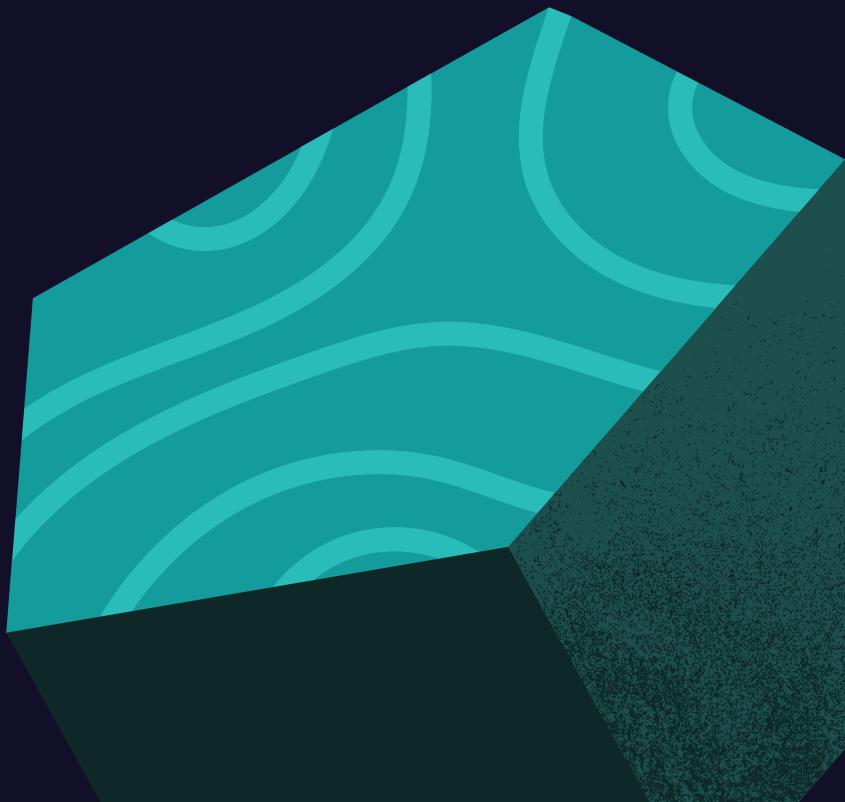


NLP: Tweets Sentiment Analysis

Cem EKE

SUMMARY



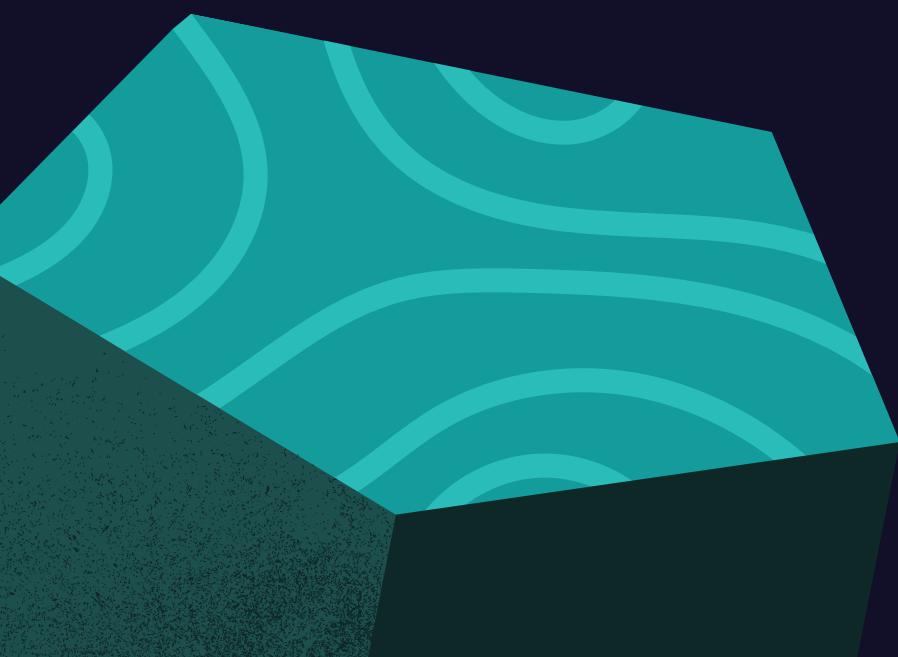
Data Pre-Processing

Data Pre-Processing

Tokenizer

Padding and
Truncating

Label Preparation

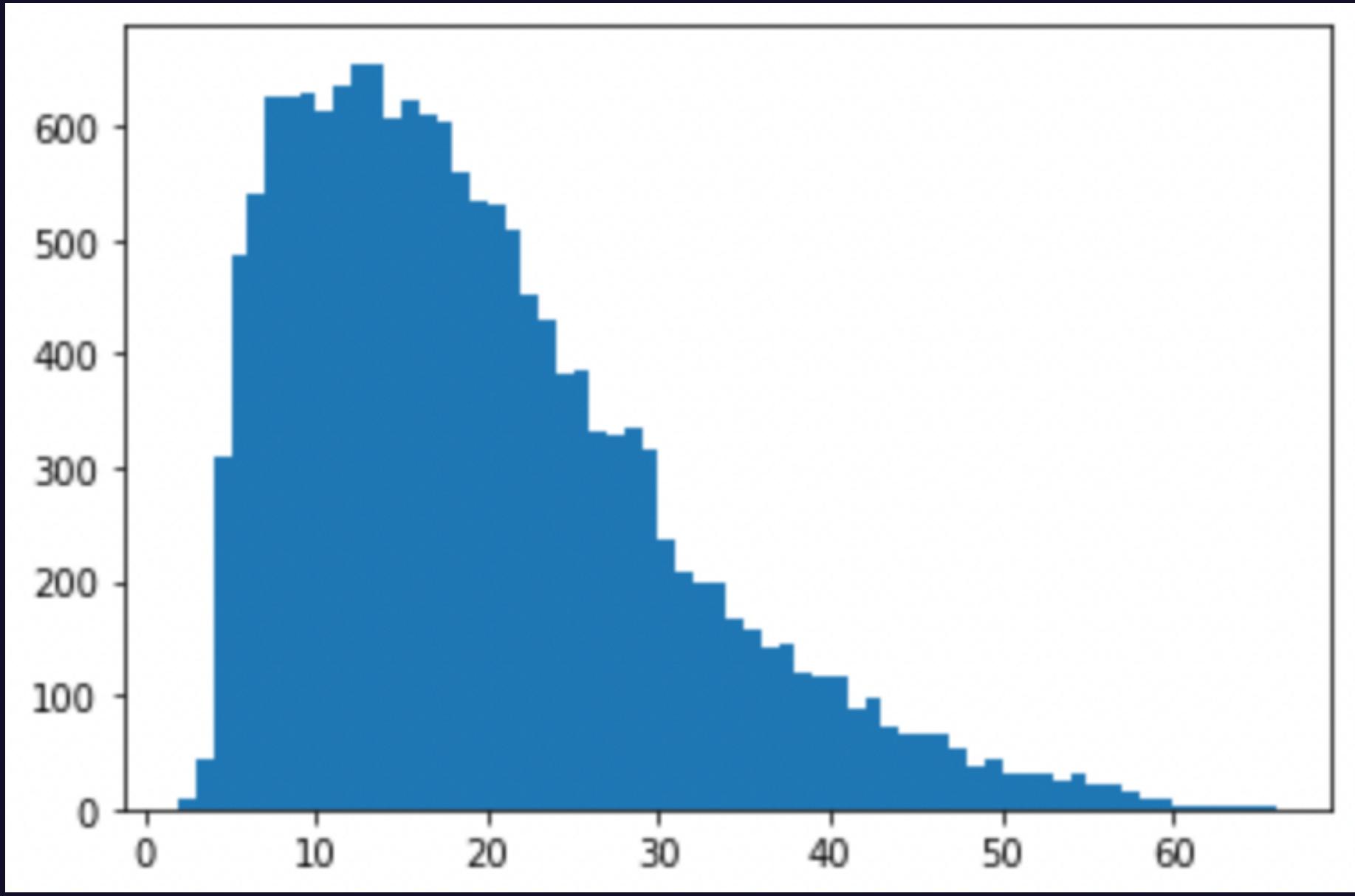


```
12] tokenizer.texts_to_sequences([tweets[1]])
```

```
[ [2,  
  40,  
 101,  
 60,  
 8,  
15,  
494,  
 5,  
15,  
3496,  
553,  
32,  
60,  
61,  
128,  
148,  
76,  
1480,  
 4,  
22,  
1255]]
```

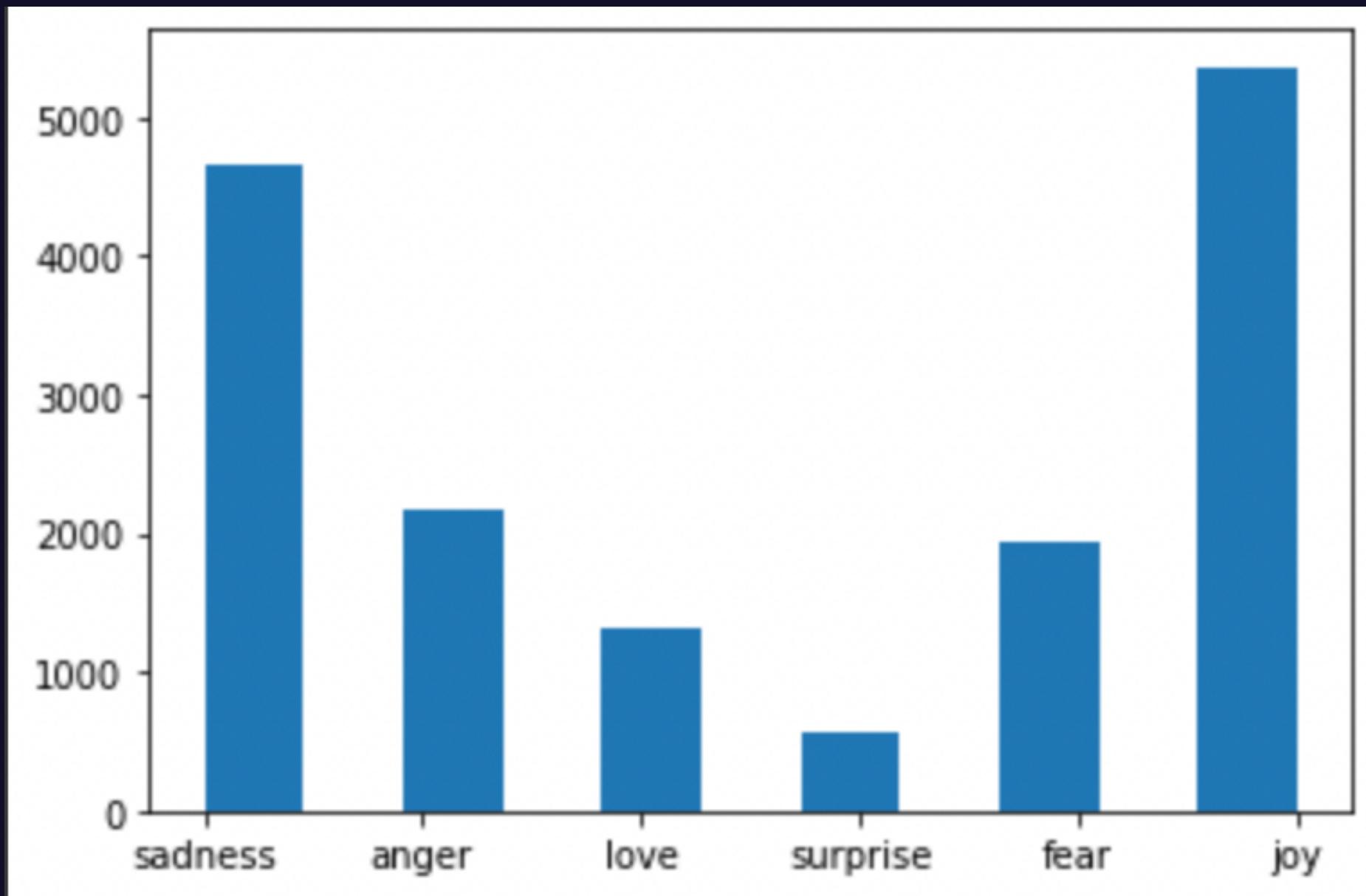
```
13] tweets[1]
```

```
'i can go from feeling so hopeless to so damned hopeful just from being around someone who cares and is awake'
```



```
[22] padded_train_sequences[1]
```

```
array([ 2,  40, 101,  60,    8,   15, 494,    5,   15, 3496, 553,
       32,  60,   61, 128, 148,   76, 1480,    4,   22, 1255,    0,
       0,    0,    0,    0,    0,    0,    0,    0,    0,    0,    0,
       0,    0,    0,    0,    0,    0,    0,    0,    0,    0,    0,
       0,    0,    0,    0,    0], dtype=int32)
```



Model Building

```
model = tf.keras.models.Sequential([
    tf.keras.layers.Embedding(10000, 16, input_length = maxlen),
    tf.keras.layers.Bidirectional(tf.keras.layers.LSTM(20, return_sequences = True)),
    tf.keras.layers.Bidirectional(tf.keras.layers.LSTM(20)),
    tf.keras.layers.Dense(6, activation = 'softmax')
])

model.compile(
    loss = 'sparse_categorical_crossentropy',
    optimizer = 'adam',
    metrics = ['accuracy']
)

model.summary()
```

Model: "sequential_3"

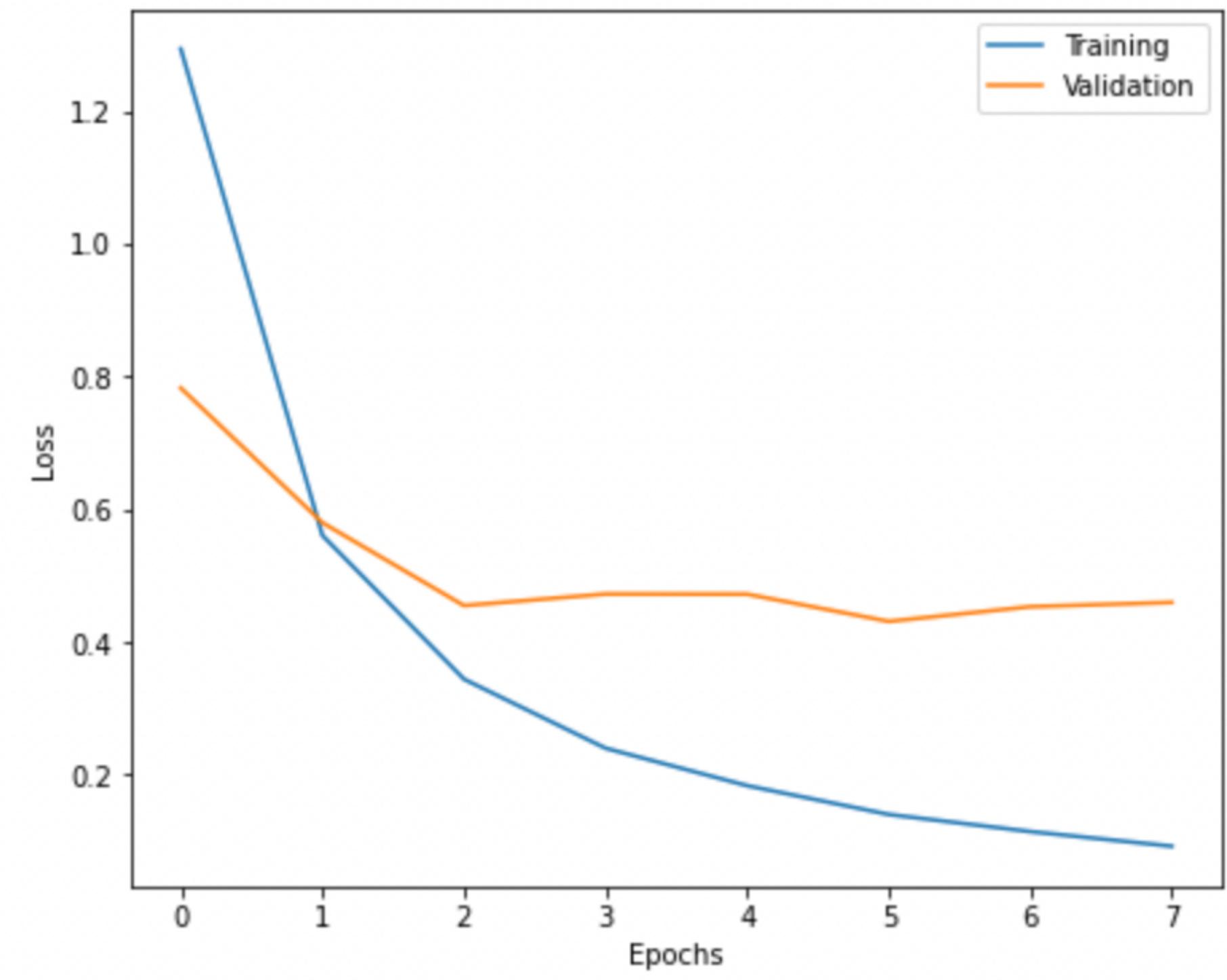
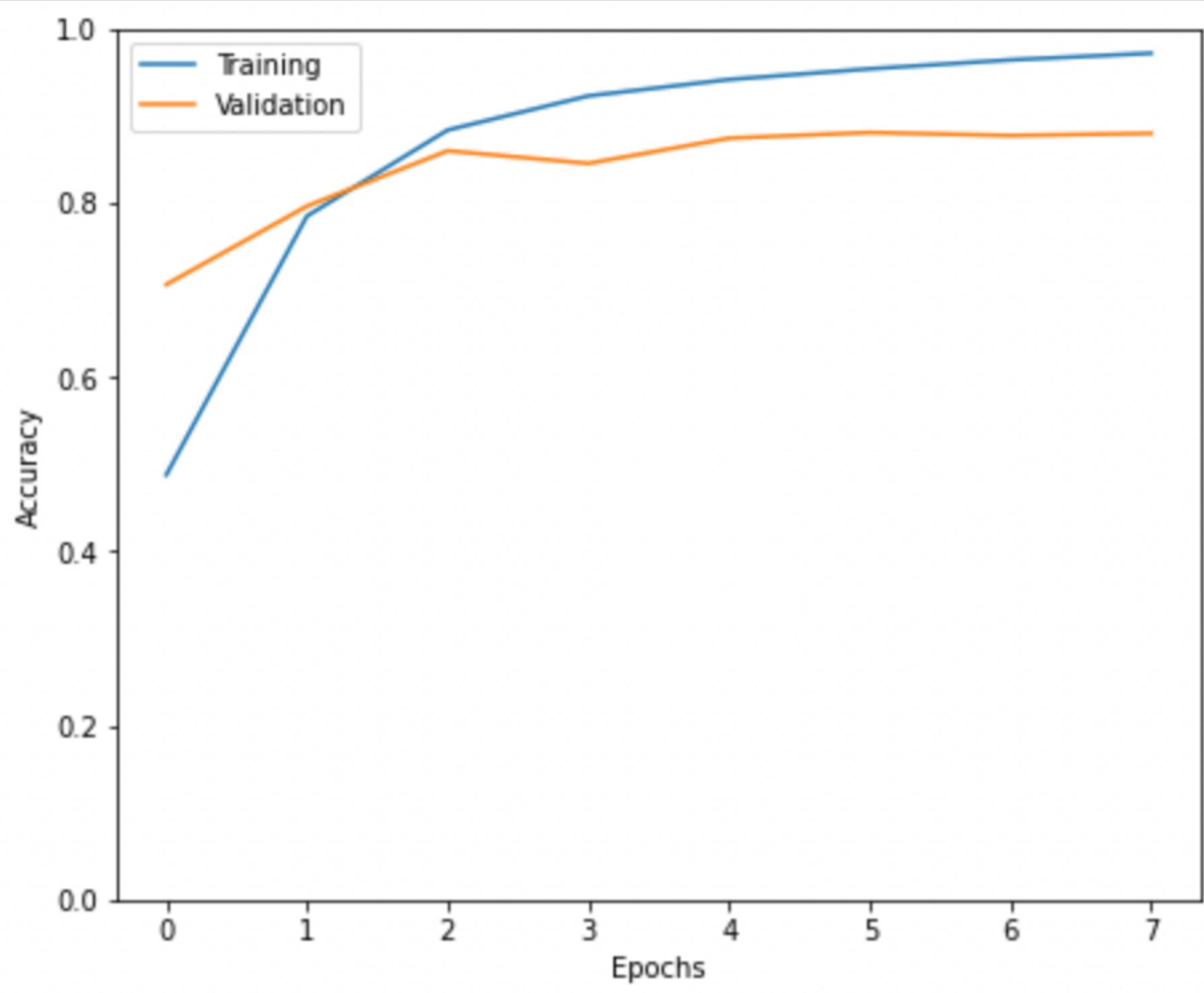
Layer (type)	Output Shape	Param #
embedding_3 (Embedding)	(None, 50, 16)	160000
bidirectional_6 (Bidirectional)	(None, 50, 40)	5920
bidirectional_7 (Bidirectional)	(None, 40)	9760
dense_3 (Dense)	(None, 6)	246

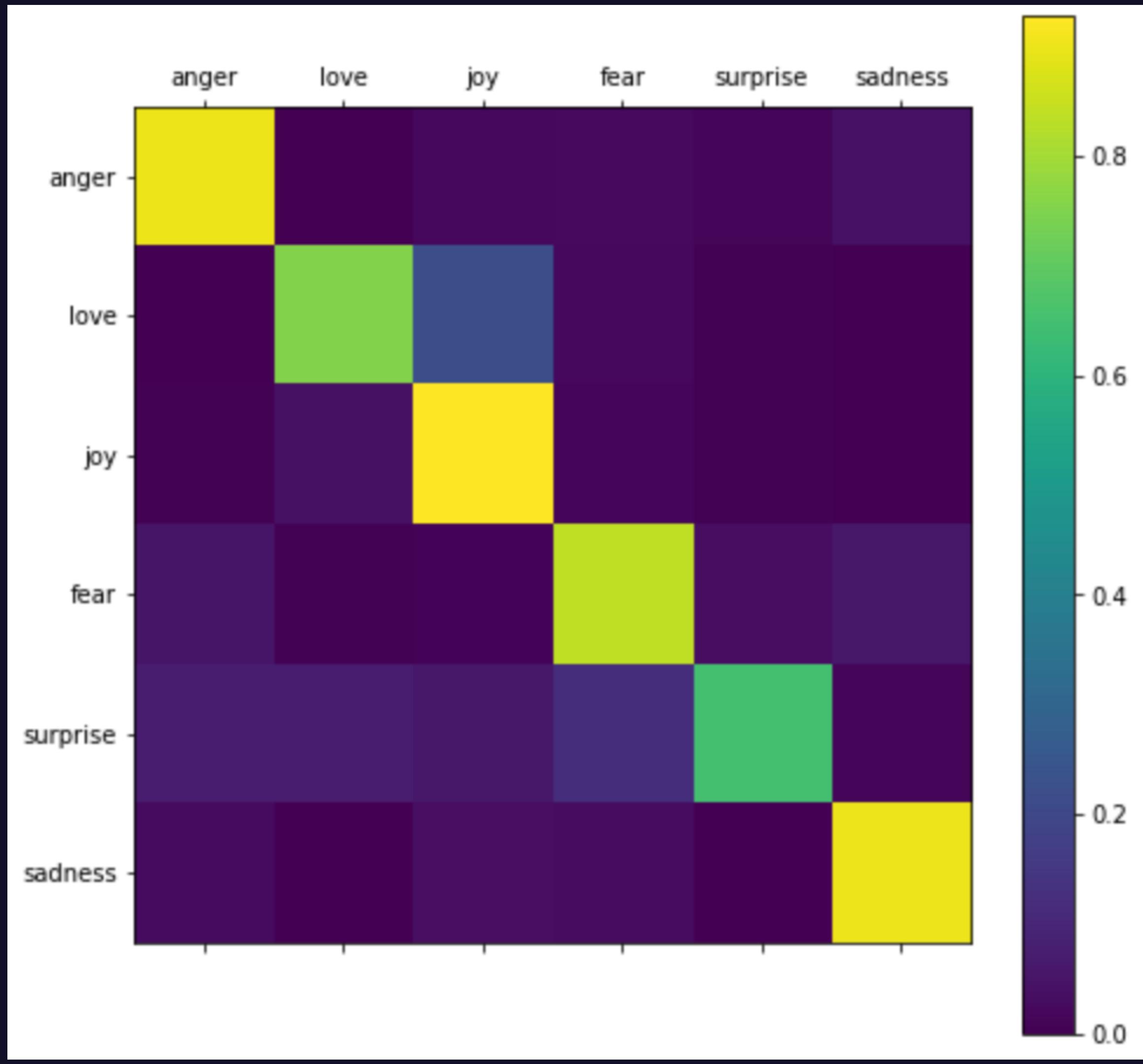
Total params: 175,926
Trainable params: 175,926
Non-trainable params: 0

Training & Evaluating

```
39] history = model.fit(  
    padded_train_sequences, train_labels,  
    validation_data = (val_sequences, val_labels),  
    epochs = 20,  
    callbacks = [  
        tf.keras.callbacks.EarlyStopping(monitor = 'val_accuracy', patience = 2)  
    ]  
)
```

```
Epoch 1/20  
500/500 [=====] - 41s 61ms/step - loss: 1.2935 - accuracy: 0.4874 - val_loss: 0.7826 - val_accuracy: 0.7060  
Epoch 2/20  
500/500 [=====] - 28s 57ms/step - loss: 0.5605 - accuracy: 0.7846 - val_loss: 0.5796 - val_accuracy: 0.7960  
Epoch 3/20  
500/500 [=====] - 28s 56ms/step - loss: 0.3439 - accuracy: 0.8833 - val_loss: 0.4547 - val_accuracy: 0.8595  
Epoch 4/20  
500/500 [=====] - 28s 56ms/step - loss: 0.2398 - accuracy: 0.9225 - val_loss: 0.4721 - val_accuracy: 0.8450  
Epoch 5/20  
500/500 [=====] - 28s 57ms/step - loss: 0.1834 - accuracy: 0.9413 - val_loss: 0.4718 - val_accuracy: 0.8740  
Epoch 6/20  
500/500 [=====] - 29s 57ms/step - loss: 0.1400 - accuracy: 0.9536 - val_loss: 0.4309 - val_accuracy: 0.8805  
Epoch 7/20  
500/500 [=====] - 28s 57ms/step - loss: 0.1142 - accuracy: 0.9638 - val_loss: 0.4530 - val_accuracy: 0.8770  
Epoch 8/20  
500/500 [=====] - 29s 57ms/step - loss: 0.0921 - accuracy: 0.9714 - val_loss: 0.4595 - val_accuracy: 0.8795
```

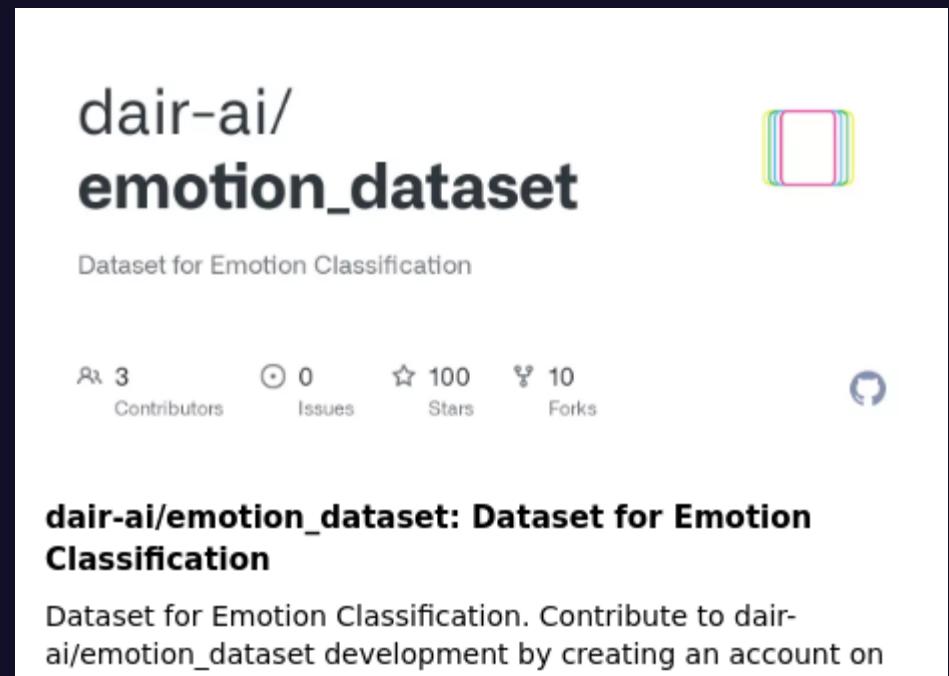




Dataset

https://github.com/dair-ai/emotion_dataset

dair-ai/
emotion_dataset



Dataset for Emotion Classification

3 Contributors 0 Issues 100 Stars 10 Forks

dair-ai/emotion_dataset: Dataset for Emotion Classification

Dataset for Emotion Classification. Contribute to dair-ai/emotion_dataset development by creating an account on GitHub.

THANKS

Blog: <https://cmkqwerty.github.io>

Twitter: @cmk_qwerty

LinkedIn: @cmkqwerty

Github: @cmkqwerty