Source Code :

from transformers import DistilBertTokenizerFast, TFDistilBertForSequenceClassification

from sklearn.model\_selection import train\_test\_split

import tensorflow as tf

tokenizer = DistilBertTokenizerFast.from\_pretrained('distilbert-base-uncased')

X = list(df['text'])

y = list(df['label'])

X\_train, X\_test, y\_train, y\_test = train\_test\_split(X, y, test\_size=0.2, random\_state=42)

train\_encodings = tokenizer(X\_train, truncation=True, padding=True, max\_length=512)

test\_encodings = tokenizer(X\_test, truncation=True, padding=True, max\_length=512)

train\_dataset = tf.data.Dataset.from\_tensor\_slices((

dict(train\_encodings),

y\_train

))

model = TFDistilBertForSequenceClassification.from\_pretrained('distilbert-base-uncased', num\_labels=2)

model.compile(optimizer=tf.keras.optimizers.Adam(), loss=model.compute\_loss, metrics=['accuracy'])

model.fit(train\_dataset.shuffle(100).batch(16), epochs=3)