1. Install Required Libraries

```
###!pip install -q comet_ml gradio
```

```
import comet_ml

comet_ml.init(project_name="text-classification-with-transfomers")
```

```
COMET WARNING: comet_ml.init() is deprecated and will be removed soon. Please use comet_ml.login()
```

2. Initialize Comet and HuggingFace

```
from huggingface_hub import notebook_login

notebook_login()
```

Copy a token from your Hugging Face tokens page and paste it below.

Immediately click login after copying your token or it might be stored in plain text in this notebook file.

Token:

Add token as git credential?

Login

Pro Tip: If you don't already have one, you can create a dedicated 'notebooks' token with 'write' access, that you can then easily reuse for all notebooks.

3. Load Data

```
###!mkdir ~/.kaggle


###! pip install kaggle


###!cp /kaggle.json ~/.kaggle/


###!chmod 600 ~/.kaggle/kaggle.json


###!pip install keras-tuner


###!kaggle datasets download -d muhammadehsan000/healthcare-dataset-2019-2024


###! unzip /content/healthcare-dataset-2019-2024.zip
####! pip install datasets


from datasets import load_dataset


raw_datasets = load_dataset("rotten_tomatoes")
```

```
/usr/local/lib/python3.10/dist-packages/huggingface_hub/utils/_token.py:89: UserWarning:
The secret `HF_TOKEN` does not exist in your Colab secrets.
To authenticate with the Hugging Face Hub, create a token in your settings tab (https://huggingface.co/settings/tokens), set it as secre
You will be able to reuse this secret in all of your notebooks.
Please note that authentication is recommended but still optional to access public models or datasets.
  warnings.warn(
```

```
raw_datasets
```

```
DatasetDict({
    train: Dataset({
        features: ['text', 'label'],
        num_rows: 8530
    })
    validation: Dataset({
        features: ['text', 'label'],
        num_rows: 1066
    })
    test: Dataset({
```

```
            features: ['text', 'label'],
            num_rows: 1066
        })
    })
```

```
raw_datasets["train"][0]
```

```
{'text': 'the rock is destined to be the 21st century\'s new " conan " and that he\'s going to make a splash even greater than arnold
schwarzenegger , jean-claud van damme or steven segal .',
 'label': 1}
```

4. Understand Data

```
import pandas as pd

raw_datasets.set_format(type="pandas")
df_train = raw_datasets["train"][:]
df_train.head(3)
```

	text	label
0	the rock is destined to be the 21st century's ...	1
1	the gorgeously elaborate continuation of " the...	1
2	effective but too-tepid biopic	1

Next steps: Generate code with df_train | View recommended plots | New interactive sheet

```
df_train["label"].value_counts()
```

label	count
1	4265
0	4265

dtype: int64

```
df_val = raw_datasets["validation"][:]
df_val.head(3)
```

	text	label
0	compassionately explores the seemingly irrecon...	1
1	the soundtrack alone is worth the price of adm...	1
2	rodriguez does a splendid job of racial profil...	1

Next steps: Generate code with df_val | View recommended plots | New interactive sheet

```
df_val["label"].value_counts()
```

label	count
1	533
0	533

dtype: int64

```
df_test = raw_datasets["test"][:]
df_test.head(3)
```

	text	label
0	lovingly photographed in the manner of a golde...	1
1	consistently clever and suspenseful .	1
2	it's like a " big chill " reunion of the baade...	1

Next steps: Generate code with df_test | View recommended plots | New interactive sheet

```
df_test["label"].value_counts()
```

label	count
1	533
0	533

```
import pandas as pd

raw_datasets.set_format(type="pandas")
df = raw_datasets["train"][:]
df.head()
```

	text	label
0	the rock is destined to be the 21st century's ...	1
1	the gorgeously elaborate continuation of " the...	1
2	effective but too-tepid biopic	1
3	if you sometimes like to go to the movies to h...	1
4	emerges as something rare . an issue movie tha...	1

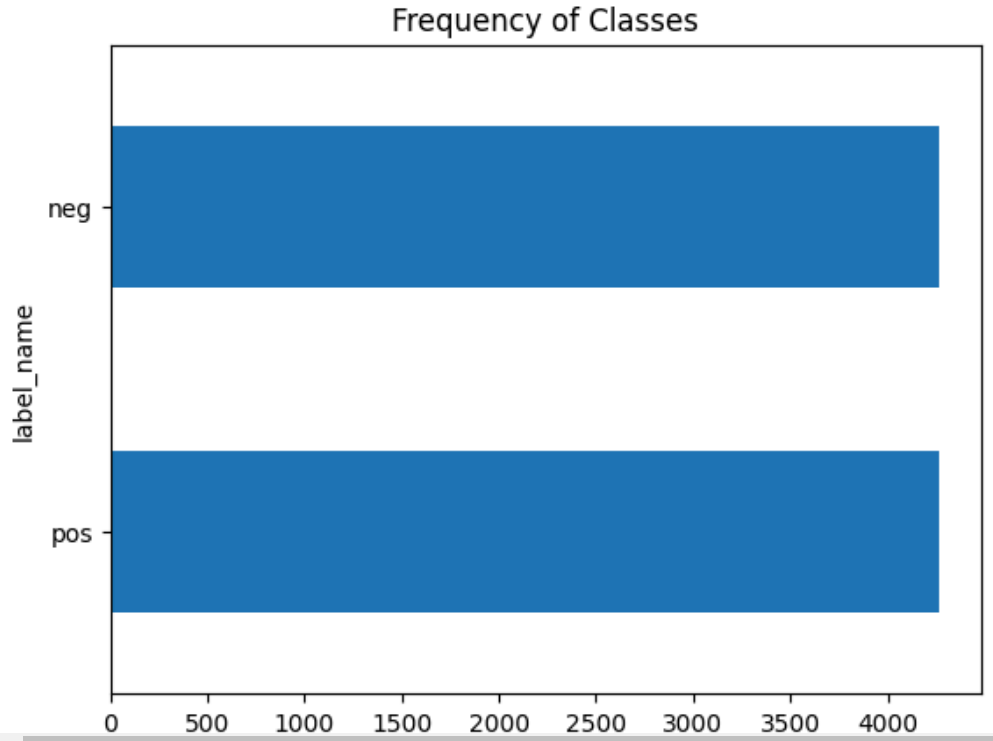
Next steps: Generate code with df | View recommended plots | New interactive sheet

```
def label_int2str(row):
    return raw_datasets["train"].features["label"].int2str(row)

df["label_name"] = df["label"].apply(label_int2str)
df.head()
```

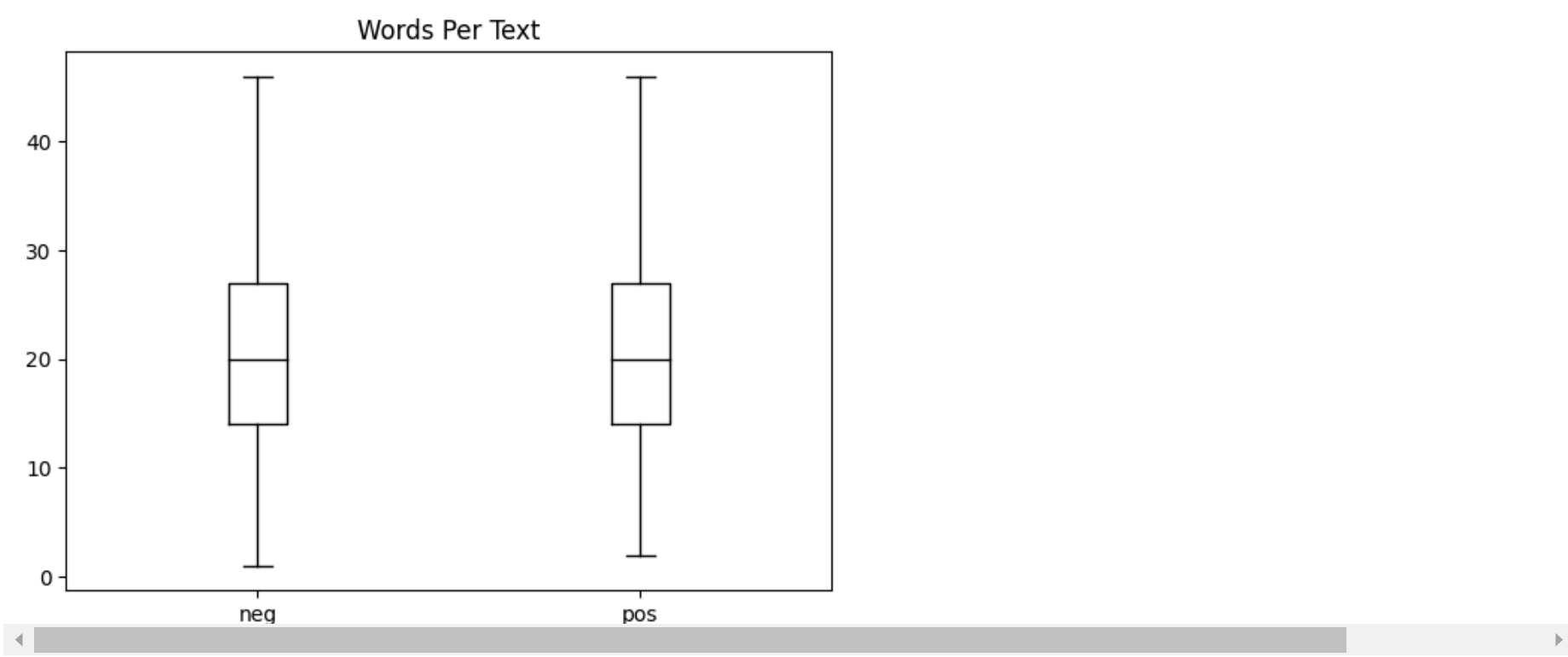
	text	label	label_name
0	the rock is destined to be the 21st century's ...	1	pos
1	the gorgeously elaborate continuation of " the...	1	pos
2	effective but too-tepid biopic	1	pos
3	if you sometimes like to go to the movies to h...	1	pos
4	emerges as something rare , an issue movie tha...	1	pos

Next steps: Generate code with df | View recommended plots | New interactive sheet

```
import matplotlib.pyplot as plt

df["label_name"].value_counts(ascending=True).plot.barh()
plt.title("Frequency of Classes")
plt.show()
```



```
df["Words Per Text"] = df["text"].str.split().apply(len)
df.boxplot("Words Per Text", by="label_name", grid=False, showfliers=False, color="black")
plt.suptitle("")
plt.xlabel("")
plt.show()
```




```
raw_datasets.reset_format()
```

5. Pre-Processing Data

```
from huggingface_hub import login
login(token = "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")
```

```
The token has not been saved to the git credentials helper. Pass `add_to_git_credential=True` in this function directly or `--add-to-git
Token is valid (permission: fineGrained).
Your token has been saved to /root/.cache/huggingface/token
Login successful
```

```
from transformers import AutoTokenizer

checkpoint = "distilbert-base-uncased"
tokenizer = AutoTokenizer.from_pretrained(checkpoint)


def tokenize_function(examples):
    return tokenizer(examples["text"], truncation=True)

tokenized_datasets = raw_datasets.map(tokenize_function, batched=True)
```

```
Map: 100% 1066/1066 [00:00<00:00, 4328.35 examples/s]
```

```
tokenized_datasets
```

```
DatasetDict({
    train: Dataset({
        features: ['text', 'label', 'input_ids', 'attention_mask'],
        num_rows: 8530
    })
    validation: Dataset({
        features: ['text', 'label', 'input_ids', 'attention_mask'],
        num_rows: 1066
    })
    test: Dataset({
        features: ['text', 'label', 'input_ids', 'attention_mask'],
        num_rows: 1066
    })
})
```

```
from transformers import DataCollatorWithPadding

data_collator = DataCollatorWithPadding(tokenizer=tokenizer)
```

6. Setup Evaluation Function

```
from sklearn.metrics import accuracy_score, precision_recall_fscore_support

def get_example(index):
    return tokenized_datasets["test"][index]["text"]


def compute_metrics(pred):
    experiment = comet_ml.get_global_experiment()

    labels = pred.label_ids
    preds = pred.predictions.argmax(-1)
    precision, recall, f1, _ = precision_recall_fscore_support(
        labels, preds, average="macro"
    )
    acc = accuracy_score(labels, preds)

    if experiment:
        epoch = int(experiment.curr_epoch) if experiment.curr_epoch is not None else 0
        experiment.set_epoch(epoch)
        experiment.log_confusion_matrix(
            y_true=labels,
            y_predicted=preds,
            file_name=f"confusion-matrix-epoch-{epoch}.json",
            labels=["negative", "postive"],
            index_to_example_function=get_example,
        )

    return {"accuracy": acc, "f1": f1, "precision": precision, "recall": recall}
```

7. Setup Transformer Model

```
from transformers import AutoModelForSequenceClassification

id2label = {0: "NEGATIVE", 1: "POSITIVE"}
label2id = {"NEGATIVE": 0, "POSITIVE": 1}

model = AutoModelForSequenceClassification.from_pretrained(checkpoint, num_labels=2, id2label=id2label, label2id=label2id)
```

```
Some weights of DistilBertForSequenceClassification were not initialized from the model checkpoint at distilbert-base-uncased and are ne
You should probably TRAIN this model on a down-stream task to be able to use it for predictions and inference.
```

8. Run Training

```
from transformers import TrainingArguments, Trainer

%env COMET_MODE=ONLINE
%env COMET_LOG_ASSETS=TRUE

training_args = TrainingArguments(
    output_dir="my_distilbert_model",
    learning_rate=2e-5,
    per_device_train_batch_size=16,
    per_device_eval_batch_size=16,
    num_train_epochs=3,
    weight_decay=0.01,
    evaluation_strategy="epoch",
    save_strategy="epoch",
    load_best_model_at_end=True,
    push_to_hub=True,
    report_to=["comet_ml"],
)
```

```
env: COMET_MODE=ONLINE
env: COMET_LOG_ASSETS=TRUE
/usr/local/lib/python3.10/dist-packages/transformers/training_args.py:1494: FutureWarning: `evaluation_strategy` is deprecated and will
  warnings.warn(
```

```
trainer = Trainer(
    model=model,
    args=training_args,
    train_dataset=tokenized_datasets["train"],
    eval_dataset=tokenized_datasets["validation"],
    compute_metrics=compute_metrics,
    data_collator=data_collator,
)

trainer.train()
```

```
COMET WARNING: As you are running in a Jupyter environment, you will need to call `experiment.end()` when finished to ensure all metri
COMET INFO: Experiment is live on comet.com https://www.comet.com/debmalyaray9989/text-classification-with-transfomers/6e575ddc7c184b

COMET INFO: Couldn't find a Git repository in '/content' nor in any parent directory. Set `COMET_GIT_DIRECTORY` if your Git Repository
```

[1602/1602 03:54, Epoch 3/3]

Epoch	Training Loss	Validation Loss	Accuracy	F1	Precision	Recall
1	0.411800	0.417395	0.818011	0.815089	0.839467	0.818011
2	0.254900	0.392083	0.849906	0.849803	0.850874	0.849906
3	0.161000	0.484692	0.846154	0.846127	0.846393	0.846154

```
COMET INFO: ---------------------------------------------------------------------------------------
COMET INFO: Comet.ml Experiment Summary
COMET INFO: ---------------------------------------------------------------------------------------
COMET INFO:   Data:
COMET INFO:     display_summary_level : 1
COMET INFO:     name                 : considerable_jigsaw_8658
COMET INFO:     url                  : https://www.comet.com/debmalyaray9989/text-classification-with-transfomers/6e575ddc7c184bca8b
COMET INFO:   Metrics [count] (min, max):
COMET INFO:     epoch [7]                   : (0.9363295880149812, 3.0)
COMET INFO:     eval_accuracy [3]           : (0.8180112570356473, 0.849906191369606)
COMET INFO:     eval_f1 [3]                 : (0.8150894134477824, 0.8498025664598985)
COMET INFO:     eval_loss [3]               : (0.3920828104019165, 0.4846917390823364)
COMET INFO:     eval_precision [3]          : (0.8394674843029613, 0.8508744992146273)
COMET INFO:     eval_recall [3]             : (0.8180112570356473, 0.849906191369606)
COMET INFO:     eval_runtime [3]            : (1.6095, 1.7682)
COMET INFO:     eval_samples_per_second [3] : (602.874, 662.312)
COMET INFO:     eval_steps_per_second [3]   : (37.892, 41.628)
COMET INFO:     grad_norm [3]               : (3.826653003692627, 25.425678253173828)
COMET INFO:     learning_rate [3]           : (1.2734082397003748e-06, 1.3757802746566792e-05)
COMET INFO:     loss [163]                  : (0.009388956241309643, 0.740016520023346)
COMET INFO:     total_flos                  : 323059008078528.0
COMET INFO:     train_loss                  : 0.2670270887653479
COMET INFO:     train_runtime               : 175.6053
COMET INFO:     train_samples_per_second    : 145.725
COMET INFO:     train_steps_per_second      : 9.123
COMET INFO:   Parameters:
COMET INFO:     args/_n_gpu                                 : 1
COMET INFO:     args/_no_sync_in_gradient_accumulation      : True
COMET INFO:     args/_setup_devices                         : cuda:0
COMET INFO:     args/accelerator_config                     : AcceleratorConfig(split_batches=False, dispatch_batches=None, even_bat
COMET INFO:     args/adafactor                              : False
COMET INFO:     args/adam_beta1                             : 0.9
COMET INFO:     args/adam_beta2                             : 0.999
COMET INFO:     args/adam_epsilon                           : 1e-08
COMET INFO:     args/auto_find_batch_size                   : False
COMET INFO:     args/batch_eval_metrics                     : False
```

```
COMET INFO:     args/batch_eval_metrics                 : False
COMET INFO:     args/bf16                               : False
COMET INFO:     args/bf16_full_eval                     : False
COMET INFO:     args/data_seed                          : None
COMET INFO:     args/dataloader_drop_last               : False
COMET INFO:     args/dataloader_num_workers             : 0
COMET INFO:     args/dataloader_persistent_workers      : False
COMET INFO:     args/dataloader_pin_memory              : True
COMET INFO:     args/dataloader_prefetch_factor         : None
COMET INFO:     args/ddp_backend                        : None
COMET INFO:     args/ddp_broadcast_buffers              : None
COMET INFO:     args/ddp_bucket_cap_mb                  : None
COMET INFO:     args/ddp_find_unused_parameters         : None
COMET INFO:     args/ddp_timeout                        : 1800
COMET INFO:     args/ddp_timeout_delta                  : 0:30:00
COMET INFO:     args/debug                              : []
COMET INFO:     args/deepspeed                          : None
COMET INFO:     args/deepspeed_plugin                   : None
COMET INFO:     args/default_optim                      : adamw_torch
COMET INFO:     args/device                             : cuda:0
COMET INFO:     args/disable_tqdm                       : False
COMET INFO:     args/dispatch_batches                   : None
COMET INFO:     args/distributed_state                  : Distributed environment: NO
Num processes: 1
Process index: 0
Local process index: 0
Device: cuda

COMET INFO:     args/do_eval                            : True
COMET INFO:     args/do_predict                         : False
COMET INFO:     args/do_train                           : False
COMET INFO:     args/eval_accumulation_steps            : None
COMET INFO:     args/eval_batch_size                    : 16
COMET INFO:     args/eval_delay                         : 0
COMET INFO:     args/eval_do_concat_batches             : True
COMET INFO:     args/eval_on_start                      : False
COMET INFO:     args/eval_steps                         : None
COMET INFO:     args/eval_strategy                      : IntervalStrategy.EPOCH
COMET INFO:     args/evaluation_strategy                : epoch
COMET INFO:     args/fp16                               : False
COMET INFO:     args/fp16_backend                       : auto
COMET INFO:     args/fp16_full_eval                     : False
COMET INFO:     args/fp16_opt_level                     : O1
COMET INFO:     args/framework                          : pt
COMET INFO:     args/fsdp                               : []
COMET INFO:     args/fsdp_config                        : {'min_num_params': 0, 'xla': False, 'xla_fsdp_v2': False, 'xla_fsdp_gra
COMET INFO:     args/fsdp_min_num_params                : 0
COMET INFO:     args/fsdp_transformer_layer_cls_to_wrap : None
COMET INFO:     args/full_determinism                   : False
```

```
COMET INFO:     args/gradient_accumulation_steps            : 1
COMET INFO:     args/gradient_checkpointing                 : False
COMET INFO:     args/gradient_checkpointing_kwargs          : None
COMET INFO:     args/greater_is_better                      : False
COMET INFO:     args/group_by_length                        : False
COMET INFO:     args/half_precision_backend                 : auto
COMET INFO:     args/hub_always_push                        : False
COMET INFO:     args/hub_model_id                           : None
COMET INFO:     args/hub_private_repo                       : False
COMET INFO:     args/hub_strategy                           : HubStrategy.EVERY_SAVE
COMET INFO:     args/hub_token                              : None
COMET INFO:     args/ignore_data_skip                       : False
COMET INFO:     args/include_inputs_for_metrics             : False
COMET INFO:     args/include_num_input_tokens_seen          : False
COMET INFO:     args/include_tokens_per_second              : False
COMET INFO:     args/jit_mode_eval                          : False
COMET INFO:     args/label_names                            : None
COMET INFO:     args/label_smoothing_factor                 : 0.0
COMET INFO:     args/learning_rate                          : 2e-05
COMET INFO:     args/length_column_name                     : length
COMET INFO:     args/load_best_model_at_end                 : True
COMET INFO:     args/local_process_index                    : 0
COMET INFO:     args/local_rank                             : 0
COMET INFO:     args/log_level                              : passive
COMET INFO:     args/log_level_replica                      : warning
COMET INFO:     args/log_on_each_node                       : True
COMET INFO:     args/logging_dir                            : my_distilbert_model/runs/Aug17_04-46-12_f5656e35329a
COMET INFO:     args/logging_first_step                     : False
COMET INFO:     args/logging_nan_inf_filter                 : True
COMET INFO:     args/logging_steps                          : 500
COMET INFO:     args/logging_strategy                       : IntervalStrategy.STEPS
COMET INFO:     args/lr_scheduler_kwargs                    : {}
COMET INFO:     args/lr_scheduler_type                      : SchedulerType.LINEAR
COMET INFO:     args/max_grad_norm                          : 1.0
COMET INFO:     args/max_steps                              : -1
COMET INFO:     args/metric_for_best_model                  : loss
COMET INFO:     args/mp_parameters                          :
COMET INFO:     args/n_gpu                                  : 1
COMET INFO:     args/neftune_noise_alpha                    : None
COMET INFO:     args/no_cuda                                : False
COMET INFO:     args/num_train_epochs                       : 3
COMET INFO:     args/optim                                  : OptimizerNames.ADAMW_TORCH
COMET INFO:     args/optim_args                             : None
COMET INFO:     args/optim_target_modules                   : None
COMET INFO:     args/output_dir                             : my_distilbert_model
COMET INFO:     args/overwrite_output_dir                   : False
COMET INFO:     args/parallel_mode                          : ParallelMode.NOT_PARALLEL
COMET INFO:     args/past_index                             : -1
COMET INFO:     args/per_device_eval_batch_size             : 16
```

```
COMET INFO:     args/per_device_train_batch_size             : 16
COMET INFO:     args/per_gpu_eval_batch_size                 : None
COMET INFO:     args/per_gpu_train_batch_size                : None
COMET INFO:     args/place_model_on_device                   : True
COMET INFO:     args/prediction_loss_only                    : False
COMET INFO:     args/process_index                           : 0
COMET INFO:     args/push_to_hub                             : True
COMET INFO:     args/push_to_hub_model_id                    : None
COMET INFO:     args/push_to_hub_organization                : None
COMET INFO:     args/push_to_hub_token                       : None
COMET INFO:     args/ray_scope                               : last
COMET INFO:     args/remove_unused_columns                   : True
COMET INFO:     args/report_to                               : ['comet_ml']
COMET INFO:     args/restore_callback_states_from_checkpoint : False
COMET INFO:     args/resume_from_checkpoint                  : None
COMET INFO:     args/run_name                                : my_distilbert_model
COMET INFO:     args/save_on_each_node                       : False
COMET INFO:     args/save_only_model                         : False
COMET INFO:     args/save_safetensors                        : True
COMET INFO:     args/save_steps                              : 500
COMET INFO:     args/save_strategy                           : IntervalStrategy.EPOCH
COMET INFO:     args/save_total_limit                        : None
COMET INFO:     args/seed                                    : 42
COMET INFO:     args/should_log                              : True
COMET INFO:     args/should_save                             : True
COMET INFO:     args/skip_memory_metrics                     : True
COMET INFO:     args/split_batches                           : None
COMET INFO:     args/tf32                                    : None
COMET INFO:     args/torch_compile                           : False
COMET INFO:     args/torch_compile_backend                   : None
COMET INFO:     args/torch_compile_mode                      : None
COMET INFO:     args/torchdynamo                             : None
COMET INFO:     args/tpu_metrics_debug                       : False
COMET INFO:     args/tpu_num_cores                           : None
COMET INFO:     args/train_batch_size                        : 16
COMET INFO:     args/use_cpu                                 : False
COMET INFO:     args/use_ipex                                : False
COMET INFO:     args/use_legacy_prediction_loop              : False
COMET INFO:     args/use_mps_device                          : False
COMET INFO:     args/warmup_ratio                            : 0.0
COMET INFO:     args/warmup_steps                            : 0
COMET INFO:     args/weight_decay                            : 0.01
COMET INFO:     args/world_size                              : 1
COMET INFO:     config/_attn_implementation                  : eager
COMET INFO:     config/_attn_implementation_internal         : None
COMET INFO:     config/_auto_class                           : None
COMET INFO:     config/_commit_hash                          : 12040accade4e8a0f71eabdb258fecc2e7e948be
COMET INFO:     config/_name_or_path                         : distilbert-base-uncased
COMET INFO:     config/activation                            : gelu
```

```
COMET INFO:     config/activation                          : gelu
COMET INFO:     config/add_cross_attention                 : False
COMET INFO:     config/architectures                       : ['DistilBertForMaskedLM']
COMET INFO:     config/attention_dropout                   : 0.1
COMET INFO:     config/attribute_map                       : {'hidden_size': 'dim', 'num_attention_heads': 'n_heads', 'num_hidden_la
COMET INFO:     config/bad_words_ids                       : None
COMET INFO:     config/begin_suppress_tokens               : None
COMET INFO:     config/bos_token_id                        : None
COMET INFO:     config/chunk_size_feed_forward             : 0
COMET INFO:     config/cross_attention_hidden_size         : None
COMET INFO:     config/decoder_start_token_id              : None
COMET INFO:     config/dim                                 : 768
COMET INFO:     config/diversity_penalty                   : 0.0
COMET INFO:     config/do_sample                           : False
COMET INFO:     config/dropout                             : 0.1
COMET INFO:     config/early_stopping                      : False
COMET INFO:     config/encoder_no_repeat_ngram_size        : 0
COMET INFO:     config/eos_token_id                        : None
COMET INFO:     config/exponential_decay_length_penalty    : None
COMET INFO:     config/finetuning_task                     : None
COMET INFO:     config/forced_bos_token_id                 : None
COMET INFO:     config/forced_eos_token_id                 : None
COMET INFO:     config/hidden_dim                          : 3072
COMET INFO:     config/id2label                            : {0: 'NEGATIVE', 1: 'POSITIVE'}
COMET INFO:     config/initializer_range                   : 0.02
COMET INFO:     config/is_composition                      : False
COMET INFO:     config/is_decoder                          : False
COMET INFO:     config/is_encoder_decoder                  : False
COMET INFO:     config/label2id                            : {'NEGATIVE': 0, 'POSITIVE': 1}
COMET INFO:     config/length_penalty                      : 1.0
COMET INFO:     config/max_length                          : 20
COMET INFO:     config/max_position_embeddings             : 512
COMET INFO:     config/min_length                          : 0
COMET INFO:     config/model_type                          : distilbert
COMET INFO:     config/n_heads                             : 12
COMET INFO:     config/n_layers                            : 6
COMET INFO:     config/name_or_path                        : distilbert-base-uncased
COMET INFO:     config/no_repeat_ngram_size                : 0
COMET INFO:     config/num_beam_groups                     : 1
COMET INFO:     config/num_beams                           : 1
COMET INFO:     config/num_labels                          : 2
COMET INFO:     config/num_return_sequences                : 1
COMET INFO:     config/output_attentions                   : False
COMET INFO:     config/output_hidden_states                : False
COMET INFO:     config/output_scores                       : False
COMET INFO:     config/pad_token_id                        : 0
COMET INFO:     config/prefix                              : None
COMET INFO:     config/problem_type                        : None
COMET INFO:     config/pruned_heads                        : {}
```

```
COMET INFO:     config/qa_dropout                          : 0.1
COMET INFO:     config/remove_invalid_values               : False
COMET INFO:     config/repetition_penalty                  : 1.0
COMET INFO:     config/return_dict                         : True
COMET INFO:     config/return_dict_in_generate             : False
COMET INFO:     config/sep_token_id                        : None
COMET INFO:     config/seq_classif_dropout                 : 0.2
COMET INFO:     config/sinusoidal_pos_embds                : False
COMET INFO:     config/suppress_tokens                     : None
COMET INFO:     config/task_specific_params                : None
COMET INFO:     config/temperature                         : 1.0
COMET INFO:     config/tf_legacy_loss                      : False
COMET INFO:     config/tie_encoder_decoder                 : False
COMET INFO:     config/tie_weights_                        : True
COMET INFO:     config/tie_word_embeddings                 : True
COMET INFO:     config/tokenizer_class                     : None
COMET INFO:     config/top_k                               : 50
COMET INFO:     config/top_p                               : 1.0
COMET INFO:     config/torch_dtype                         : None
COMET INFO:     config/torchscript                         : False
COMET INFO:     config/transformers_version                : 4.10.0.dev0
COMET INFO:     config/typical_p                           : 1.0
COMET INFO:     config/use_bfloat16                        : False
COMET INFO:     config/use_return_dict                     : True
COMET INFO:     config/vocab_size                          : 30522
COMET INFO:   Uploads:
COMET INFO:     asset               : 25 (2.49 GB)
COMET INFO:     confusion-matrix    : 3
COMET INFO:     environment details : 1
COMET INFO:     filename            : 1
COMET INFO:     installed packages  : 1
COMET INFO:     model graph         : 1
COMET INFO:     notebook            : 2
COMET INFO:     os packages         : 1
COMET INFO:     source_code         : 1
COMET INFO:
COMET INFO: Please wait for metadata to finish uploading (timeout is 3600 seconds)
COMET INFO: Uploading 263 metrics, params and output messages
COMET INFO: Please wait for assets to finish uploading (timeout is 10800 seconds)
COMET INFO: Still uploading 13 file(s), remaining 1.29 GB/2.00 GB
COMET INFO: Still uploading 3 asset(s), remaining 472.24 MB/1.25 GB, Throughput 56.36 MB/s, ETA ~9s
COMET INFO: Still uploading 1 asset(s), remaining 10.91 MB/510.91 MB, Throughput 30.72 MB/s, ETA ~1s
TrainOutput(global_step=1602, training_loss=0.2670270887653479, metrics={'train_runtime': 175.6053, 'train_samples_per_second':
145.725, 'train_steps_per_second': 9.123, 'total_flos': 323059008078528.0, 'train_loss': 0.2670270887653479, 'epoch': 3.0})
```

```
trainer.push_to_hub()
```

```
CommitInfo(commit_url='https://huggingface.co/Debmalya/my_distilbert_model/commit/b0b46c8b1de346876382f1511618b4b376372ea0',
```

9. Inference

```
text = "This is a great movie. It may be my favourite."


from transformers import pipeline

classifier = pipeline("sentiment-analysis", model="Tirendaz/my_distilbert_model")
classifier(text)
```

```
Hardware accelerator e.g. GPU is available in the environment, but no `device` argument is passed to the `Pipeline` object. Model will b
[{'label': 'POSITIVE', 'score': 0.976166307926178}]
```

10. Deploy

```
import gradio as gr
from transformers import pipeline

classifier = pipeline("sentiment-analysis", model="Tirendaz/my_distilbert_model")

def text_classification(text):
    result= classifier(text)
    sentiment_label = result[0]['label']
    sentiment_score = result[0]['score']
    formatted_output = f"This sentiment is {sentiment_label} with the probability {sentiment_score*100:.2f}%"
    return formatted_output

examples=["This is wonderful movie!", "The movie was really bad; I didn't like it."]

io = gr.Interface(fn=text_classification,
                          inputs= gr.Textbox(lines=2, label="Text", placeholder="Enter title here..."),
                          outputs=gr.Textbox(lines=2, label="Text Classification Result"),
                          title="Text Classification",
                          description="Enter a text and see the text classification result!",
                          examples=examples)

io.launch(inline=False, share=True)


experiment = comet_ml.Experiment()
experiment.add_tag("text-classifier")

io.integrate(comet_ml=experiment)
```

```
Hardware accelerator e.g. GPU is available in the environment, but no `device` argument is passed to the `Pipeline` object. Model will b
Colab notebook detected. To show errors in colab notebook, set debug=True in launch()
COMET WARNING: As you are running in a Jupyter environment, you will need to call `experiment.end()` when finished to ensure all metrics
Running on public URL: https://9ab920337e2d9fbfae.gradio.live

This share link expires in 72 hours. For free permanent hosting and GPU upgrades, run `gradio deploy` from Terminal to deploy to Spaces
COMET INFO: Experiment is live on comet.com https://www.comet.com/debmalyaray9989/text-classification-with-transfomers/323046bac114430a9

COMET INFO: Couldn't find a Git repository in '/content' nor in any parent directory. Set `COMET_GIT_DIRECTORY` if your Git Repository i
COMET INFO: ---------------------------------------------------------------------------------------
COMET INFO: Comet.ml Experiment Summary
COMET INFO: ---------------------------------------------------------------------------------------
COMET INFO:   Data:
COMET INFO:     display_summary_level : 1
```

```
COMET INFO:     name                    : determined flitch 6133
```

```
###! pip install joblib
```

```
COMET INFO:     Created from : Gradio
```

```
import joblib
```

```
COMET INFO:     filename                : 1
```

```
joblib.dump(classifier, 'classifier_model.joblib')
```

```
['classifier_model.joblib']
```

```
COMET INFO:     text_sample             : 1
```

```
from comet_ml import Experiment
exp = Experiment()
exp.log_model("model_name", "/content/text_classifier_model.joblib")
```

```
COMET WARNING: As you are running in a Jupyter environment, you will need to call `experiment.end()` when finished to ensure all metrics
COMET INFO: Experiment is live on comet.com https://www.comet.com/debmalyaray9989/text-classification-with-transfomers/932c4ace1dc04b128

COMET INFO: Couldn't find a Git repository in '/content' nor in any parent directory. Set `COMET_GIT_DIRECTORY` if your Git Repository i
{'web': 'https://www.comet.com/api/asset/download?
assetId=ac11ba66954b44bb83f0a7ebc2b428d3&experimentKey=932c4ace1dc04b12854e432852391f2d',
 'api': 'https://www.comet.com/api/rest/v2/experiment/asset/get-asset?
```