

```
In [ ]: from simpletransformers.classification import ClassificationModel
import pandas as pd
import logging
import sklearn
from sklearn.model_selection import train_test_split
from sklearn.preprocessing import LabelEncoder

In [ ]: df = pd.read_csv("sentiment_analysis.csv")

In [ ]: # Drop unnecessary columns
df.drop(columns=['Year', 'Month', 'Day', 'Time of Tweet', 'Platform'], inplace=True)

In [ ]: df = df.dropna()

In [ ]: df['sentiment'].unique()

Out[ ]: array(['positive', 'negative', 'neutral'], dtype=object)

In [ ]: df = df[df['sentiment'] != 'neutral']

In [ ]: # Rename the columns to match expected names
df = df.rename(columns={"text": "text", "sentiment": "labels"})

In [ ]: # Initialize LabelEncoder and encode the sentiments
label_encoder = LabelEncoder()
df['labels'] = label_encoder.fit_transform(df['labels'])

In [ ]: df['labels'].unique()

Out[ ]: array([1, 0])

In [ ]: # Split the data into training and testing sets (80% train, 20% test)
train_df, test_df = train_test_split(df, test_size=0.2, random_state=42)

In [ ]: model = ClassificationModel(
    'distilbert', 'distilbert-base-uncased',
    use_cuda=False,
    args={'reprocess_input_data': True, 'overwrite_output_dir': True}
)
model.train_model(train_df)

Some weights of DistilBertForSequenceClassification were not initialized from the model checkpoint at distilbert-base-uncased and are newly initialized: ['classifier.bias', 'classifier.weight', 'pre_classifier.bias', 'pre_classifier.weight']
You should probably TRAIN this model on a down-stream task to be able to use it for predictions and inference.
1it [00:19, 19.03s/it]
Epochs 1/1. Running Loss: 0.5189: 100%[██████████] 30/30 [00:43<00:00, 1.43s/it]
Epoch 1 of 1: 100%[██████████] 1/1 [00:43<00:00, 43.92s/it]

Out[ ]: (30, 0.640367333094279)

In [ ]: #model = ClassificationModel('roberta', 'roberta-base', use_cuda=False, args={'repr

# Train the model
#model.train_model(train_df)

In [ ]: test_df
```

Out[]:

		text	labels
351		Sorry RB is on PS3 for me	0
454		Still jealous	0
273		Went to a party last night. Dindin and I showe...	1
13		i want to go to music tonight but i lost my vo...	0
406		Have a good one	1
390		sorry friends - I`m swamped with deadlines r...	0
342		Ok so I`ve now got a bit of a bad back after l...	0
187		I am tres depressed	0
6		2 am feedings for the baby are fun when he is ...	1
309		I know It was worth a shot, though!	1
412		_x_ATL u mean jack barakat`s?! wow so have u e...	1
91		Its amazing game, I playing everyday	1
376		aww I miss driving down elmwood	0
70		Those who want to go, let them go	0
320		Sorry to hear ur flight got cancelled that b...	0
379		_30439 I really wish I could go!	1
483		I`m missing crab legs and attending my going a...	0
363		aww you loooove me	1
264		Well good morning all, What a wonderful day in...	1
288		You know your neck is jacked up when you are f...	0
131		failed inspection.	0
191		That`s awesome dude, yay for surprise celebri...	1
428		Yay, three followers! Good to know more than ...	1
429		'my name is Tony!!!!!! ...not hey!!!!' - poor...	0
181		thank you for teaching me values and to be a b...	1
67		I feel sorry, I miss you here in the sea beach	1
472		250 miles down, only 1750 to go. Thats ok, u2 ...	1
492		Sorry, we`ll try to keep it down.	0
278		Thanks	1
413		I sure do hope it becomes 4:20 this afternoon ...	1
26		I`m sorry.	0
286		it makes me happy to hear a girl talk, or twe...	1
55		I AM SUCH A CREEPER I feel disappointed becau...	0
36		Born and raised in NYC and living in Texas for...	0
371		I feel your pain. Mine is the same way	0
202		what`s goin on hun? I`m worried about you	0
11		I really really like the song Cardigan by Tayl...	1
154		I'm studying in psychology	0
72		2 am feedings for the baby are fun when he is ...	1
119		Lets go, its time to play cricket match	1

	text	labels
160	your attitude looks like autistics child	0
124	feeling gloomy	0
480	Bumping dj opus in the drunk in the car. lmao...	0
96	How looks like our company new logo?	1
129	Walking to class. I hate not having a bike.....	0
99	That`s very funny. Cute kids.	1
407	yep infact she is popular, miss india 99, tal...	1
397	I hate the dentist	0
189	Happy Mothers Day people. i love my mom a lot ...	1
399	I`m taking a twitter break. Cell is dying	0
317	Laying in bed til workkk... Oh the life. Defin...	0
254	Trying to decide on a movie with the friends.....	0
414	it means that you are now FAMOUS. Congrats!	1
123	Feeling smooth like chrome	1
495	the plan to not spend money is not going well	0
469	That`s just weird... :\ Oh, and what was it ...	0
167	writing report cards soooo tired but what an ...	1
158	almost died. Laptop screen was set to 100% bri...	1
337	My frist post... Off to find a new car for my ...	1
38	i`m soooooo sleeeeeeepy!!! the last day o` schoo...	0

```
In [ ]: # Make predictions on the test set (pass the 'text' column as a list of strings)
        predictions, raw_outputs = model.predict(test_df['text'].tolist())

        # Print the predictions (you can modify this as needed)
        print(predictions)
```

```
1it [00:18, 18.97s/it]
100%|██████████| 1/1 [00:02<00:00, 2.64s/it]
[1 0 1 0 1 0 1 0 1 1 1 1 0 1 1 1 0 0 1 1 1 1 1 0 1 1 0 1 1 0 1
 1 1 1 1 0 0 1 0 1 1 0 1 0 1 1 1 1 0 1 1 1 1 0]
```

```
In [ ]: import matplotlib.pyplot as plt
        import seaborn as sns
        from sklearn.metrics import confusion_matrix

        # Ensure the predictions and actual labels are in the DataFrame
        test_df['predictions'] = predictions

        # Plot the distribution of actual and predicted sentiments
        plt.figure(figsize=(10, 5))

        # Plot actual sentiment distribution
        plt.subplot(1, 2, 1)
        sns.countplot(x='labels', data=test_df, palette="viridis")
        plt.title("Actual Sentiment Distribution")
        plt.xlabel("Sentiment")
        plt.ylabel("Count")

        # Plot predicted sentiment distribution
        plt.subplot(1, 2, 2)
        sns.countplot(x='predictions', data=test_df, palette="viridis")
        plt.title("Predicted Sentiment Distribution")
        plt.xlabel("Sentiment")
        plt.ylabel("Count")
```

```
plt.tight_layout()  
plt.show()
```

```
C:\Users\Asus\AppData\Local\Temp\ipykernel_34528\1706612813.py:13: FutureWarning:  
  
Passing `palette` without assigning `hue` is deprecated and will be removed in v0.1  
4.0. Assign the `x` variable to `hue` and set `legend=False` for the same effect.  
  
sns.countplot(x='labels', data=test_df, palette="viridis")  
C:\Users\Asus\AppData\Local\Temp\ipykernel_34528\1706612813.py:20: FutureWarning:  
  
Passing `palette` without assigning `hue` is deprecated and will be removed in v0.1  
4.0. Assign the `x` variable to `hue` and set `legend=False` for the same effect.  
  
sns.countplot(x='predictions', data=test_df, palette="viridis")
```

