
out_features=768, bias=True) (dropout): Dropout(p=0.1, inplace=False) ) (output): BertSelfOutput(
(dense): Linear(in_features=768, out_features=768, bias=True) (LayerNorm): LayerNorm((768,), eps=1e-12, elementwise_affine=True) (dropout): Dropout(p=0.1, inplace=False) ) ) (intermediate): BertIntermediate( (dense): Linear(in_features=768, out_features=3072, bias=True) (intermediate_act_fn): GELUActivation() ) (output): BertOutput( (dense): Linear(in_features=3072, out_features=768, bias=True) (LayerNorm): LayerNorm((768,), eps=1e-12, elementwise_affine=True) (dropout): Dropout(p=0.1, inplace=False) ) ) ) ) (pooler): BertPooler( (dense): Linear(in_features=768, out_features=768, bias=True) (activation): Tanh() ) ) (dropout): Dropout(p=0.1, inplace=False) (classifier): Linear(in_features=768, out_features=2, bias=True) )
```

model_path

str

34 chars

'/content/drive/MyDrive/000/pt/A.pt'

optimizer

AdamW

```
AdamW ( Parameter Group 0 amsgrad: False betas: (0.9, 0.999) capturable: False differentiable: False eps: 1e-08 foreach: None fused: None lr: 1e-05 maximize: False weight_decay: 0.01 )
```

original_model

BertForSequenceClassification

```
BertForSequenceClassification( (bert): BertModel( (embeddings): BertEmbeddings( (word_embeddings): Embedding(30522, 768, padding_idx=0) (position_embeddings): Embedding(512, 768) (token_type_embeddings): Embedding(2, 768) (LayerNorm): LayerNorm((768,), eps=1e-12, elementwise_affine=True) (dropout): Dropout(p=0.1, inplace=False) ) (encoder): BertEncoder( (layer): ModuleList( (0-11): 12 x BertLayer( (attention): BertAttention( (self): BertSelfAttention( (query): Linear(in_features=768, out_features=768, bias=True) (key):
```

```
Linear(in_features=768, out_features=768, bias=True) (value): Linear(in_features=768,
out_features=768, bias=True) (dropout): Dropout(p=0.1, inplace=False) ) (output): BertSelfOutput(
(dense): Linear(in_features=768, out_features=768, bias=True) (LayerNorm): LayerNorm((768,),
eps=1e-12, elementwise_affine=True) (dropout): Dropout(p=0.1, inplace=False) ) ) (intermediate):
BertIntermediate( (dense): Linear(in_features=768, out_features=3072, bias=True)
(intermediate_act_fn): GELUActivation() ) (output): BertOutput( (dense): Linear(in_features=3072,
out_features=768, bias=True) (LayerNorm): LayerNorm((768,), eps=1e-12,
elementwise_affine=True) (dropout): Dropout(p=0.1, inplace=False) ) ) ) ) (pooler): BertPooler(
(dense): Linear(in_features=768, out_features=768, bias=True) (activation): Tanh() ) ) (dropout):
Dropout(p=0.1, inplace=False) (classifier): Linear(in_features=768, out_features=2, bias=True) )
```

outputs

SequenceClassifierOutput

```
SequenceClassifierOutput(loss=tensor(0.0115, device='cuda:0', grad_fn=<NllLossBackward0>),
logits=tensor([[ 2.6082, -1.9123], [ 2.5235, -1.9518], [ 2.9701, -1.9957], [ 2.5651, -1.8357], [ 1.9125,
-1.4096], [ 2.7099, -2.1329], [ 2.5209, -2.0388], [ 2.6560, -1.8715], [ 2.6667, -1.8876], [ 2.5396,
-1.9172], [ 2.5758, -1.8686], [ 2.8509, -2.1623], [ 2.6637, -1.8730], [ 2.7988, -1.7300], [ 2.4641,
-2.1801], [ 2.3826, -2.0551], [ 2.6262, -1.9004], [ 2.3763, -2.1112], [ 2.6413, -1.9029], [ 2.6327,
-1.9391], [ 2.5397, -2.1128], [ 2.5278, -1.6447], [ 2.5663, -2.0353], [ 2.7413, -1.8982]],
device='cuda:0', grad_fn=<AddmmBackward0>), hidden_states=None, attentions=None)
```

params

float

```
85648130.0
```

prof

profile

```
<torch.profiler.profiler.profile object at 0x7de413f328c0>
root
Element

<Element 'ArrayOfFeatureExtractor' at 0x7de406565b20>
saved_state_dict
OrderedDict

OrderedDict([('bert.embeddings.word...ngs.weight', tensor([[-0.0102, -0... 0.0753]])),
('bert.embeddings.posi...ngs.weight', tensor([[ 1.7500e-02...656e-02]])),
('bert.embeddings.toke...ngs.weight', tensor([[ 0.0004, 0...-0.0112]])),
('bert.embeddings.Laye...orm.weight', tensor([0.9261, 0.88..., 0.9371])),
('bert.embeddings.LayerNorm.bias', tensor([-2.5922e-02,...0525e-02])),
('bert.encoder.layer.0...ery.weight', tensor([[-0.0164, 0...-0.0084]])),
('bert.encoder.layer.0...query.bias', tensor([ 5.8489e-01,...6869e-02])),
('bert.encoder.layer.0...key.weight', tensor([[ 0.0081, -0...-0.0017]])),
('bert.encoder.layer.0...f.key.bias', tensor([ 1.0431e-03,...9301e-03])),
('bert.encoder.layer.0...lue.weight', tensor([[ 1.1470e-02...745e-02]])),
('bert.encoder.layer.0...value.bias', tensor([-1.1800e-02,...1531e-03])),
('bert.encoder.layer.0...nse.weight', tensor([[ 0.0058, -0...-0.0059]])),
('bert.encoder.layer.0...dense.bias', tensor([ 5.1317e-03,...8560e-02])),
('bert.encoder.layer.0...orm.weight', tensor([0.9804, 0.95..., 0.9843])), ...])
tokenized_data
BatchEncoding
```