

## learn-co-curriculum / dsc-tuning-neural-networks-with-regularization-lab-v2-1

Public

View license

1 star 155 forks

Star

Watch ▾

Code

Issues

Pull requests

Actions

Projects

Security

Insights

solution ▾

...

This branch is [11 commits ahead](#), [14 commits behind](#) master.



Cheffrey2000 fixed env conflicts ...

on Oct 22, 2020

12

[View code](#)

README.md

# Tuning Neural Networks with Regularization - Lab

## Introduction

In this lab, you'll use a train-test partition as well as a validation set to get better insights about how to tune neural networks using regularization techniques. You'll start by repeating the process from the last section: importing the data and performing preprocessing including one-hot encoding. From there, you'll define and compile the model like before.

## Objectives

You will be able to:

- Apply early stopping criteria with a neural network

- Apply L1, L2, and dropout regularization on a neural network
- Examine the effects of training with more data on a neural network

## Load the Data

---

Run the following cell to import some of the libraries and classes you'll need in this lab.

```
import pandas as pd
import numpy as np
import random
import matplotlib.pyplot as plt
%matplotlib inline
from sklearn.model_selection import train_test_split
from keras.utils.np_utils import to_categorical
from sklearn.preprocessing import LabelBinarizer
from keras.preprocessing.text import Tokenizer

import warnings
warnings.filterwarnings(action='ignore', category=FutureWarning)
```

The data is stored in the file 'Bank\_complaints.csv'. Load and preview the dataset.

```
# Load and preview the dataset
df = pd.read_csv('Bank_complaints.csv')
df.head()
```

<style scoped> .dataframe tbody tr th:only-of-type { vertical-align: middle; }

.dataframe tbody tr th {
 vertical-align: top;
}

.dataframe thead th {
 text-align: right;
}

</style>

	Product	Consumer complaint narrative
0	Student loan	In XX/XX/XXXX I filled out the Fedlaon applica...
1	Student loan	I am being contacted by a debt collector for p...

	Product	Consumer complaint narrative
2	Student loan	I cosigned XXXX student loans at SallieMae for...
3	Student loan	Navient has systematically and illegally failed...
4	Student loan	My wife became eligible for XXXX Loan Forgiven...

## Preprocessing Overview

---

Before you begin to practice some of your new tools such as regularization and optimization, let's practice munging some data as you did in the previous section with bank complaints. Recall some techniques:

- Sampling in order to reduce training time (investigate model accuracy vs data size later on)
- Train - test split
- One-hot encoding your complaint text
- Transforming your category labels

## Preprocessing: Generate a Random Sample

---

Since you have quite a bit of data and training neural networks takes a substantial amount of time and resources, downsample in order to test your initial pipeline. Going forward, these can be interesting areas of investigation: how does your model's performance change as you increase (or decrease) the size of your dataset?

- Generate a random sample of 10,000 observations using seed 123 for consistency of results.
- Split this sample into `x` and `y`

```
# Downsample the data
df_sample = df.sample(10000, random_state=123)

# Split the data into X and y
y = df_sample['Product']
x = df_sample['Consumer complaint narrative']
```

## Train-test split

- Split the data into training and test sets
- Assign 1500 observations to the test set and use 42 as the seed

```
# Split data into training and test sets
X_train, X_test, y_train, y_test = train_test_split(X, y, test_size=1500, random_st
```

## Validation set

As mentioned in the previous lesson, it is good practice to set aside a validation set, which is then used during hyperparameter tuning. Afterwards, when you have decided upon a final model, the test set can then be used to determine an unbiased performance of the model.

Run the cell below to further divide the training data into training and validation sets.

```
# Split the data into training and validation sets
X_train_final, X_val, y_train_final, y_val = train_test_split(X_train, y_train, test
```

## Preprocessing: One-hot Encoding the Complaints

As before, you need to do some preprocessing before building a neural network model.

- Keep the 2,000 most common words and use one-hot encoding to reformat the complaints into a matrix of vectors
- Transform the training, validate, and test sets

```
# Use one-hot encoding to reformat the complaints into a matrix of vectors
# Only keep the 2000 most common words

tokenizer = Tokenizer(num_words=2000)
tokenizer.fit_on_texts(X_train_final)

X_train_tokens = tokenizer.texts_to_matrix(X_train_final, mode='binary')
X_val_tokens = tokenizer.texts_to_matrix(X_val, mode='binary')
X_test_tokens = tokenizer.texts_to_matrix(X_test, mode='binary')
```

## Preprocessing: Encoding the Products

Similarly, now transform the descriptive product labels to integers labels. After transforming them to integer labels, retransform them into a matrix of binary flags, one for each of the various product labels.

**Note:** This is similar to your previous work with dummy variables. Each of the various product categories will be its own column, and each observation will be a row. In turn, each of these observation rows will have a 1 in the column associated with it's label, and all other entries for the row will be zero.

Transform the training, validate, and test sets.

```
# Transform the product labels to numerical values
lb = LabelBinarizer()
lb.fit(y_train_final)

y_train_lb = to_categorical(lb.transform(y_train_final))[:, :, 1]
y_val_lb = to_categorical(lb.transform(y_val))[:, :, 1]
y_test_lb = to_categorical(lb.transform(y_test))[:, :, 1]
```

## A Baseline Model

Rebuild a fully connected (Dense) layer network:

- Use 2 hidden layers with 50 units in the first and 25 in the second layer, both with 'relu' activation functions (since you are dealing with a multiclass problem, classifying the complaints into 7 classes)
- Use a 'softmax' activation function for the output layer

```
# Build a baseline neural network model using Keras
random.seed(123)
from keras import models
from keras import layers
baseline_model = models.Sequential()
baseline_model.add(layers.Dense(50, activation='relu', input_shape=(2000,)))
baseline_model.add(layers.Dense(25, activation='relu'))
baseline_model.add(layers.Dense(7, activation='softmax'))
```

## Compile the Model

Compile this model with:

- a stochastic gradient descent optimizer
- 'categorical\_crossentropy' as the loss function
- a focus on 'accuracy'

```
# Compile the model
baseline_model.compile(optimizer='SGD',
                      loss='categorical_crossentropy',
                      metrics=['acc'])
```

## Train the Model

- Train the model for 150 epochs in mini-batches of 256 samples
- Include the `validation_data` argument to ensure you keep track of the validation loss

```
# Train the model
baseline_model_val = baseline_model.fit(X_train_tokens,
                                         y_train_lb,
                                         epochs=150,
                                         batch_size=256,
                                         validation_data=(X_val_tokens, y_val_lb))
```

```
Epoch 1/150
30/30 [=====] - 0s 6ms/step - loss: 1.9482 - acc: 0.1567
- val_loss: 1.9364 - val_acc: 0.1590
Epoch 2/150
30/30 [=====] - 0s 2ms/step - loss: 1.9260 - acc: 0.1809
- val_loss: 1.9186 - val_acc: 0.1850
Epoch 3/150
30/30 [=====] - 0s 3ms/step - loss: 1.9042 - acc: 0.1980
- val_loss: 1.9013 - val_acc: 0.1960
Epoch 4/150
30/30 [=====] - 0s 2ms/step - loss: 1.8829 - acc: 0.2091
- val_loss: 1.8834 - val_acc: 0.1970
Epoch 5/150
30/30 [=====] - 0s 2ms/step - loss: 1.8605 - acc: 0.2211
- val_loss: 1.8642 - val_acc: 0.2010
Epoch 6/150
30/30 [=====] - 0s 2ms/step - loss: 1.8364 - acc: 0.2349
- val_loss: 1.8424 - val_acc: 0.2170
Epoch 7/150
30/30 [=====] - 0s 3ms/step - loss: 1.8098 - acc: 0.2549
- val_loss: 1.8166 - val_acc: 0.2320
Epoch 8/150
30/30 [=====] - 0s 3ms/step - loss: 1.7792 - acc: 0.2856
```

```
- val_loss: 1.7864 - val_acc: 0.2680
Epoch 9/150
30/30 [=====] - 0s 3ms/step - loss: 1.7436 - acc: 0.3247
- val_loss: 1.7509 - val_acc: 0.3110
Epoch 10/150
30/30 [=====] - 0s 3ms/step - loss: 1.7036 - acc: 0.3611
- val_loss: 1.7102 - val_acc: 0.3480
Epoch 11/150
30/30 [=====] - 0s 3ms/step - loss: 1.6600 - acc: 0.4025
- val_loss: 1.6673 - val_acc: 0.3760
Epoch 12/150
30/30 [=====] - 0s 3ms/step - loss: 1.6137 - acc: 0.4308
- val_loss: 1.6219 - val_acc: 0.4110
Epoch 13/150
30/30 [=====] - 0s 3ms/step - loss: 1.5649 - acc: 0.4524
- val_loss: 1.5753 - val_acc: 0.4180
Epoch 14/150
30/30 [=====] - 0s 3ms/step - loss: 1.5151 - acc: 0.4729
- val_loss: 1.5268 - val_acc: 0.4530
Epoch 15/150
30/30 [=====] - 0s 3ms/step - loss: 1.4646 - acc: 0.4980
- val_loss: 1.4778 - val_acc: 0.4760
Epoch 16/150
30/30 [=====] - 0s 2ms/step - loss: 1.4141 - acc: 0.5199
- val_loss: 1.4284 - val_acc: 0.5020
Epoch 17/150
30/30 [=====] - 0s 3ms/step - loss: 1.3649 - acc: 0.5425
- val_loss: 1.3801 - val_acc: 0.5290
Epoch 18/150
30/30 [=====] - 0s 3ms/step - loss: 1.3169 - acc: 0.5632
- val_loss: 1.3348 - val_acc: 0.5540
Epoch 19/150
30/30 [=====] - 0s 3ms/step - loss: 1.2706 - acc: 0.5881
- val_loss: 1.2901 - val_acc: 0.5730
Epoch 20/150
30/30 [=====] - 0s 2ms/step - loss: 1.2262 - acc: 0.6091
- val_loss: 1.2470 - val_acc: 0.5900
Epoch 21/150
30/30 [=====] - 0s 3ms/step - loss: 1.1843 - acc: 0.6285
- val_loss: 1.2084 - val_acc: 0.6040
Epoch 22/150
30/30 [=====] - 0s 3ms/step - loss: 1.1448 - acc: 0.6387
- val_loss: 1.1713 - val_acc: 0.6110
Epoch 23/150
30/30 [=====] - 0s 3ms/step - loss: 1.1078 - acc: 0.6560
- val_loss: 1.1355 - val_acc: 0.6270
Epoch 24/150
30/30 [=====] - 0s 3ms/step - loss: 1.0731 - acc: 0.6659
- val_loss: 1.1033 - val_acc: 0.6320
```

```
Epoch 25/150
30/30 [=====] - 0s 2ms/step - loss: 1.0414 - acc: 0.6760
- val_loss: 1.0744 - val_acc: 0.6380
Epoch 26/150
30/30 [=====] - 0s 3ms/step - loss: 1.0116 - acc: 0.6844
- val_loss: 1.0443 - val_acc: 0.6500
Epoch 27/150
30/30 [=====] - 0s 3ms/step - loss: 0.9835 - acc: 0.6929
- val_loss: 1.0183 - val_acc: 0.6540
Epoch 28/150
30/30 [=====] - 0s 3ms/step - loss: 0.9578 - acc: 0.6985
- val_loss: 0.9941 - val_acc: 0.6500
Epoch 29/150
30/30 [=====] - 0s 3ms/step - loss: 0.9339 - acc: 0.7031
- val_loss: 0.9716 - val_acc: 0.6620
Epoch 30/150
30/30 [=====] - 0s 3ms/step - loss: 0.9113 - acc: 0.7099
- val_loss: 0.9543 - val_acc: 0.6580
Epoch 31/150
30/30 [=====] - 0s 3ms/step - loss: 0.8904 - acc: 0.7163
- val_loss: 0.9340 - val_acc: 0.6680
Epoch 32/150
30/30 [=====] - 0s 2ms/step - loss: 0.8714 - acc: 0.7189
- val_loss: 0.9137 - val_acc: 0.6760
Epoch 33/150
30/30 [=====] - 0s 3ms/step - loss: 0.8530 - acc: 0.7260
- val_loss: 0.9020 - val_acc: 0.6690
Epoch 34/150
30/30 [=====] - 0s 3ms/step - loss: 0.8362 - acc: 0.7289
- val_loss: 0.8829 - val_acc: 0.6790
Epoch 35/150
30/30 [=====] - 0s 3ms/step - loss: 0.8203 - acc: 0.7304
- val_loss: 0.8693 - val_acc: 0.6820
Epoch 36/150
30/30 [=====] - 0s 3ms/step - loss: 0.8049 - acc: 0.7373
- val_loss: 0.8569 - val_acc: 0.6790
Epoch 37/150
30/30 [=====] - 0s 3ms/step - loss: 0.7905 - acc: 0.7383
- val_loss: 0.8442 - val_acc: 0.6900
Epoch 38/150
30/30 [=====] - 0s 2ms/step - loss: 0.7774 - acc: 0.7432
- val_loss: 0.8364 - val_acc: 0.6860
Epoch 39/150
30/30 [=====] - 0s 3ms/step - loss: 0.7647 - acc: 0.7427
- val_loss: 0.8238 - val_acc: 0.6970
Epoch 40/150
30/30 [=====] - 0s 3ms/step - loss: 0.7523 - acc: 0.7497
- val_loss: 0.8122 - val_acc: 0.6990
Epoch 41/150
```

```
30/30 [=====] - 0s 2ms/step - loss: 0.7412 - acc: 0.7537
- val_loss: 0.8041 - val_acc: 0.6960
Epoch 42/150
30/30 [=====] - 0s 3ms/step - loss: 0.7300 - acc: 0.7556
- val_loss: 0.7974 - val_acc: 0.6950
Epoch 43/150
30/30 [=====] - 0s 3ms/step - loss: 0.7194 - acc: 0.7583
- val_loss: 0.7877 - val_acc: 0.7000
Epoch 44/150
30/30 [=====] - 0s 3ms/step - loss: 0.7095 - acc: 0.7633
- val_loss: 0.7810 - val_acc: 0.7040
Epoch 45/150
30/30 [=====] - 0s 3ms/step - loss: 0.7001 - acc: 0.7640
- val_loss: 0.7766 - val_acc: 0.7040
Epoch 46/150
30/30 [=====] - 0s 3ms/step - loss: 0.6907 - acc: 0.7688
- val_loss: 0.7741 - val_acc: 0.7100
Epoch 47/150
30/30 [=====] - 0s 3ms/step - loss: 0.6824 - acc: 0.7681
- val_loss: 0.7627 - val_acc: 0.7100
Epoch 48/150
30/30 [=====] - 0s 2ms/step - loss: 0.6740 - acc: 0.7697
- val_loss: 0.7584 - val_acc: 0.7130
Epoch 49/150
30/30 [=====] - 0s 3ms/step - loss: 0.6658 - acc: 0.7763
- val_loss: 0.7513 - val_acc: 0.7120
Epoch 50/150
30/30 [=====] - 0s 3ms/step - loss: 0.6579 - acc: 0.7787
- val_loss: 0.7454 - val_acc: 0.7070
Epoch 51/150
30/30 [=====] - 0s 2ms/step - loss: 0.6504 - acc: 0.7835
- val_loss: 0.7401 - val_acc: 0.7110
Epoch 52/150
30/30 [=====] - 0s 3ms/step - loss: 0.6432 - acc: 0.7820
- val_loss: 0.7399 - val_acc: 0.7170
Epoch 53/150
30/30 [=====] - 0s 3ms/step - loss: 0.6357 - acc: 0.7849
- val_loss: 0.7309 - val_acc: 0.7120
Epoch 54/150
30/30 [=====] - 0s 2ms/step - loss: 0.6289 - acc: 0.7904
- val_loss: 0.7255 - val_acc: 0.7140
Epoch 55/150
30/30 [=====] - 0s 3ms/step - loss: 0.6221 - acc: 0.7923
- val_loss: 0.7252 - val_acc: 0.7210
Epoch 56/150
30/30 [=====] - 0s 3ms/step - loss: 0.6155 - acc: 0.7909
- val_loss: 0.7221 - val_acc: 0.7160
Epoch 57/150
30/30 [=====] - 0s 2ms/step - loss: 0.6094 - acc: 0.7956
```

```
- val_loss: 0.7196 - val_acc: 0.7120
Epoch 58/150
30/30 [=====] - 0s 2ms/step - loss: 0.6037 - acc: 0.7981
- val_loss: 0.7145 - val_acc: 0.7180
Epoch 59/150
30/30 [=====] - 0s 2ms/step - loss: 0.5972 - acc: 0.8016
- val_loss: 0.7147 - val_acc: 0.7260
Epoch 60/150
30/30 [=====] - 0s 3ms/step - loss: 0.5915 - acc: 0.8019
- val_loss: 0.7079 - val_acc: 0.7230
Epoch 61/150
30/30 [=====] - 0s 3ms/step - loss: 0.5853 - acc: 0.8064
- val_loss: 0.7150 - val_acc: 0.7320
Epoch 62/150
30/30 [=====] - 0s 3ms/step - loss: 0.5800 - acc: 0.8039
- val_loss: 0.7037 - val_acc: 0.7230
Epoch 63/150
30/30 [=====] - 0s 3ms/step - loss: 0.5744 - acc: 0.8072
- val_loss: 0.6962 - val_acc: 0.7240
Epoch 64/150
30/30 [=====] - 0s 3ms/step - loss: 0.5690 - acc: 0.8081
- val_loss: 0.6999 - val_acc: 0.7210
Epoch 65/150
30/30 [=====] - 0s 3ms/step - loss: 0.5642 - acc: 0.8123
- val_loss: 0.6980 - val_acc: 0.7290
Epoch 66/150
30/30 [=====] - 0s 3ms/step - loss: 0.5592 - acc: 0.8111
- val_loss: 0.6902 - val_acc: 0.7250
Epoch 67/150
30/30 [=====] - 0s 3ms/step - loss: 0.5539 - acc: 0.8156
- val_loss: 0.6931 - val_acc: 0.7280
Epoch 68/150
30/30 [=====] - 0s 3ms/step - loss: 0.5490 - acc: 0.8157
- val_loss: 0.6865 - val_acc: 0.7210
Epoch 69/150
30/30 [=====] - 0s 2ms/step - loss: 0.5444 - acc: 0.8192
- val_loss: 0.6910 - val_acc: 0.7370
Epoch 70/150
30/30 [=====] - 0s 3ms/step - loss: 0.5392 - acc: 0.8203
- val_loss: 0.6855 - val_acc: 0.7290
Epoch 71/150
30/30 [=====] - 0s 3ms/step - loss: 0.5344 - acc: 0.8229
- val_loss: 0.6810 - val_acc: 0.7230
Epoch 72/150
30/30 [=====] - 0s 3ms/step - loss: 0.5302 - acc: 0.8247
- val_loss: 0.6846 - val_acc: 0.7350
Epoch 73/150
30/30 [=====] - 0s 3ms/step - loss: 0.5253 - acc: 0.8253
- val_loss: 0.6764 - val_acc: 0.7360
```

```
Epoch 74/150
30/30 [=====] - 0s 2ms/step - loss: 0.5213 - acc: 0.8284
- val_loss: 0.6790 - val_acc: 0.7350
Epoch 75/150
30/30 [=====] - 0s 2ms/step - loss: 0.5168 - acc: 0.8307
- val_loss: 0.6769 - val_acc: 0.7360
Epoch 76/150
30/30 [=====] - 0s 3ms/step - loss: 0.5121 - acc: 0.8296
- val_loss: 0.6742 - val_acc: 0.7300
Epoch 77/150
30/30 [=====] - 0s 2ms/step - loss: 0.5082 - acc: 0.8325
- val_loss: 0.6769 - val_acc: 0.7410
Epoch 78/150
30/30 [=====] - 0s 2ms/step - loss: 0.5039 - acc: 0.8341
- val_loss: 0.6727 - val_acc: 0.7290
Epoch 79/150
30/30 [=====] - 0s 3ms/step - loss: 0.5000 - acc: 0.8337
- val_loss: 0.6740 - val_acc: 0.7320
Epoch 80/150
30/30 [=====] - 0s 3ms/step - loss: 0.4955 - acc: 0.8377
- val_loss: 0.6706 - val_acc: 0.7380
Epoch 81/150
30/30 [=====] - 0s 3ms/step - loss: 0.4914 - acc: 0.8387
- val_loss: 0.6760 - val_acc: 0.7350
Epoch 82/150
30/30 [=====] - 0s 3ms/step - loss: 0.4872 - acc: 0.8393
- val_loss: 0.6688 - val_acc: 0.7300
Epoch 83/150
30/30 [=====] - 0s 2ms/step - loss: 0.4840 - acc: 0.8397
- val_loss: 0.6668 - val_acc: 0.7410
Epoch 84/150
30/30 [=====] - 0s 3ms/step - loss: 0.4798 - acc: 0.8436
- val_loss: 0.6681 - val_acc: 0.7260
Epoch 85/150
30/30 [=====] - 0s 2ms/step - loss: 0.4767 - acc: 0.8427
- val_loss: 0.6669 - val_acc: 0.7370
Epoch 86/150
30/30 [=====] - 0s 2ms/step - loss: 0.4722 - acc: 0.8463
- val_loss: 0.6673 - val_acc: 0.7270
Epoch 87/150
30/30 [=====] - 0s 3ms/step - loss: 0.4686 - acc: 0.8464
- val_loss: 0.6665 - val_acc: 0.7400
Epoch 88/150
30/30 [=====] - 0s 2ms/step - loss: 0.4652 - acc: 0.8455
- val_loss: 0.6660 - val_acc: 0.7430
Epoch 89/150
30/30 [=====] - 0s 2ms/step - loss: 0.4617 - acc: 0.8497
- val_loss: 0.6633 - val_acc: 0.7430
Epoch 90/150
```

```
30/30 [=====] - 0s 3ms/step - loss: 0.4576 - acc: 0.8504
- val_loss: 0.6629 - val_acc: 0.7420
Epoch 91/150
30/30 [=====] - 0s 2ms/step - loss: 0.4542 - acc: 0.8512
- val_loss: 0.6619 - val_acc: 0.7330
Epoch 92/150
30/30 [=====] - 0s 2ms/step - loss: 0.4511 - acc: 0.8531
- val_loss: 0.6660 - val_acc: 0.7440
Epoch 93/150
30/30 [=====] - 0s 2ms/step - loss: 0.4479 - acc: 0.8555
- val_loss: 0.6622 - val_acc: 0.7460
Epoch 94/150
30/30 [=====] - 0s 3ms/step - loss: 0.4443 - acc: 0.8541
- val_loss: 0.6596 - val_acc: 0.7310
Epoch 95/150
30/30 [=====] - 0s 2ms/step - loss: 0.4407 - acc: 0.8569
- val_loss: 0.6642 - val_acc: 0.7370
Epoch 96/150
30/30 [=====] - 0s 2ms/step - loss: 0.4375 - acc: 0.8577
- val_loss: 0.6635 - val_acc: 0.7420
Epoch 97/150
30/30 [=====] - 0s 3ms/step - loss: 0.4339 - acc: 0.8593
- val_loss: 0.6600 - val_acc: 0.7350
Epoch 98/150
30/30 [=====] - 0s 3ms/step - loss: 0.4310 - acc: 0.8631
- val_loss: 0.6635 - val_acc: 0.7420
Epoch 99/150
30/30 [=====] - 0s 2ms/step - loss: 0.4277 - acc: 0.8635
- val_loss: 0.6618 - val_acc: 0.7440
Epoch 100/150
30/30 [=====] - 0s 3ms/step - loss: 0.4239 - acc: 0.8628
- val_loss: 0.6619 - val_acc: 0.7450
Epoch 101/150
30/30 [=====] - 0s 2ms/step - loss: 0.4211 - acc: 0.8669
- val_loss: 0.6642 - val_acc: 0.7410
Epoch 102/150
30/30 [=====] - 0s 2ms/step - loss: 0.4180 - acc: 0.8680
- val_loss: 0.6600 - val_acc: 0.7450
Epoch 103/150
30/30 [=====] - 0s 2ms/step - loss: 0.4151 - acc: 0.8687
- val_loss: 0.6627 - val_acc: 0.7440
Epoch 104/150
30/30 [=====] - 0s 2ms/step - loss: 0.4120 - acc: 0.8703
- val_loss: 0.6554 - val_acc: 0.7450
Epoch 105/150
30/30 [=====] - 0s 3ms/step - loss: 0.4085 - acc: 0.8712
- val_loss: 0.6574 - val_acc: 0.7320
Epoch 106/150
30/30 [=====] - 0s 2ms/step - loss: 0.4060 - acc: 0.8729
```

```
- val_loss: 0.6583 - val_acc: 0.7420
Epoch 107/150
30/30 [=====] - 0s 2ms/step - loss: 0.4027 - acc: 0.8743
- val_loss: 0.6632 - val_acc: 0.7490
Epoch 108/150
30/30 [=====] - 0s 2ms/step - loss: 0.3996 - acc: 0.8751
- val_loss: 0.6636 - val_acc: 0.7440
Epoch 109/150
30/30 [=====] - 0s 2ms/step - loss: 0.3965 - acc: 0.8767
- val_loss: 0.6590 - val_acc: 0.7430
Epoch 110/150
30/30 [=====] - 0s 2ms/step - loss: 0.3943 - acc: 0.8755
- val_loss: 0.6556 - val_acc: 0.7330
Epoch 111/150
30/30 [=====] - 0s 2ms/step - loss: 0.3913 - acc: 0.8789
- val_loss: 0.6592 - val_acc: 0.7490
Epoch 112/150
30/30 [=====] - 0s 3ms/step - loss: 0.3883 - acc: 0.8799
- val_loss: 0.6639 - val_acc: 0.7510
Epoch 113/150
30/30 [=====] - 0s 3ms/step - loss: 0.3853 - acc: 0.8803
- val_loss: 0.6583 - val_acc: 0.7440
Epoch 114/150
30/30 [=====] - 0s 2ms/step - loss: 0.3825 - acc: 0.8824
- val_loss: 0.6579 - val_acc: 0.7470
Epoch 115/150
30/30 [=====] - 0s 2ms/step - loss: 0.3796 - acc: 0.8839
- val_loss: 0.6602 - val_acc: 0.7430
Epoch 116/150
30/30 [=====] - 0s 3ms/step - loss: 0.3767 - acc: 0.8839
- val_loss: 0.6569 - val_acc: 0.7410
Epoch 117/150
30/30 [=====] - 0s 2ms/step - loss: 0.3740 - acc: 0.8840
- val_loss: 0.6633 - val_acc: 0.7520
Epoch 118/150
30/30 [=====] - 0s 2ms/step - loss: 0.3714 - acc: 0.8856
- val_loss: 0.6584 - val_acc: 0.7370
Epoch 119/150
30/30 [=====] - 0s 2ms/step - loss: 0.3690 - acc: 0.8860
- val_loss: 0.6608 - val_acc: 0.7410
Epoch 120/150
30/30 [=====] - 0s 3ms/step - loss: 0.3659 - acc: 0.8888
- val_loss: 0.6648 - val_acc: 0.7490
Epoch 121/150
30/30 [=====] - 0s 2ms/step - loss: 0.3637 - acc: 0.8887
- val_loss: 0.6652 - val_acc: 0.7430
Epoch 122/150
30/30 [=====] - 0s 2ms/step - loss: 0.3610 - acc: 0.8891
- val_loss: 0.6664 - val_acc: 0.7490
```

```
Epoch 123/150
30/30 [=====] - 0s 2ms/step - loss: 0.3582 - acc: 0.8928
- val_loss: 0.6638 - val_acc: 0.7410
Epoch 124/150
30/30 [=====] - 0s 3ms/step - loss: 0.3555 - acc: 0.8936
- val_loss: 0.6658 - val_acc: 0.7450
Epoch 125/150
30/30 [=====] - 0s 2ms/step - loss: 0.3531 - acc: 0.8944
- val_loss: 0.6705 - val_acc: 0.7480
Epoch 126/150
30/30 [=====] - 0s 3ms/step - loss: 0.3510 - acc: 0.8957
- val_loss: 0.6625 - val_acc: 0.7480
Epoch 127/150
30/30 [=====] - 0s 3ms/step - loss: 0.3481 - acc: 0.8951
- val_loss: 0.6648 - val_acc: 0.7510
Epoch 128/150
30/30 [=====] - 0s 3ms/step - loss: 0.3456 - acc: 0.8960
- val_loss: 0.6650 - val_acc: 0.7420
Epoch 129/150
30/30 [=====] - 0s 3ms/step - loss: 0.3427 - acc: 0.8975
- val_loss: 0.6623 - val_acc: 0.7500
Epoch 130/150
30/30 [=====] - 0s 2ms/step - loss: 0.3406 - acc: 0.8973
- val_loss: 0.6644 - val_acc: 0.7480
Epoch 131/150
30/30 [=====] - 0s 3ms/step - loss: 0.3381 - acc: 0.9003
- val_loss: 0.6686 - val_acc: 0.7480
Epoch 132/150
30/30 [=====] - 0s 2ms/step - loss: 0.3356 - acc: 0.9021
- val_loss: 0.6753 - val_acc: 0.7450
Epoch 133/150
30/30 [=====] - 0s 2ms/step - loss: 0.3336 - acc: 0.9015
- val_loss: 0.6697 - val_acc: 0.7480
Epoch 134/150
30/30 [=====] - 0s 3ms/step - loss: 0.3310 - acc: 0.9027
- val_loss: 0.6708 - val_acc: 0.7440
Epoch 135/150
30/30 [=====] - 0s 3ms/step - loss: 0.3288 - acc: 0.9037
- val_loss: 0.6693 - val_acc: 0.7470
Epoch 136/150
30/30 [=====] - 0s 2ms/step - loss: 0.3267 - acc: 0.9049
- val_loss: 0.6717 - val_acc: 0.7440
Epoch 137/150
30/30 [=====] - 0s 3ms/step - loss: 0.3243 - acc: 0.9040
- val_loss: 0.6699 - val_acc: 0.7430
Epoch 138/150
30/30 [=====] - 0s 3ms/step - loss: 0.3218 - acc: 0.9076
- val_loss: 0.6716 - val_acc: 0.7420
Epoch 139/150
```

```
30/30 [=====] - 0s 3ms/step - loss: 0.3194 - acc: 0.9068
- val_loss: 0.6711 - val_acc: 0.7470
Epoch 140/150
30/30 [=====] - 0s 2ms/step - loss: 0.3172 - acc: 0.9083
- val_loss: 0.6727 - val_acc: 0.7430
Epoch 141/150
30/30 [=====] - 0s 2ms/step - loss: 0.3149 - acc: 0.9095
- val_loss: 0.6720 - val_acc: 0.7430
Epoch 142/150
30/30 [=====] - 0s 3ms/step - loss: 0.3133 - acc: 0.9100
- val_loss: 0.6730 - val_acc: 0.7390
Epoch 143/150
30/30 [=====] - 0s 2ms/step - loss: 0.3102 - acc: 0.9113
- val_loss: 0.6798 - val_acc: 0.7460
Epoch 144/150
30/30 [=====] - 0s 2ms/step - loss: 0.3084 - acc: 0.9112
- val_loss: 0.6754 - val_acc: 0.7390
Epoch 145/150
30/30 [=====] - 0s 2ms/step - loss: 0.3061 - acc: 0.9131
- val_loss: 0.6808 - val_acc: 0.7410
Epoch 146/150
30/30 [=====] - 0s 3ms/step - loss: 0.3043 - acc: 0.9135
- val_loss: 0.6764 - val_acc: 0.7460
Epoch 147/150
30/30 [=====] - 0s 3ms/step - loss: 0.3020 - acc: 0.9141
- val_loss: 0.6781 - val_acc: 0.7420
Epoch 148/150
30/30 [=====] - 0s 3ms/step - loss: 0.2998 - acc: 0.9144
- val_loss: 0.6846 - val_acc: 0.7480
Epoch 149/150
30/30 [=====] - 0s 2ms/step - loss: 0.2974 - acc: 0.9172
- val_loss: 0.6773 - val_acc: 0.7440
Epoch 150/150
30/30 [=====] - 0s 3ms/step - loss: 0.2959 - acc: 0.9175
- val_loss: 0.6794 - val_acc: 0.7420
```

```
# Access the history attribute and store the dictionary
baseline_model_val_dict = baseline_model_val.history
```

```
# Print the keys
baseline_model_val_dict.keys()

dict_keys(['loss', 'acc', 'val_loss', 'val_acc'])
```

Evaluate this model on the training data:

```
results_train = baseline_model.evaluate(X_train_tokens, y_train_lb)
print('-----')
print(f'Training Loss: {results_train[0]:.3} \nTraining Accuracy: {results_train[1]}:
```

```
235/235 [=====] - 0s 578us/step - loss: 0.2923 - acc: 0.9189
-----
Training Loss: 0.292
Training Accuracy: 0.919
```

Evaluate this model on the test data:

```
results_test = baseline_model.evaluate(X_test_tokens, y_test_lb)
print('-----')
print(f'Test Loss: {results_test[0]:.3} \nTest Accuracy: {results_test[1]:.3}')
```

```
47/47 [=====] - 0s 646us/step - loss: 0.6197 - acc: 0.7753
-----
Test Loss: 0.62
Test Accuracy: 0.775
```

## Plot the Results

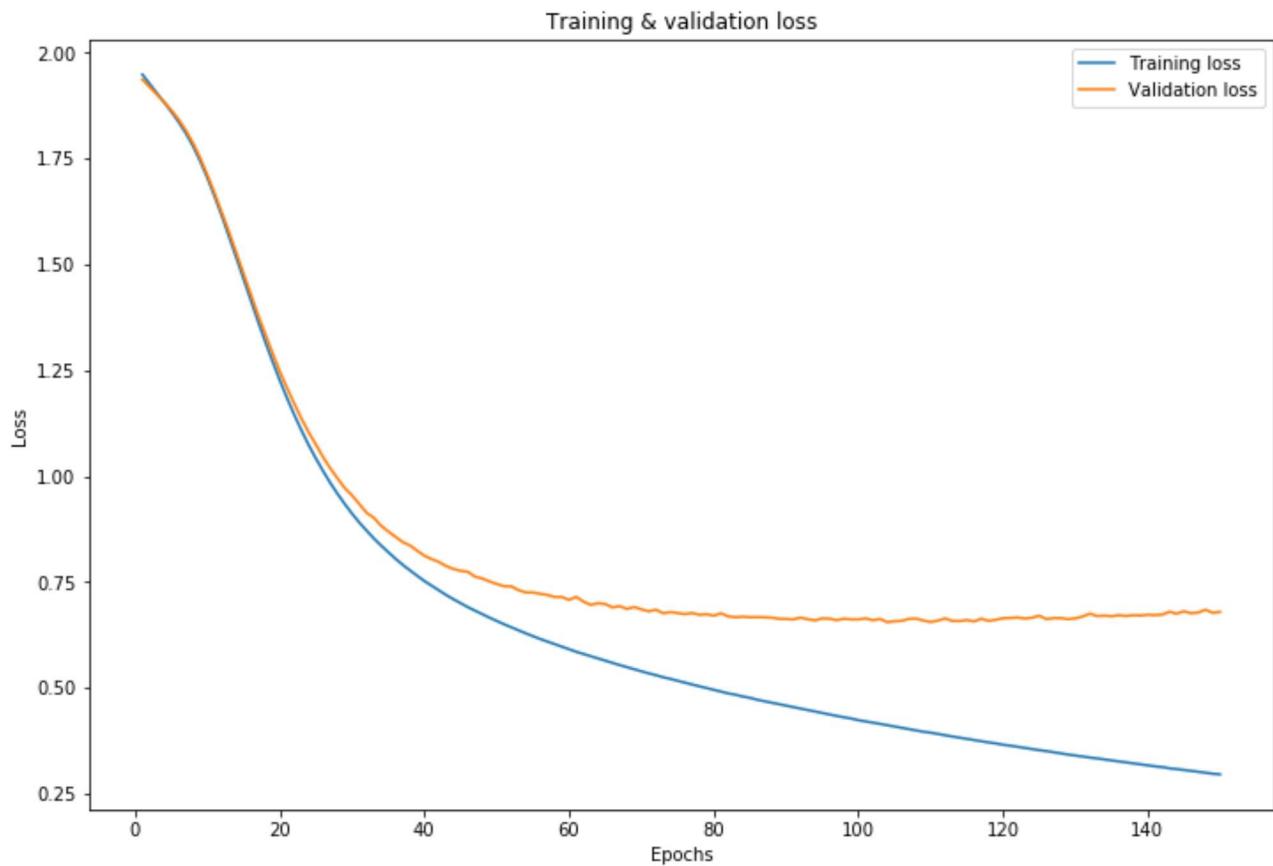
Plot the loss versus the number of epochs. Be sure to include the training and the validation loss in the same plot.

```
fig, ax = plt.subplots(figsize=(12, 8))

loss_values = baseline_model_val_dict['loss']
val_loss_values = baseline_model_val_dict['val_loss']
```

```
epochs = range(1, len(loss_values) + 1)
ax.plot(epochs, loss_values, label='Training loss')
ax.plot(epochs, val_loss_values, label='Validation loss')

ax.set_title('Training & validation loss')
ax.set_xlabel('Epochs')
ax.set_ylabel('Loss')
ax.legend();
```

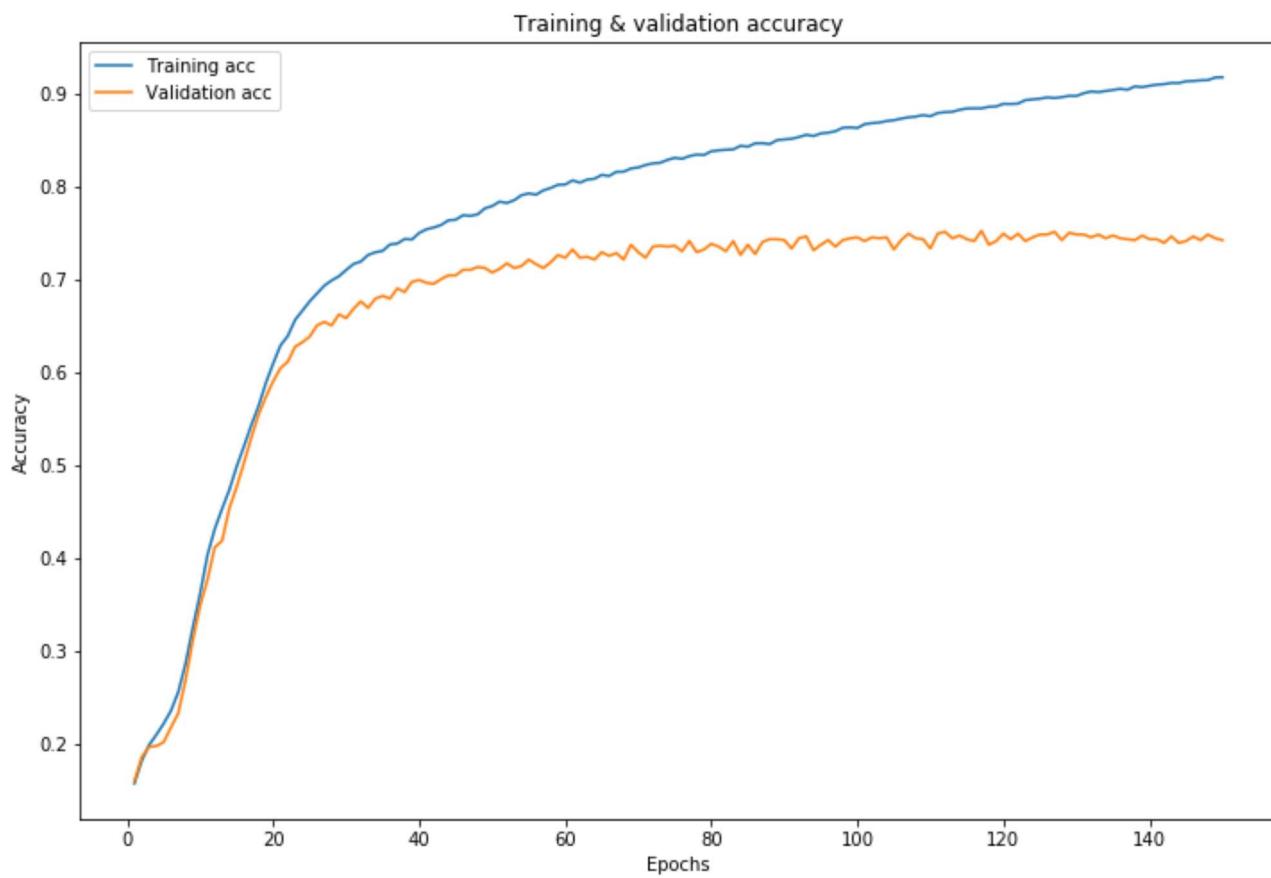


Create a second plot comparing training and validation accuracy to the number of epochs.

```
fig, ax = plt.subplots(figsize=(12, 8))

acc_values = baseline_model_val_dict['acc']
val_acc_values = baseline_model_val_dict['val_acc']

ax.plot(epochs, acc_values, label='Training acc')
ax.plot(epochs, val_acc_values, label='Validation acc')
ax.set_title('Training & validation accuracy')
ax.set_xlabel('Epochs')
ax.set_ylabel('Accuracy')
ax.legend();
```



Did you notice an interesting pattern here? Although the training accuracy keeps increasing when going through more epochs, and the training loss keeps decreasing, the validation accuracy and loss don't necessarily do the same. After a certain point, validation accuracy keeps swinging, which means that you're probably **overfitting** the model to the training data when you train for many epochs past a certain dropoff point. Let's tackle this now. You will now specify an early stopping point when training your model.

## Early Stopping

Overfitting neural networks is something you *want* to avoid at all costs. However, it's not possible to know in advance how many *epochs* you need to train your model on, and running the model multiple times with varying number of *epochs* maybe helpful, but is a time-consuming process.

We've defined a model with the same architecture as above. This time specify an early stopping point when training the model.

```
random.seed(123)
model_2 = models.Sequential()
model_2.add(layers.Dense(50, activation='relu', input_shape=(2000,)))
model_2.add(layers.Dense(25, activation='relu'))
model_2.add(layers.Dense(7, activation='softmax'))
```

```
model_2.compile(optimizer='SGD',
                 loss='categorical_crossentropy',
                 metrics=['acc'])
```

- Import EarlyStopping and ModelCheckpoint from keras.callbacks
- Define a list, early\_stopping :
  - Monitor 'val\_loss' and continue training for 10 epochs before stopping
  - Save the best model while monitoring 'val\_loss'

If you need help, consult [documentation](#).

```
# Import EarlyStopping and ModelCheckpoint
from keras.callbacks import EarlyStopping, ModelCheckpoint

# Define the callbacks
early_stopping = [EarlyStopping(monitor='val_loss', patience=10),
                  ModelCheckpoint(filepath='best_model.h5', monitor='val_loss', save
```

Train model\_2 . Make sure you set the callbacks argument to early\_stopping .

```
model_2_val = model_2.fit(X_train_tokens,
                           y_train_lb,
                           epochs=150,
                           callbacks=early_stopping,
                           batch_size=256,
                           validation_data=(X_val_tokens, y_val_lb))
```

```
Epoch 1/150
30/30 [=====] - 0s 6ms/step - loss: 1.9168 - acc: 0.1969
- val_loss: 1.9202 - val_acc: 0.1790
Epoch 2/150
30/30 [=====] - 0s 3ms/step - loss: 1.8987 - acc: 0.2083
- val_loss: 1.9046 - val_acc: 0.1940
Epoch 3/150
30/30 [=====] - 0s 3ms/step - loss: 1.8801 - acc: 0.2159
- val_loss: 1.8884 - val_acc: 0.1980
Epoch 4/150
30/30 [=====] - 0s 3ms/step - loss: 1.8597 - acc: 0.2259
- val_loss: 1.8704 - val_acc: 0.2060
Epoch 5/150
30/30 [=====] - 0s 3ms/step - loss: 1.8374 - acc: 0.2372
- val_loss: 1.8487 - val_acc: 0.2160
```

```
Epoch 6/150
30/30 [=====] - 0s 3ms/step - loss: 1.8114 - acc: 0.2581
- val_loss: 1.8226 - val_acc: 0.2360
Epoch 7/150
30/30 [=====] - 0s 3ms/step - loss: 1.7802 - acc: 0.2847
- val_loss: 1.7896 - val_acc: 0.2820
Epoch 8/150
30/30 [=====] - 0s 3ms/step - loss: 1.7436 - acc: 0.3241
- val_loss: 1.7548 - val_acc: 0.2930
Epoch 9/150
30/30 [=====] - 0s 3ms/step - loss: 1.7027 - acc: 0.3556
- val_loss: 1.7139 - val_acc: 0.3390
Epoch 10/150
30/30 [=====] - 0s 3ms/step - loss: 1.6574 - acc: 0.3872
- val_loss: 1.6670 - val_acc: 0.3880
Epoch 11/150
30/30 [=====] - 0s 3ms/step - loss: 1.6091 - acc: 0.4239
- val_loss: 1.6201 - val_acc: 0.4020
Epoch 12/150
30/30 [=====] - 0s 3ms/step - loss: 1.5582 - acc: 0.4543
- val_loss: 1.5699 - val_acc: 0.4360
Epoch 13/150
30/30 [=====] - 0s 3ms/step - loss: 1.5060 - acc: 0.4879
- val_loss: 1.5196 - val_acc: 0.4700
Epoch 14/150
30/30 [=====] - 0s 3ms/step - loss: 1.4534 - acc: 0.5175
- val_loss: 1.4679 - val_acc: 0.4960
Epoch 15/150
30/30 [=====] - 0s 3ms/step - loss: 1.4010 - acc: 0.5423
- val_loss: 1.4181 - val_acc: 0.5250
Epoch 16/150
30/30 [=====] - 0s 3ms/step - loss: 1.3495 - acc: 0.5652
- val_loss: 1.3699 - val_acc: 0.5420
Epoch 17/150
30/30 [=====] - 0s 3ms/step - loss: 1.2992 - acc: 0.5881
- val_loss: 1.3220 - val_acc: 0.5580
Epoch 18/150
30/30 [=====] - 0s 3ms/step - loss: 1.2508 - acc: 0.6100
- val_loss: 1.2776 - val_acc: 0.5850
Epoch 19/150
30/30 [=====] - 0s 3ms/step - loss: 1.2040 - acc: 0.6295
- val_loss: 1.2319 - val_acc: 0.5960
Epoch 20/150
30/30 [=====] - 0s 3ms/step - loss: 1.1592 - acc: 0.6419
- val_loss: 1.1893 - val_acc: 0.6070
Epoch 21/150
30/30 [=====] - 0s 3ms/step - loss: 1.1169 - acc: 0.6585
- val_loss: 1.1489 - val_acc: 0.6230
Epoch 22/150
```

```
30/30 [=====] - 0s 3ms/step - loss: 1.0770 - acc: 0.6700
- val_loss: 1.1114 - val_acc: 0.6360
Epoch 23/150
30/30 [=====] - 0s 3ms/step - loss: 1.0394 - acc: 0.6792
- val_loss: 1.0766 - val_acc: 0.6500
Epoch 24/150
30/30 [=====] - 0s 3ms/step - loss: 1.0052 - acc: 0.6889
- val_loss: 1.0474 - val_acc: 0.6550
Epoch 25/150
30/30 [=====] - 0s 3ms/step - loss: 0.9736 - acc: 0.6979
- val_loss: 1.0187 - val_acc: 0.6710
Epoch 26/150
30/30 [=====] - 0s 3ms/step - loss: 0.9446 - acc: 0.7059
- val_loss: 0.9898 - val_acc: 0.6660
Epoch 27/150
30/30 [=====] - 0s 3ms/step - loss: 0.9175 - acc: 0.7113
- val_loss: 0.9636 - val_acc: 0.6720
Epoch 28/150
30/30 [=====] - 0s 3ms/step - loss: 0.8931 - acc: 0.7159
- val_loss: 0.9420 - val_acc: 0.6750
Epoch 29/150
30/30 [=====] - 0s 3ms/step - loss: 0.8704 - acc: 0.7196
- val_loss: 0.9233 - val_acc: 0.6780
Epoch 30/150
30/30 [=====] - 0s 3ms/step - loss: 0.8494 - acc: 0.7272
- val_loss: 0.9026 - val_acc: 0.6840
Epoch 31/150
30/30 [=====] - 0s 3ms/step - loss: 0.8299 - acc: 0.7333
- val_loss: 0.8863 - val_acc: 0.6830
Epoch 32/150
30/30 [=====] - 0s 3ms/step - loss: 0.8121 - acc: 0.7360
- val_loss: 0.8738 - val_acc: 0.6860
Epoch 33/150
30/30 [=====] - 0s 3ms/step - loss: 0.7951 - acc: 0.7416
- val_loss: 0.8577 - val_acc: 0.6900
Epoch 34/150
30/30 [=====] - 0s 3ms/step - loss: 0.7799 - acc: 0.7451
- val_loss: 0.8453 - val_acc: 0.6900
Epoch 35/150
30/30 [=====] - 0s 3ms/step - loss: 0.7651 - acc: 0.7475
- val_loss: 0.8348 - val_acc: 0.6930
Epoch 36/150
30/30 [=====] - 0s 3ms/step - loss: 0.7515 - acc: 0.7519
- val_loss: 0.8242 - val_acc: 0.6870
Epoch 37/150
30/30 [=====] - 0s 3ms/step - loss: 0.7388 - acc: 0.7551
- val_loss: 0.8120 - val_acc: 0.6960
Epoch 38/150
30/30 [=====] - 0s 3ms/step - loss: 0.7269 - acc: 0.7569
```

```
- val_loss: 0.8067 - val_acc: 0.6990
Epoch 39/150
30/30 [=====] - 0s 3ms/step - loss: 0.7153 - acc: 0.7593
- val_loss: 0.7940 - val_acc: 0.6950
Epoch 40/150
30/30 [=====] - 0s 3ms/step - loss: 0.7046 - acc: 0.7628
- val_loss: 0.7866 - val_acc: 0.6940
Epoch 41/150
30/30 [=====] - 0s 3ms/step - loss: 0.6943 - acc: 0.7652
- val_loss: 0.7783 - val_acc: 0.7020
Epoch 42/150
30/30 [=====] - 0s 3ms/step - loss: 0.6845 - acc: 0.7693
- val_loss: 0.7711 - val_acc: 0.7030
Epoch 43/150
30/30 [=====] - 0s 3ms/step - loss: 0.6755 - acc: 0.7715
- val_loss: 0.7632 - val_acc: 0.7040
Epoch 44/150
30/30 [=====] - 0s 3ms/step - loss: 0.6664 - acc: 0.7743
- val_loss: 0.7628 - val_acc: 0.7080
Epoch 45/150
30/30 [=====] - 0s 3ms/step - loss: 0.6579 - acc: 0.7765
- val_loss: 0.7544 - val_acc: 0.7090
Epoch 46/150
30/30 [=====] - 0s 3ms/step - loss: 0.6500 - acc: 0.7785
- val_loss: 0.7469 - val_acc: 0.7090
Epoch 47/150
30/30 [=====] - 0s 3ms/step - loss: 0.6423 - acc: 0.7835
- val_loss: 0.7426 - val_acc: 0.7040
Epoch 48/150
30/30 [=====] - 0s 3ms/step - loss: 0.6346 - acc: 0.7839
- val_loss: 0.7352 - val_acc: 0.7120
Epoch 49/150
30/30 [=====] - 0s 3ms/step - loss: 0.6275 - acc: 0.7861
- val_loss: 0.7311 - val_acc: 0.7080
Epoch 50/150
30/30 [=====] - 0s 3ms/step - loss: 0.6201 - acc: 0.7904
- val_loss: 0.7289 - val_acc: 0.7160
Epoch 51/150
30/30 [=====] - 0s 3ms/step - loss: 0.6132 - acc: 0.7896
- val_loss: 0.7268 - val_acc: 0.7100
Epoch 52/150
30/30 [=====] - 0s 3ms/step - loss: 0.6072 - acc: 0.7933
- val_loss: 0.7210 - val_acc: 0.7160
Epoch 53/150
30/30 [=====] - 0s 3ms/step - loss: 0.6006 - acc: 0.7969
- val_loss: 0.7152 - val_acc: 0.7190
Epoch 54/150
30/30 [=====] - 0s 3ms/step - loss: 0.5943 - acc: 0.7991
- val_loss: 0.7126 - val_acc: 0.7170
```

```
Epoch 55/150
30/30 [=====] - 0s 3ms/step - loss: 0.5884 - acc: 0.7993
- val_loss: 0.7126 - val_acc: 0.7190
Epoch 56/150
30/30 [=====] - 0s 3ms/step - loss: 0.5825 - acc: 0.8017
- val_loss: 0.7059 - val_acc: 0.7250
Epoch 57/150
30/30 [=====] - 0s 3ms/step - loss: 0.5774 - acc: 0.8001
- val_loss: 0.7067 - val_acc: 0.7200
Epoch 58/150
30/30 [=====] - 0s 3ms/step - loss: 0.5716 - acc: 0.8060
- val_loss: 0.7049 - val_acc: 0.7200
Epoch 59/150
30/30 [=====] - 0s 3ms/step - loss: 0.5662 - acc: 0.8061
- val_loss: 0.6985 - val_acc: 0.7220
Epoch 60/150
30/30 [=====] - 0s 3ms/step - loss: 0.5611 - acc: 0.8104
- val_loss: 0.6945 - val_acc: 0.7260
Epoch 61/150
30/30 [=====] - 0s 3ms/step - loss: 0.5561 - acc: 0.8101
- val_loss: 0.6919 - val_acc: 0.7280
Epoch 62/150
30/30 [=====] - 0s 2ms/step - loss: 0.5507 - acc: 0.8107
- val_loss: 0.6920 - val_acc: 0.7270
Epoch 63/150
30/30 [=====] - 0s 3ms/step - loss: 0.5459 - acc: 0.8157
- val_loss: 0.6909 - val_acc: 0.7240
Epoch 64/150
30/30 [=====] - 0s 3ms/step - loss: 0.5415 - acc: 0.8156
- val_loss: 0.6863 - val_acc: 0.7280
Epoch 65/150
30/30 [=====] - 0s 3ms/step - loss: 0.5367 - acc: 0.8176
- val_loss: 0.6830 - val_acc: 0.7280
Epoch 66/150
30/30 [=====] - 0s 3ms/step - loss: 0.5319 - acc: 0.8193
- val_loss: 0.6849 - val_acc: 0.7260
Epoch 67/150
30/30 [=====] - 0s 2ms/step - loss: 0.5276 - acc: 0.8216
- val_loss: 0.6831 - val_acc: 0.7270
Epoch 68/150
30/30 [=====] - 0s 3ms/step - loss: 0.5226 - acc: 0.8229
- val_loss: 0.6810 - val_acc: 0.7280
Epoch 69/150
30/30 [=====] - 0s 3ms/step - loss: 0.5181 - acc: 0.8253
- val_loss: 0.6818 - val_acc: 0.7260
Epoch 70/150
30/30 [=====] - 0s 3ms/step - loss: 0.5145 - acc: 0.8233
- val_loss: 0.6767 - val_acc: 0.7310
Epoch 71/150
```

```
30/30 [=====] - 0s 3ms/step - loss: 0.5099 - acc: 0.8281
- val_loss: 0.6747 - val_acc: 0.7320
Epoch 72/150
30/30 [=====] - 0s 2ms/step - loss: 0.5059 - acc: 0.8288
- val_loss: 0.6752 - val_acc: 0.7310
Epoch 73/150
30/30 [=====] - 0s 3ms/step - loss: 0.5016 - acc: 0.8307
- val_loss: 0.6739 - val_acc: 0.7300
Epoch 74/150
30/30 [=====] - 0s 3ms/step - loss: 0.4981 - acc: 0.8343
- val_loss: 0.6728 - val_acc: 0.7300
Epoch 75/150
30/30 [=====] - 0s 3ms/step - loss: 0.4938 - acc: 0.8344
- val_loss: 0.6753 - val_acc: 0.7280
Epoch 76/150
30/30 [=====] - 0s 2ms/step - loss: 0.4901 - acc: 0.8355
- val_loss: 0.6757 - val_acc: 0.7340
Epoch 77/150
30/30 [=====] - 0s 3ms/step - loss: 0.4862 - acc: 0.8355
- val_loss: 0.6694 - val_acc: 0.7340
Epoch 78/150
30/30 [=====] - 0s 3ms/step - loss: 0.4824 - acc: 0.8379
- val_loss: 0.6700 - val_acc: 0.7390
Epoch 79/150
30/30 [=====] - 0s 3ms/step - loss: 0.4786 - acc: 0.8404
- val_loss: 0.6728 - val_acc: 0.7410
Epoch 80/150
30/30 [=====] - 0s 3ms/step - loss: 0.4749 - acc: 0.8399
- val_loss: 0.6680 - val_acc: 0.7360
Epoch 81/150
30/30 [=====] - 0s 3ms/step - loss: 0.4714 - acc: 0.8415
- val_loss: 0.6647 - val_acc: 0.7330
Epoch 82/150
30/30 [=====] - 0s 3ms/step - loss: 0.4677 - acc: 0.8431
- val_loss: 0.6619 - val_acc: 0.7390
Epoch 83/150
30/30 [=====] - 0s 3ms/step - loss: 0.4639 - acc: 0.8451
- val_loss: 0.6636 - val_acc: 0.7390
Epoch 84/150
30/30 [=====] - 0s 3ms/step - loss: 0.4603 - acc: 0.8468
- val_loss: 0.6667 - val_acc: 0.7420
Epoch 85/150
30/30 [=====] - 0s 2ms/step - loss: 0.4571 - acc: 0.8469
- val_loss: 0.6639 - val_acc: 0.7450
Epoch 86/150
30/30 [=====] - 0s 3ms/step - loss: 0.4536 - acc: 0.8511
- val_loss: 0.6629 - val_acc: 0.7400
Epoch 87/150
30/30 [=====] - 0s 2ms/step - loss: 0.4507 - acc: 0.8515
```

```
- val_loss: 0.6624 - val_acc: 0.7420
Epoch 88/150
30/30 [=====] - 0s 2ms/step - loss: 0.4473 - acc: 0.8503
- val_loss: 0.6628 - val_acc: 0.7430
Epoch 89/150
30/30 [=====] - 0s 3ms/step - loss: 0.4438 - acc: 0.8536
- val_loss: 0.6595 - val_acc: 0.7370
Epoch 90/150
30/30 [=====] - 0s 2ms/step - loss: 0.4404 - acc: 0.8544
- val_loss: 0.6635 - val_acc: 0.7430
Epoch 91/150
30/30 [=====] - 0s 3ms/step - loss: 0.4366 - acc: 0.8557
- val_loss: 0.6617 - val_acc: 0.7400
Epoch 92/150
30/30 [=====] - 0s 2ms/step - loss: 0.4339 - acc: 0.8569
- val_loss: 0.6640 - val_acc: 0.7460
Epoch 93/150
30/30 [=====] - 0s 3ms/step - loss: 0.4305 - acc: 0.8571
- val_loss: 0.6623 - val_acc: 0.7420
Epoch 94/150
30/30 [=====] - 0s 3ms/step - loss: 0.4269 - acc: 0.8608
- val_loss: 0.6597 - val_acc: 0.7410
Epoch 95/150
30/30 [=====] - 0s 3ms/step - loss: 0.4248 - acc: 0.8605
- val_loss: 0.6583 - val_acc: 0.7470
Epoch 96/150
30/30 [=====] - 0s 3ms/step - loss: 0.4213 - acc: 0.8639
- val_loss: 0.6657 - val_acc: 0.7440
Epoch 97/150
30/30 [=====] - 0s 3ms/step - loss: 0.4183 - acc: 0.8635
- val_loss: 0.6622 - val_acc: 0.7460
Epoch 98/150
30/30 [=====] - 0s 2ms/step - loss: 0.4150 - acc: 0.8660
- val_loss: 0.6629 - val_acc: 0.7490
Epoch 99/150
30/30 [=====] - 0s 3ms/step - loss: 0.4122 - acc: 0.8676
- val_loss: 0.6603 - val_acc: 0.7440
Epoch 100/150
30/30 [=====] - 0s 3ms/step - loss: 0.4091 - acc: 0.8664
- val_loss: 0.6606 - val_acc: 0.7410
Epoch 101/150
30/30 [=====] - 0s 2ms/step - loss: 0.4064 - acc: 0.8693
- val_loss: 0.6634 - val_acc: 0.7470
Epoch 102/150
30/30 [=====] - 0s 2ms/step - loss: 0.4031 - acc: 0.8707
- val_loss: 0.6612 - val_acc: 0.7460
Epoch 103/150
30/30 [=====] - 0s 2ms/step - loss: 0.4004 - acc: 0.8712
- val_loss: 0.6612 - val_acc: 0.7440
```

```
Epoch 104/150
30/30 [=====] - 0s 2ms/step - loss: 0.3976 - acc: 0.8720
- val_loss: 0.6606 - val_acc: 0.7430
Epoch 105/150
30/30 [=====] - 0s 3ms/step - loss: 0.3947 - acc: 0.8729
- val_loss: 0.6613 - val_acc: 0.7440
```

```
# Load the best (saved) model
from keras.models import load_model
saved_model = load_model('best_model.h5')
```

Now, use this model to calculate the training and test accuracy:

```
results_train = saved_model.evaluate(X_train_tokens, y_train_lb)
print(f'Training Loss: {results_train[0]:.3} \nTraining Accuracy: {results_train[1]:.3}')
print('-----')

results_test = saved_model.evaluate(X_test_tokens, y_test_lb)
print(f'Test Loss: {results_test[0]:.3} \nTest Accuracy: {results_test[1]:.3}')
```

A screenshot of a Jupyter Notebook cell. The cell contains Python code for evaluating a saved model on training and test datasets. The output shows training and test loss and accuracy values.

```
235/235 [=====] - 0s 599us/step - loss: 0.4198 - acc: 0.8648
Training Loss: 0.42
Training Accuracy: 0.865
-----
47/47 [=====] - 0s 624us/step - loss: 0.6169 - acc: 0.7780
Test Loss: 0.617
Test Accuracy: 0.778
```

Nicely done! Did you notice that the model didn't train for all 150 epochs? You reduced your training time.

Now, take a look at how regularization techniques can further improve your model performance.

## L2 Regularization

First, take a look at L2 regularization. Keras makes L2 regularization easy. Simply add the `kernel_regularizer=keras.regularizers.l2(lambda_coeff)` parameter to any model layer. The `lambda_coeff` parameter determines the strength of the regularization you wish to perform.

- Use 2 hidden layers with 50 units in the first and 25 in the second layer, both with '`relu`' activation functions
- Add L2 regularization to both the hidden layers with 0.005 as the `lambda_coeff`

```
# Import regularizers
from keras import regularizers
random.seed(123)
L2_model = models.Sequential()

# Add the input and first hidden layer
L2_model.add(layers.Dense(50, activation='relu', kernel_regularizer=regularizers.l2(
    0.005))

# Add another hidden layer
L2_model.add(layers.Dense(25, kernel_regularizer=regularizers.l2(0.005), activation='relu'))

# Add an output layer
L2_model.add(layers.Dense(7, activation='softmax'))

# Compile the model
L2_model.compile(optimizer='SGD',
                  loss='categorical_crossentropy',
                  metrics=['acc'])

# Train the model
L2_model_val = L2_model.fit(X_train_tokens,
                            y_train_lb,
                            epochs=150,
                            batch_size=256,
                            validation_data=(X_val_tokens, y_val_lb))
```



```
Epoch 1/150
30/30 [=====] - 0s 6ms/step - loss: 2.6010 - acc: 0.1543
- val_loss: 2.5849 - val_acc: 0.1780
Epoch 2/150
30/30 [=====] - 0s 3ms/step - loss: 2.5697 - acc: 0.1877
- val_loss: 2.5637 - val_acc: 0.2040
Epoch 3/150
```

```
30/30 [=====] - 0s 3ms/step - loss: 2.5452 - acc: 0.2149
- val_loss: 2.5407 - val_acc: 0.2220
Epoch 4/150
30/30 [=====] - 0s 3ms/step - loss: 2.5180 - acc: 0.2404
- val_loss: 2.5131 - val_acc: 0.2450
Epoch 5/150
30/30 [=====] - 0s 3ms/step - loss: 2.4865 - acc: 0.2671
- val_loss: 2.4801 - val_acc: 0.2710
Epoch 6/150
30/30 [=====] - 0s 3ms/step - loss: 2.4502 - acc: 0.3019
- val_loss: 2.4419 - val_acc: 0.3150
Epoch 7/150
30/30 [=====] - 0s 3ms/step - loss: 2.4089 - acc: 0.3372
- val_loss: 2.3990 - val_acc: 0.3450
Epoch 8/150
30/30 [=====] - 0s 3ms/step - loss: 2.3627 - acc: 0.3763
- val_loss: 2.3528 - val_acc: 0.3740
Epoch 9/150
30/30 [=====] - 0s 3ms/step - loss: 2.3127 - acc: 0.4092
- val_loss: 2.3034 - val_acc: 0.4020
Epoch 10/150
30/30 [=====] - 0s 3ms/step - loss: 2.2597 - acc: 0.4481
- val_loss: 2.2495 - val_acc: 0.4460
Epoch 11/150
30/30 [=====] - 0s 3ms/step - loss: 2.2041 - acc: 0.4811
- val_loss: 2.1958 - val_acc: 0.4760
Epoch 12/150
30/30 [=====] - 0s 3ms/step - loss: 2.1474 - acc: 0.5163
- val_loss: 2.1396 - val_acc: 0.5100
Epoch 13/150
30/30 [=====] - 0s 3ms/step - loss: 2.0896 - acc: 0.5471
- val_loss: 2.0849 - val_acc: 0.5370
Epoch 14/150
30/30 [=====] - 0s 3ms/step - loss: 2.0320 - acc: 0.5731
- val_loss: 2.0310 - val_acc: 0.5500
Epoch 15/150
30/30 [=====] - 0s 3ms/step - loss: 1.9751 - acc: 0.5892
- val_loss: 1.9757 - val_acc: 0.5790
Epoch 16/150
30/30 [=====] - 0s 3ms/step - loss: 1.9204 - acc: 0.6136
- val_loss: 1.9252 - val_acc: 0.5910
Epoch 17/150
30/30 [=====] - 0s 3ms/step - loss: 1.8678 - acc: 0.6277
- val_loss: 1.8728 - val_acc: 0.6110
Epoch 18/150
30/30 [=====] - 0s 3ms/step - loss: 1.8182 - acc: 0.6457
- val_loss: 1.8270 - val_acc: 0.6360
Epoch 19/150
30/30 [=====] - 0s 3ms/step - loss: 1.7720 - acc: 0.6583
```

```
- val_loss: 1.7826 - val_acc: 0.6420
Epoch 20/150
30/30 [=====] - 0s 3ms/step - loss: 1.7289 - acc: 0.6741
- val_loss: 1.7428 - val_acc: 0.6480
Epoch 21/150
30/30 [=====] - 0s 3ms/step - loss: 1.6893 - acc: 0.6824
- val_loss: 1.7056 - val_acc: 0.6620
Epoch 22/150
30/30 [=====] - 0s 3ms/step - loss: 1.6519 - acc: 0.6904
- val_loss: 1.6733 - val_acc: 0.6620
Epoch 23/150
30/30 [=====] - 0s 3ms/step - loss: 1.6180 - acc: 0.6973
- val_loss: 1.6391 - val_acc: 0.6680
Epoch 24/150
30/30 [=====] - 0s 3ms/step - loss: 1.5866 - acc: 0.7039
- val_loss: 1.6109 - val_acc: 0.6760
Epoch 25/150
30/30 [=====] - 0s 3ms/step - loss: 1.5570 - acc: 0.7103
- val_loss: 1.5836 - val_acc: 0.6790
Epoch 26/150
30/30 [=====] - 0s 3ms/step - loss: 1.5302 - acc: 0.7191
- val_loss: 1.5595 - val_acc: 0.6850
Epoch 27/150
30/30 [=====] - 0s 3ms/step - loss: 1.5051 - acc: 0.7201
- val_loss: 1.5376 - val_acc: 0.6900
Epoch 28/150
30/30 [=====] - 0s 3ms/step - loss: 1.4820 - acc: 0.7257
- val_loss: 1.5162 - val_acc: 0.6970
Epoch 29/150
30/30 [=====] - 0s 3ms/step - loss: 1.4605 - acc: 0.7296
- val_loss: 1.4978 - val_acc: 0.6990
Epoch 30/150
30/30 [=====] - 0s 3ms/step - loss: 1.4406 - acc: 0.7352
- val_loss: 1.4803 - val_acc: 0.6930
Epoch 31/150
30/30 [=====] - 0s 3ms/step - loss: 1.4219 - acc: 0.7404
- val_loss: 1.4661 - val_acc: 0.6970
Epoch 32/150
30/30 [=====] - 0s 3ms/step - loss: 1.4045 - acc: 0.7412
- val_loss: 1.4471 - val_acc: 0.7040
Epoch 33/150
30/30 [=====] - 0s 3ms/step - loss: 1.3881 - acc: 0.7463
- val_loss: 1.4353 - val_acc: 0.6990
Epoch 34/150
30/30 [=====] - 0s 3ms/step - loss: 1.3727 - acc: 0.7491
- val_loss: 1.4195 - val_acc: 0.7070
Epoch 35/150
30/30 [=====] - 0s 3ms/step - loss: 1.3573 - acc: 0.7528
- val_loss: 1.4066 - val_acc: 0.7060
```

```
Epoch 36/150
30/30 [=====] - 0s 3ms/step - loss: 1.3437 - acc: 0.7553
- val_loss: 1.3956 - val_acc: 0.7110
Epoch 37/150
30/30 [=====] - 0s 3ms/step - loss: 1.3304 - acc: 0.7580
- val_loss: 1.3865 - val_acc: 0.7070
Epoch 38/150
30/30 [=====] - 0s 3ms/step - loss: 1.3180 - acc: 0.7616
- val_loss: 1.3763 - val_acc: 0.7130
Epoch 39/150
30/30 [=====] - 0s 3ms/step - loss: 1.3056 - acc: 0.7643
- val_loss: 1.3692 - val_acc: 0.7060
Epoch 40/150
30/30 [=====] - 0s 3ms/step - loss: 1.2938 - acc: 0.7696
- val_loss: 1.3559 - val_acc: 0.7160
Epoch 41/150
30/30 [=====] - 0s 3ms/step - loss: 1.2827 - acc: 0.7683
- val_loss: 1.3466 - val_acc: 0.7230
Epoch 42/150
30/30 [=====] - 0s 3ms/step - loss: 1.2725 - acc: 0.7712
- val_loss: 1.3391 - val_acc: 0.7160
Epoch 43/150
30/30 [=====] - 0s 3ms/step - loss: 1.2618 - acc: 0.7755
- val_loss: 1.3356 - val_acc: 0.7160
Epoch 44/150
30/30 [=====] - 0s 3ms/step - loss: 1.2523 - acc: 0.7779
- val_loss: 1.3239 - val_acc: 0.7210
Epoch 45/150
30/30 [=====] - 0s 3ms/step - loss: 1.2427 - acc: 0.7803
- val_loss: 1.3167 - val_acc: 0.7250
Epoch 46/150
30/30 [=====] - 0s 3ms/step - loss: 1.2337 - acc: 0.7823
- val_loss: 1.3079 - val_acc: 0.7270
Epoch 47/150
30/30 [=====] - 0s 3ms/step - loss: 1.2241 - acc: 0.7840
- val_loss: 1.3034 - val_acc: 0.7160
Epoch 48/150
30/30 [=====] - 0s 3ms/step - loss: 1.2158 - acc: 0.7857
- val_loss: 1.2952 - val_acc: 0.7280
Epoch 49/150
30/30 [=====] - 0s 3ms/step - loss: 1.2069 - acc: 0.7875
- val_loss: 1.2944 - val_acc: 0.7310
Epoch 50/150
30/30 [=====] - 0s 3ms/step - loss: 1.1989 - acc: 0.7912
- val_loss: 1.2846 - val_acc: 0.7300
Epoch 51/150
30/30 [=====] - 0s 3ms/step - loss: 1.1906 - acc: 0.7925
- val_loss: 1.2787 - val_acc: 0.7270
Epoch 52/150
```

```
30/30 [=====] - 0s 3ms/step - loss: 1.1828 - acc: 0.7932
- val_loss: 1.2733 - val_acc: 0.7240
Epoch 53/150
30/30 [=====] - 0s 3ms/step - loss: 1.1753 - acc: 0.7957
- val_loss: 1.2701 - val_acc: 0.7240
Epoch 54/150
30/30 [=====] - 0s 3ms/step - loss: 1.1677 - acc: 0.7960
- val_loss: 1.2631 - val_acc: 0.7300
Epoch 55/150
30/30 [=====] - 0s 3ms/step - loss: 1.1600 - acc: 0.8009
- val_loss: 1.2599 - val_acc: 0.7290
Epoch 56/150
30/30 [=====] - 0s 3ms/step - loss: 1.1534 - acc: 0.8009
- val_loss: 1.2520 - val_acc: 0.7370
Epoch 57/150
30/30 [=====] - 0s 3ms/step - loss: 1.1460 - acc: 0.8049
- val_loss: 1.2464 - val_acc: 0.7260
Epoch 58/150
30/30 [=====] - 0s 3ms/step - loss: 1.1388 - acc: 0.8037
- val_loss: 1.2419 - val_acc: 0.7350
Epoch 59/150
30/30 [=====] - 0s 3ms/step - loss: 1.1319 - acc: 0.8039
- val_loss: 1.2388 - val_acc: 0.7450
Epoch 60/150
30/30 [=====] - 0s 3ms/step - loss: 1.1255 - acc: 0.8093
- val_loss: 1.2318 - val_acc: 0.7370
Epoch 61/150
30/30 [=====] - 0s 3ms/step - loss: 1.1190 - acc: 0.8105
- val_loss: 1.2347 - val_acc: 0.7270
Epoch 62/150
30/30 [=====] - 0s 2ms/step - loss: 1.1130 - acc: 0.8103
- val_loss: 1.2263 - val_acc: 0.7360
Epoch 63/150
30/30 [=====] - 0s 3ms/step - loss: 1.1063 - acc: 0.8121
- val_loss: 1.2238 - val_acc: 0.7340
Epoch 64/150
30/30 [=====] - 0s 3ms/step - loss: 1.1003 - acc: 0.8139
- val_loss: 1.2166 - val_acc: 0.7400
Epoch 65/150
30/30 [=====] - 0s 3ms/step - loss: 1.0941 - acc: 0.8144
- val_loss: 1.2151 - val_acc: 0.7320
Epoch 66/150
30/30 [=====] - 0s 2ms/step - loss: 1.0881 - acc: 0.8161
- val_loss: 1.2109 - val_acc: 0.7340
Epoch 67/150
30/30 [=====] - 0s 3ms/step - loss: 1.0819 - acc: 0.8183
- val_loss: 1.2041 - val_acc: 0.7440
Epoch 68/150
30/30 [=====] - 0s 3ms/step - loss: 1.0761 - acc: 0.8196
```

```
- val_loss: 1.2043 - val_acc: 0.7430
Epoch 69/150
30/30 [=====] - 0s 3ms/step - loss: 1.0705 - acc: 0.8227
- val_loss: 1.1973 - val_acc: 0.7420
Epoch 70/150
30/30 [=====] - 0s 2ms/step - loss: 1.0649 - acc: 0.8237
- val_loss: 1.1962 - val_acc: 0.7380
Epoch 71/150
30/30 [=====] - 0s 3ms/step - loss: 1.0589 - acc: 0.8240
- val_loss: 1.1923 - val_acc: 0.7480
Epoch 72/150
30/30 [=====] - 0s 2ms/step - loss: 1.0540 - acc: 0.8241
- val_loss: 1.1902 - val_acc: 0.7420
Epoch 73/150
30/30 [=====] - 0s 3ms/step - loss: 1.0485 - acc: 0.8252
- val_loss: 1.1859 - val_acc: 0.7380
Epoch 74/150
30/30 [=====] - 0s 3ms/step - loss: 1.0430 - acc: 0.8284
- val_loss: 1.1826 - val_acc: 0.7410
Epoch 75/150
30/30 [=====] - 0s 3ms/step - loss: 1.0375 - acc: 0.8295
- val_loss: 1.1790 - val_acc: 0.7480
Epoch 76/150
30/30 [=====] - 0s 3ms/step - loss: 1.0324 - acc: 0.8315
- val_loss: 1.1738 - val_acc: 0.7420
Epoch 77/150
30/30 [=====] - 0s 3ms/step - loss: 1.0272 - acc: 0.8323
- val_loss: 1.1743 - val_acc: 0.7450
Epoch 78/150
30/30 [=====] - 0s 3ms/step - loss: 1.0217 - acc: 0.8357
- val_loss: 1.1718 - val_acc: 0.7440
Epoch 79/150
30/30 [=====] - 0s 3ms/step - loss: 1.0169 - acc: 0.8372
- val_loss: 1.1691 - val_acc: 0.7440
Epoch 80/150
30/30 [=====] - 0s 3ms/step - loss: 1.0120 - acc: 0.8361
- val_loss: 1.1640 - val_acc: 0.7440
Epoch 81/150
30/30 [=====] - 0s 2ms/step - loss: 1.0073 - acc: 0.8384
- val_loss: 1.1619 - val_acc: 0.7430
Epoch 82/150
30/30 [=====] - 0s 3ms/step - loss: 1.0022 - acc: 0.8411
- val_loss: 1.1577 - val_acc: 0.7480
Epoch 83/150
30/30 [=====] - 0s 3ms/step - loss: 0.9975 - acc: 0.8388
- val_loss: 1.1568 - val_acc: 0.7410
Epoch 84/150
30/30 [=====] - 0s 3ms/step - loss: 0.9926 - acc: 0.8424
- val_loss: 1.1512 - val_acc: 0.7430
```

```
Epoch 85/150
30/30 [=====] - 0s 3ms/step - loss: 0.9878 - acc: 0.8441
- val_loss: 1.1491 - val_acc: 0.7450
Epoch 86/150
30/30 [=====] - 0s 3ms/step - loss: 0.9830 - acc: 0.8456
- val_loss: 1.1488 - val_acc: 0.7450
Epoch 87/150
30/30 [=====] - 0s 3ms/step - loss: 0.9785 - acc: 0.8451
- val_loss: 1.1463 - val_acc: 0.7450
Epoch 88/150
30/30 [=====] - 0s 3ms/step - loss: 0.9742 - acc: 0.8464
- val_loss: 1.1450 - val_acc: 0.7420
Epoch 89/150
30/30 [=====] - 0s 3ms/step - loss: 0.9697 - acc: 0.8488
- val_loss: 1.1406 - val_acc: 0.7450
Epoch 90/150
30/30 [=====] - 0s 2ms/step - loss: 0.9643 - acc: 0.8503
- val_loss: 1.1375 - val_acc: 0.7520
Epoch 91/150
30/30 [=====] - 0s 2ms/step - loss: 0.9604 - acc: 0.8505
- val_loss: 1.1373 - val_acc: 0.7480
Epoch 92/150
30/30 [=====] - 0s 3ms/step - loss: 0.9560 - acc: 0.8517
- val_loss: 1.1319 - val_acc: 0.7520
Epoch 93/150
30/30 [=====] - 0s 3ms/step - loss: 0.9514 - acc: 0.8525
- val_loss: 1.1315 - val_acc: 0.7490
Epoch 94/150
30/30 [=====] - 0s 2ms/step - loss: 0.9473 - acc: 0.8531
- val_loss: 1.1274 - val_acc: 0.7470
Epoch 95/150
30/30 [=====] - 0s 3ms/step - loss: 0.9433 - acc: 0.8548
- val_loss: 1.1264 - val_acc: 0.7450
Epoch 96/150
30/30 [=====] - 0s 3ms/step - loss: 0.9387 - acc: 0.8564
- val_loss: 1.1249 - val_acc: 0.7510
Epoch 97/150
30/30 [=====] - 0s 2ms/step - loss: 0.9343 - acc: 0.8561
- val_loss: 1.1231 - val_acc: 0.7490
Epoch 98/150
30/30 [=====] - 0s 3ms/step - loss: 0.9303 - acc: 0.8580
- val_loss: 1.1211 - val_acc: 0.7520
Epoch 99/150
30/30 [=====] - 0s 3ms/step - loss: 0.9264 - acc: 0.8589
- val_loss: 1.1184 - val_acc: 0.7480
Epoch 100/150
30/30 [=====] - 0s 3ms/step - loss: 0.9221 - acc: 0.8613
- val_loss: 1.1144 - val_acc: 0.7480
Epoch 101/150
```

```
30/30 [=====] - 0s 3ms/step - loss: 0.9183 - acc: 0.8605
- val_loss: 1.1152 - val_acc: 0.7510
Epoch 102/150
30/30 [=====] - 0s 3ms/step - loss: 0.9140 - acc: 0.8620
- val_loss: 1.1101 - val_acc: 0.7430
Epoch 103/150
30/30 [=====] - 0s 3ms/step - loss: 0.9102 - acc: 0.8635
- val_loss: 1.1119 - val_acc: 0.7470
Epoch 104/150
30/30 [=====] - 0s 3ms/step - loss: 0.9065 - acc: 0.8660
- val_loss: 1.1075 - val_acc: 0.7480
Epoch 105/150
30/30 [=====] - 0s 3ms/step - loss: 0.9022 - acc: 0.8640
- val_loss: 1.1087 - val_acc: 0.7540
Epoch 106/150
30/30 [=====] - 0s 3ms/step - loss: 0.8985 - acc: 0.8656
- val_loss: 1.1066 - val_acc: 0.7540
Epoch 107/150
30/30 [=====] - 0s 2ms/step - loss: 0.8946 - acc: 0.8672
- val_loss: 1.1091 - val_acc: 0.7510
Epoch 108/150
30/30 [=====] - 0s 3ms/step - loss: 0.8909 - acc: 0.8675
- val_loss: 1.1002 - val_acc: 0.7520
Epoch 109/150
30/30 [=====] - 0s 3ms/step - loss: 0.8871 - acc: 0.8680
- val_loss: 1.0991 - val_acc: 0.7460
Epoch 110/150
30/30 [=====] - 0s 3ms/step - loss: 0.8836 - acc: 0.8693
- val_loss: 1.0979 - val_acc: 0.7540
Epoch 111/150
30/30 [=====] - 0s 3ms/step - loss: 0.8798 - acc: 0.8697
- val_loss: 1.0958 - val_acc: 0.7500
Epoch 112/150
30/30 [=====] - 0s 3ms/step - loss: 0.8760 - acc: 0.8699
- val_loss: 1.0956 - val_acc: 0.7500
Epoch 113/150
30/30 [=====] - 0s 3ms/step - loss: 0.8727 - acc: 0.8707
- val_loss: 1.0937 - val_acc: 0.7510
Epoch 114/150
30/30 [=====] - 0s 3ms/step - loss: 0.8686 - acc: 0.8737
- val_loss: 1.0897 - val_acc: 0.7540
Epoch 115/150
30/30 [=====] - 0s 2ms/step - loss: 0.8648 - acc: 0.8727
- val_loss: 1.0890 - val_acc: 0.7460
Epoch 116/150
30/30 [=====] - 0s 3ms/step - loss: 0.8619 - acc: 0.8736
- val_loss: 1.0895 - val_acc: 0.7530
Epoch 117/150
30/30 [=====] - 0s 3ms/step - loss: 0.8579 - acc: 0.8744
```

```
- val_loss: 1.0844 - val_acc: 0.7520
Epoch 118/150
30/30 [=====] - 0s 3ms/step - loss: 0.8546 - acc: 0.8771
- val_loss: 1.0830 - val_acc: 0.7540
Epoch 119/150
30/30 [=====] - 0s 2ms/step - loss: 0.8512 - acc: 0.8755
- val_loss: 1.0799 - val_acc: 0.7510
Epoch 120/150
30/30 [=====] - 0s 2ms/step - loss: 0.8483 - acc: 0.8772
- val_loss: 1.0776 - val_acc: 0.7540
Epoch 121/150
30/30 [=====] - 0s 3ms/step - loss: 0.8443 - acc: 0.8777
- val_loss: 1.0783 - val_acc: 0.7550
Epoch 122/150
30/30 [=====] - 0s 3ms/step - loss: 0.8408 - acc: 0.8799
- val_loss: 1.0746 - val_acc: 0.7600
Epoch 123/150
30/30 [=====] - 0s 3ms/step - loss: 0.8376 - acc: 0.8805
- val_loss: 1.0721 - val_acc: 0.7560
Epoch 124/150
30/30 [=====] - 0s 3ms/step - loss: 0.8345 - acc: 0.8805
- val_loss: 1.0729 - val_acc: 0.7570
Epoch 125/150
30/30 [=====] - 0s 3ms/step - loss: 0.8312 - acc: 0.8819
- val_loss: 1.0709 - val_acc: 0.7520
Epoch 126/150
30/30 [=====] - 0s 3ms/step - loss: 0.8275 - acc: 0.8839
- val_loss: 1.0690 - val_acc: 0.7540
Epoch 127/150
30/30 [=====] - 0s 3ms/step - loss: 0.8242 - acc: 0.8848
- val_loss: 1.0711 - val_acc: 0.7550
Epoch 128/150
30/30 [=====] - 0s 3ms/step - loss: 0.8215 - acc: 0.8840
- val_loss: 1.0695 - val_acc: 0.7530
Epoch 129/150
30/30 [=====] - 0s 3ms/step - loss: 0.8180 - acc: 0.8855
- val_loss: 1.0655 - val_acc: 0.7530
Epoch 130/150
30/30 [=====] - 0s 2ms/step - loss: 0.8149 - acc: 0.8880
- val_loss: 1.0650 - val_acc: 0.7550
Epoch 131/150
30/30 [=====] - 0s 3ms/step - loss: 0.8120 - acc: 0.8887
- val_loss: 1.0644 - val_acc: 0.7540
Epoch 132/150
30/30 [=====] - 0s 3ms/step - loss: 0.8085 - acc: 0.8887
- val_loss: 1.0637 - val_acc: 0.7540
Epoch 133/150
30/30 [=====] - 0s 3ms/step - loss: 0.8057 - acc: 0.8888
- val_loss: 1.0595 - val_acc: 0.7570
```

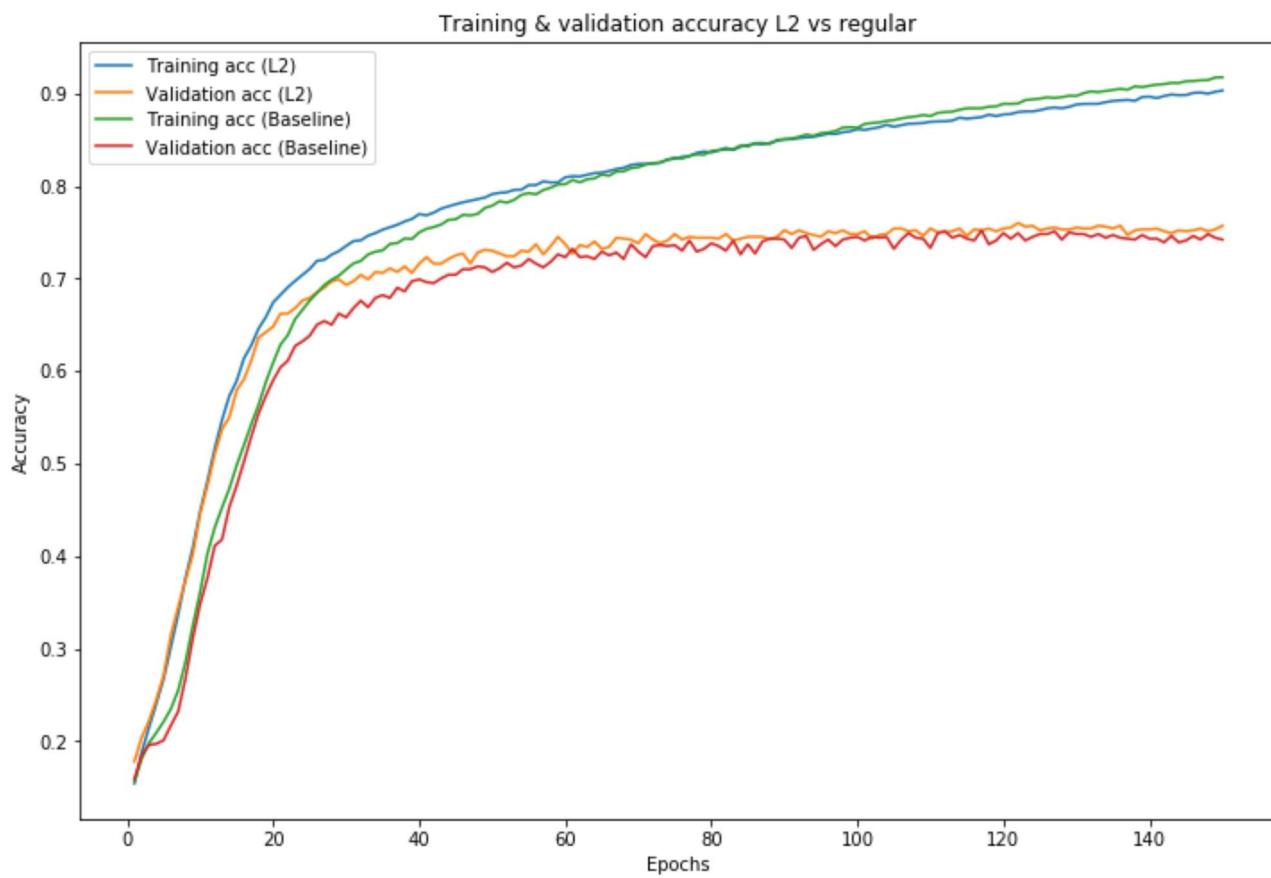
```
Epoch 134/150
30/30 [=====] - 0s 3ms/step - loss: 0.8022 - acc: 0.8905
- val_loss: 1.0587 - val_acc: 0.7560
Epoch 135/150
30/30 [=====] - 0s 2ms/step - loss: 0.7993 - acc: 0.8919
- val_loss: 1.0641 - val_acc: 0.7540
Epoch 136/150
30/30 [=====] - 0s 3ms/step - loss: 0.7963 - acc: 0.8924
- val_loss: 1.0567 - val_acc: 0.7570
Epoch 137/150
30/30 [=====] - 0s 3ms/step - loss: 0.7931 - acc: 0.8929
- val_loss: 1.0542 - val_acc: 0.7470
Epoch 138/150
30/30 [=====] - 0s 2ms/step - loss: 0.7905 - acc: 0.8920
- val_loss: 1.0544 - val_acc: 0.7520
Epoch 139/150
30/30 [=====] - 0s 2ms/step - loss: 0.7872 - acc: 0.8961
- val_loss: 1.0516 - val_acc: 0.7530
Epoch 140/150
30/30 [=====] - 0s 3ms/step - loss: 0.7847 - acc: 0.8965
- val_loss: 1.0505 - val_acc: 0.7530
Epoch 141/150
30/30 [=====] - 0s 3ms/step - loss: 0.7813 - acc: 0.8949
- val_loss: 1.0491 - val_acc: 0.7540
Epoch 142/150
30/30 [=====] - 0s 3ms/step - loss: 0.7789 - acc: 0.8972
- val_loss: 1.0475 - val_acc: 0.7510
Epoch 143/150
30/30 [=====] - 0s 3ms/step - loss: 0.7756 - acc: 0.8988
- val_loss: 1.0484 - val_acc: 0.7490
Epoch 144/150
30/30 [=====] - 0s 3ms/step - loss: 0.7732 - acc: 0.8981
- val_loss: 1.0464 - val_acc: 0.7520
Epoch 145/150
30/30 [=====] - 0s 3ms/step - loss: 0.7699 - acc: 0.8983
- val_loss: 1.0455 - val_acc: 0.7510
Epoch 146/150
30/30 [=====] - 0s 3ms/step - loss: 0.7671 - acc: 0.9005
- val_loss: 1.0459 - val_acc: 0.7520
Epoch 147/150
30/30 [=====] - 0s 2ms/step - loss: 0.7646 - acc: 0.9009
- val_loss: 1.0431 - val_acc: 0.7540
Epoch 148/150
30/30 [=====] - 0s 3ms/step - loss: 0.7618 - acc: 0.8997
- val_loss: 1.0408 - val_acc: 0.7510
Epoch 149/150
30/30 [=====] - 0s 3ms/step - loss: 0.7590 - acc: 0.9017
- val_loss: 1.0421 - val_acc: 0.7530
Epoch 150/150
```

```
30/30 [=====] - 0s 3ms/step - loss: 0.7565 - acc: 0.9031
- val_loss: 1.0395 - val_acc: 0.7570
```

```
# L2 model details
L2_model_dict = L2_model_val.history
L2_acc_values = L2_model_dict['acc']
L2_val_acc_values = L2_model_dict['val_acc']

# Baseline model
baseline_model_acc = baseline_model_val_dict['acc']
baseline_model_val_acc = baseline_model_val_dict['val_acc']

# Plot the accuracy for these models
fig, ax = plt.subplots(figsize=(12, 8))
epochs = range(1, len(acc_values) + 1)
ax.plot(epochs, L2_acc_values, label='Training acc (L2)')
ax.plot(epochs, L2_val_acc_values, label='Validation acc (L2)')
ax.plot(epochs, baseline_model_acc, label='Training acc (Baseline)')
ax.plot(epochs, baseline_model_val_acc, label='Validation acc (Baseline)')
ax.set_title('Training & validation accuracy L2 vs regular')
ax.set_xlabel('Epochs')
ax.set_ylabel('Accuracy')
ax.legend();
```



The results of L2 regularization are quite disappointing here. Notice the discrepancy between validation and training accuracy seems to have decreased slightly, but the end result is definitely not getting better.

## L1 Regularization

Now have a look at L1 regularization. Will this work better?

- Use 2 hidden layers with 50 units in the first and 25 in the second layer, both with 'relu' activation functions
- Add L1 regularization to both the hidden layers with 0.005 as the `lambda_coeff`

```
random.seed(123)
L1_model = models.Sequential()

# Add the input and first hidden layer
L1_model.add(layers.Dense(50, activation='relu', kernel_regularizer=regularizers.l1(
    0.005))

# Add a hidden layer
L1_model.add(layers.Dense(25, kernel_regularizer=regularizers.l1(0.005), activation='relu'))

# Add an output layer
L1_model.add(layers.Dense(7, activation='softmax'))
```

```
# Compile the model
L1_model.compile(optimizer='SGD',
                  loss='categorical_crossentropy',
                  metrics=['acc'])

# Train the model
L1_model_val = L1_model.fit(X_train_tokens,
                             y_train_lb,
                             epochs=150,
                             batch_size=256,
                             validation_data=(X_val_tokens, y_val_lb))

Epoch 1/150
30/30 [=====] - 0s 6ms/step - loss: 15.9847 - acc: 0.1743 - val_loss: 15.5659 - val_acc: 0.2080
Epoch 2/150
30/30 [=====] - 0s 3ms/step - loss: 15.2170 - acc: 0.2164 - val_loss: 14.8178 - val_acc: 0.2460
Epoch 3/150
30/30 [=====] - 0s 3ms/step - loss: 14.4777 - acc: 0.2517 - val_loss: 14.0937 - val_acc: 0.2660
Epoch 4/150
30/30 [=====] - 0s 3ms/step - loss: 13.7607 - acc: 0.2805 - val_loss: 13.3905 - val_acc: 0.2740
Epoch 5/150
30/30 [=====] - 0s 3ms/step - loss: 13.0644 - acc: 0.2973 - val_loss: 12.7083 - val_acc: 0.2840
Epoch 6/150
30/30 [=====] - 0s 3ms/step - loss: 12.3885 - acc: 0.3145 - val_loss: 12.0453 - val_acc: 0.2940
Epoch 7/150
30/30 [=====] - 0s 3ms/step - loss: 11.7320 - acc: 0.3292 - val_loss: 11.4015 - val_acc: 0.3050
Epoch 8/150
30/30 [=====] - 0s 3ms/step - loss: 11.0952 - acc: 0.3517 - val_loss: 10.7770 - val_acc: 0.3310
Epoch 9/150
30/30 [=====] - 0s 3ms/step - loss: 10.4767 - acc: 0.3805 - val_loss: 10.1712 - val_acc: 0.3540
Epoch 10/150
30/30 [=====] - 0s 3ms/step - loss: 9.8774 - acc: 0.4139 - val_loss: 9.5847 - val_acc: 0.3880
Epoch 11/150
30/30 [=====] - 0s 3ms/step - loss: 9.2985 - acc: 0.4519 - val_loss: 9.0198 - val_acc: 0.4160
Epoch 12/150
```

```
30/30 [=====] - 0s 3ms/step - loss: 8.7408 - acc: 0.4749
- val_loss: 8.4760 - val_acc: 0.4280
Epoch 13/150
30/30 [=====] - 0s 3ms/step - loss: 8.2059 - acc: 0.4888
- val_loss: 7.9545 - val_acc: 0.4550
Epoch 14/150
30/30 [=====] - 0s 3ms/step - loss: 7.6937 - acc: 0.5175
- val_loss: 7.4554 - val_acc: 0.4790
Epoch 15/150
30/30 [=====] - 0s 3ms/step - loss: 7.2039 - acc: 0.5337
- val_loss: 6.9780 - val_acc: 0.5160
Epoch 16/150
30/30 [=====] - 0s 3ms/step - loss: 6.7368 - acc: 0.5599
- val_loss: 6.5251 - val_acc: 0.5210
Epoch 17/150
30/30 [=====] - 0s 3ms/step - loss: 6.2928 - acc: 0.5760
- val_loss: 6.0927 - val_acc: 0.5430
Epoch 18/150
30/30 [=====] - 0s 3ms/step - loss: 5.8716 - acc: 0.5919
- val_loss: 5.6853 - val_acc: 0.5550
Epoch 19/150
30/30 [=====] - 0s 3ms/step - loss: 5.4739 - acc: 0.6083
- val_loss: 5.3001 - val_acc: 0.5760
Epoch 20/150
30/30 [=====] - 0s 3ms/step - loss: 5.0996 - acc: 0.6207
- val_loss: 4.9393 - val_acc: 0.5870
Epoch 21/150
30/30 [=====] - 0s 3ms/step - loss: 4.7492 - acc: 0.6296
- val_loss: 4.6000 - val_acc: 0.6090
Epoch 22/150
30/30 [=====] - 0s 3ms/step - loss: 4.4211 - acc: 0.6415
- val_loss: 4.2846 - val_acc: 0.6150
Epoch 23/150
30/30 [=====] - 0s 3ms/step - loss: 4.1159 - acc: 0.6464
- val_loss: 3.9925 - val_acc: 0.6290
Epoch 24/150
30/30 [=====] - 0s 3ms/step - loss: 3.8333 - acc: 0.6536
- val_loss: 3.7203 - val_acc: 0.6440
Epoch 25/150
30/30 [=====] - 0s 3ms/step - loss: 3.5738 - acc: 0.6607
- val_loss: 3.4736 - val_acc: 0.6380
Epoch 26/150
30/30 [=====] - 0s 3ms/step - loss: 3.3356 - acc: 0.6660
- val_loss: 3.2464 - val_acc: 0.6410
Epoch 27/150
30/30 [=====] - 0s 3ms/step - loss: 3.1185 - acc: 0.6665
- val_loss: 3.0400 - val_acc: 0.6560
Epoch 28/150
30/30 [=====] - 0s 3ms/step - loss: 2.9227 - acc: 0.6677
```

```
- val_loss: 2.8532 - val_acc: 0.6630
Epoch 29/150
30/30 [=====] - ETA: 0s - loss: 2.7597 - acc: 0.675 - 0s
3ms/step - loss: 2.7480 - acc: 0.6719 - val_loss: 2.6889 - val_acc: 0.6620
Epoch 30/150
30/30 [=====] - 0s 3ms/step - loss: 2.5944 - acc: 0.6724
- val_loss: 2.5462 - val_acc: 0.6590
Epoch 31/150
30/30 [=====] - 0s 3ms/step - loss: 2.4610 - acc: 0.6720
- val_loss: 2.4222 - val_acc: 0.6620
Epoch 32/150
30/30 [=====] - 0s 3ms/step - loss: 2.3476 - acc: 0.6761
- val_loss: 2.3183 - val_acc: 0.6600
Epoch 33/150
30/30 [=====] - 0s 3ms/step - loss: 2.2537 - acc: 0.6768
- val_loss: 2.2328 - val_acc: 0.6640
Epoch 34/150
30/30 [=====] - 0s 3ms/step - loss: 2.1777 - acc: 0.6772
- val_loss: 2.1644 - val_acc: 0.6730
Epoch 35/150
30/30 [=====] - 0s 3ms/step - loss: 2.1188 - acc: 0.6807
- val_loss: 2.1118 - val_acc: 0.6750
Epoch 36/150
30/30 [=====] - 0s 3ms/step - loss: 2.0744 - acc: 0.6825
- val_loss: 2.0740 - val_acc: 0.6750
Epoch 37/150
30/30 [=====] - 0s 3ms/step - loss: 2.0410 - acc: 0.6815
- val_loss: 2.0429 - val_acc: 0.6750
Epoch 38/150
30/30 [=====] - 0s 3ms/step - loss: 2.0136 - acc: 0.6840
- val_loss: 2.0163 - val_acc: 0.6760
Epoch 39/150
30/30 [=====] - 0s 3ms/step - loss: 1.9900 - acc: 0.6835
- val_loss: 1.9942 - val_acc: 0.6740
Epoch 40/150
30/30 [=====] - 0s 3ms/step - loss: 1.9685 - acc: 0.6840
- val_loss: 1.9748 - val_acc: 0.6710
Epoch 41/150
30/30 [=====] - 0s 2ms/step - loss: 1.9489 - acc: 0.6859
- val_loss: 1.9545 - val_acc: 0.6730
Epoch 42/150
30/30 [=====] - 0s 3ms/step - loss: 1.9302 - acc: 0.6867
- val_loss: 1.9349 - val_acc: 0.6780
Epoch 43/150
30/30 [=====] - 0s 3ms/step - loss: 1.9133 - acc: 0.6879
- val_loss: 1.9171 - val_acc: 0.6820
Epoch 44/150
30/30 [=====] - 0s 3ms/step - loss: 1.8964 - acc: 0.6887
- val_loss: 1.9065 - val_acc: 0.6730
```

```
Epoch 45/150
30/30 [=====] - 0s 3ms/step - loss: 1.8811 - acc: 0.6875
- val_loss: 1.8892 - val_acc: 0.6740
Epoch 46/150
30/30 [=====] - 0s 3ms/step - loss: 1.8662 - acc: 0.6884
- val_loss: 1.8740 - val_acc: 0.6780
Epoch 47/150
30/30 [=====] - 0s 3ms/step - loss: 1.8519 - acc: 0.6888
- val_loss: 1.8586 - val_acc: 0.6730
Epoch 48/150
30/30 [=====] - 0s 3ms/step - loss: 1.8374 - acc: 0.6917
- val_loss: 1.8444 - val_acc: 0.6770
Epoch 49/150
30/30 [=====] - 0s 3ms/step - loss: 1.8241 - acc: 0.6913
- val_loss: 1.8293 - val_acc: 0.6810
Epoch 50/150
30/30 [=====] - 0s 3ms/step - loss: 1.8111 - acc: 0.6932
- val_loss: 1.8194 - val_acc: 0.6730
Epoch 51/150
30/30 [=====] - 0s 3ms/step - loss: 1.7985 - acc: 0.6912
- val_loss: 1.8056 - val_acc: 0.6790
Epoch 52/150
30/30 [=====] - 0s 3ms/step - loss: 1.7870 - acc: 0.6927
- val_loss: 1.7931 - val_acc: 0.6760
Epoch 53/150
30/30 [=====] - 0s 3ms/step - loss: 1.7751 - acc: 0.6927
- val_loss: 1.7800 - val_acc: 0.6810
Epoch 54/150
30/30 [=====] - 0s 3ms/step - loss: 1.7635 - acc: 0.6943
- val_loss: 1.7714 - val_acc: 0.6820
Epoch 55/150
30/30 [=====] - 0s 3ms/step - loss: 1.7520 - acc: 0.6949
- val_loss: 1.7595 - val_acc: 0.6810
Epoch 56/150
30/30 [=====] - 0s 3ms/step - loss: 1.7414 - acc: 0.6985
- val_loss: 1.7498 - val_acc: 0.6790
Epoch 57/150
30/30 [=====] - 0s 3ms/step - loss: 1.7306 - acc: 0.6957
- val_loss: 1.7388 - val_acc: 0.6830
Epoch 58/150
30/30 [=====] - 0s 3ms/step - loss: 1.7204 - acc: 0.6989
- val_loss: 1.7337 - val_acc: 0.6780
Epoch 59/150
30/30 [=====] - 0s 3ms/step - loss: 1.7107 - acc: 0.6992
- val_loss: 1.7205 - val_acc: 0.6800
Epoch 60/150
30/30 [=====] - 0s 3ms/step - loss: 1.7006 - acc: 0.6999
- val_loss: 1.7090 - val_acc: 0.6840
Epoch 61/150
```

```
30/30 [=====] - 0s 3ms/step - loss: 1.6907 - acc: 0.6999
- val_loss: 1.6974 - val_acc: 0.6800
Epoch 62/150
30/30 [=====] - 0s 3ms/step - loss: 1.6809 - acc: 0.7012
- val_loss: 1.6877 - val_acc: 0.6890
Epoch 63/150
30/30 [=====] - 0s 3ms/step - loss: 1.6718 - acc: 0.7016
- val_loss: 1.6808 - val_acc: 0.6840
Epoch 64/150
30/30 [=====] - 0s 2ms/step - loss: 1.6623 - acc: 0.7029
- val_loss: 1.6704 - val_acc: 0.6850
Epoch 65/150
30/30 [=====] - 0s 3ms/step - loss: 1.6532 - acc: 0.7028
- val_loss: 1.6614 - val_acc: 0.6880
Epoch 66/150
30/30 [=====] - 0s 3ms/step - loss: 1.6447 - acc: 0.7028
- val_loss: 1.6528 - val_acc: 0.6860
Epoch 67/150
30/30 [=====] - 0s 3ms/step - loss: 1.6354 - acc: 0.7043
- val_loss: 1.6431 - val_acc: 0.6890
Epoch 68/150
30/30 [=====] - 0s 3ms/step - loss: 1.6268 - acc: 0.7048
- val_loss: 1.6345 - val_acc: 0.6890
Epoch 69/150
30/30 [=====] - 0s 3ms/step - loss: 1.6183 - acc: 0.7061
- val_loss: 1.6278 - val_acc: 0.6880
Epoch 70/150
30/30 [=====] - 0s 3ms/step - loss: 1.6098 - acc: 0.7056
- val_loss: 1.6226 - val_acc: 0.6860
Epoch 71/150
30/30 [=====] - 0s 2ms/step - loss: 1.6019 - acc: 0.7061
- val_loss: 1.6158 - val_acc: 0.6820
Epoch 72/150
30/30 [=====] - 0s 3ms/step - loss: 1.5936 - acc: 0.7061
- val_loss: 1.6071 - val_acc: 0.6820
Epoch 73/150
30/30 [=====] - 0s 3ms/step - loss: 1.5856 - acc: 0.7073
- val_loss: 1.5992 - val_acc: 0.6880
Epoch 74/150
30/30 [=====] - 0s 3ms/step - loss: 1.5781 - acc: 0.7069
- val_loss: 1.5875 - val_acc: 0.6880
Epoch 75/150
30/30 [=====] - 0s 3ms/step - loss: 1.5694 - acc: 0.7063
- val_loss: 1.5819 - val_acc: 0.6890
Epoch 76/150
30/30 [=====] - 0s 3ms/step - loss: 1.5619 - acc: 0.7097
- val_loss: 1.5740 - val_acc: 0.6900
Epoch 77/150
30/30 [=====] - 0s 2ms/step - loss: 1.5541 - acc: 0.7092
```

```
- val_loss: 1.5656 - val_acc: 0.6890
Epoch 78/150
30/30 [=====] - 0s 3ms/step - loss: 1.5459 - acc: 0.7099
- val_loss: 1.5578 - val_acc: 0.6950
Epoch 79/150
30/30 [=====] - 0s 2ms/step - loss: 1.5385 - acc: 0.7119
- val_loss: 1.5514 - val_acc: 0.6900
Epoch 80/150
30/30 [=====] - 0s 3ms/step - loss: 1.5314 - acc: 0.7117
- val_loss: 1.5434 - val_acc: 0.6910
Epoch 81/150
30/30 [=====] - 0s 2ms/step - loss: 1.5238 - acc: 0.7129
- val_loss: 1.5430 - val_acc: 0.6820
Epoch 82/150
30/30 [=====] - 0s 2ms/step - loss: 1.5173 - acc: 0.7123
- val_loss: 1.5280 - val_acc: 0.6910
Epoch 83/150
30/30 [=====] - 0s 3ms/step - loss: 1.5097 - acc: 0.7120
- val_loss: 1.5309 - val_acc: 0.6830
Epoch 84/150
30/30 [=====] - 0s 3ms/step - loss: 1.5031 - acc: 0.7125
- val_loss: 1.5168 - val_acc: 0.6900
Epoch 85/150
30/30 [=====] - 0s 3ms/step - loss: 1.4960 - acc: 0.7145
- val_loss: 1.5104 - val_acc: 0.6890
Epoch 86/150
30/30 [=====] - 0s 3ms/step - loss: 1.4889 - acc: 0.7149
- val_loss: 1.5016 - val_acc: 0.6890
Epoch 87/150
30/30 [=====] - 0s 3ms/step - loss: 1.4823 - acc: 0.7143
- val_loss: 1.4972 - val_acc: 0.6980
Epoch 88/150
30/30 [=====] - 0s 3ms/step - loss: 1.4757 - acc: 0.7149
- val_loss: 1.4887 - val_acc: 0.6930
Epoch 89/150
30/30 [=====] - 0s 3ms/step - loss: 1.4690 - acc: 0.7165
- val_loss: 1.4832 - val_acc: 0.6900
Epoch 90/150
30/30 [=====] - 0s 3ms/step - loss: 1.4629 - acc: 0.7143
- val_loss: 1.4751 - val_acc: 0.6930
Epoch 91/150
30/30 [=====] - 0s 3ms/step - loss: 1.4562 - acc: 0.7164
- val_loss: 1.4740 - val_acc: 0.6910
Epoch 92/150
30/30 [=====] - 0s 3ms/step - loss: 1.4500 - acc: 0.7171
- val_loss: 1.4659 - val_acc: 0.6920
Epoch 93/150
30/30 [=====] - 0s 3ms/step - loss: 1.4436 - acc: 0.7188
- val_loss: 1.4580 - val_acc: 0.6900
```

```
Epoch 94/150
30/30 [=====] - 0s 3ms/step - loss: 1.4375 - acc: 0.7172
- val_loss: 1.4515 - val_acc: 0.6960
Epoch 95/150
30/30 [=====] - 0s 3ms/step - loss: 1.4312 - acc: 0.7189
- val_loss: 1.4451 - val_acc: 0.6950
Epoch 96/150
30/30 [=====] - 0s 3ms/step - loss: 1.4249 - acc: 0.7185
- val_loss: 1.4394 - val_acc: 0.6940
Epoch 97/150
30/30 [=====] - 0s 2ms/step - loss: 1.4189 - acc: 0.7183
- val_loss: 1.4329 - val_acc: 0.6970
Epoch 98/150
30/30 [=====] - 0s 3ms/step - loss: 1.4131 - acc: 0.7188
- val_loss: 1.4268 - val_acc: 0.6960
Epoch 99/150
30/30 [=====] - 0s 3ms/step - loss: 1.4071 - acc: 0.7192
- val_loss: 1.4248 - val_acc: 0.6960
Epoch 100/150
30/30 [=====] - 0s 3ms/step - loss: 1.4016 - acc: 0.7197
- val_loss: 1.4172 - val_acc: 0.7000
Epoch 101/150
30/30 [=====] - 0s 3ms/step - loss: 1.3957 - acc: 0.7179
- val_loss: 1.4134 - val_acc: 0.6910
Epoch 102/150
30/30 [=====] - 0s 3ms/step - loss: 1.3907 - acc: 0.7205
- val_loss: 1.4075 - val_acc: 0.6970
Epoch 103/150
30/30 [=====] - 0s 3ms/step - loss: 1.3847 - acc: 0.7216
- val_loss: 1.4021 - val_acc: 0.6920
Epoch 104/150
30/30 [=====] - 0s 3ms/step - loss: 1.3793 - acc: 0.7215
- val_loss: 1.3962 - val_acc: 0.6980
Epoch 105/150
30/30 [=====] - 0s 3ms/step - loss: 1.3729 - acc: 0.7213
- val_loss: 1.3869 - val_acc: 0.7000
Epoch 106/150
30/30 [=====] - 0s 3ms/step - loss: 1.3678 - acc: 0.7209
- val_loss: 1.3854 - val_acc: 0.6950
Epoch 107/150
30/30 [=====] - 0s 3ms/step - loss: 1.3619 - acc: 0.7208
- val_loss: 1.3777 - val_acc: 0.7010
Epoch 108/150
30/30 [=====] - 0s 3ms/step - loss: 1.3565 - acc: 0.7223
- val_loss: 1.3759 - val_acc: 0.6990
Epoch 109/150
30/30 [=====] - 0s 3ms/step - loss: 1.3516 - acc: 0.7236
- val_loss: 1.3692 - val_acc: 0.7010
Epoch 110/150
```

```
30/30 [=====] - 0s 3ms/step - loss: 1.3464 - acc: 0.7243
- val_loss: 1.3654 - val_acc: 0.6970
Epoch 111/150
30/30 [=====] - 0s 3ms/step - loss: 1.3412 - acc: 0.7233
- val_loss: 1.3614 - val_acc: 0.7010
Epoch 112/150
30/30 [=====] - 0s 3ms/step - loss: 1.3360 - acc: 0.7244
- val_loss: 1.3515 - val_acc: 0.7020
Epoch 113/150
30/30 [=====] - 0s 3ms/step - loss: 1.3303 - acc: 0.7235
- val_loss: 1.3482 - val_acc: 0.7020
Epoch 114/150
30/30 [=====] - 0s 2ms/step - loss: 1.3256 - acc: 0.7261
- val_loss: 1.3435 - val_acc: 0.7010
Epoch 115/150
30/30 [=====] - 0s 3ms/step - loss: 1.3205 - acc: 0.7239
- val_loss: 1.3425 - val_acc: 0.7000
Epoch 116/150
30/30 [=====] - 0s 3ms/step - loss: 1.3157 - acc: 0.7259
- val_loss: 1.3369 - val_acc: 0.6980
Epoch 117/150
30/30 [=====] - 0s 3ms/step - loss: 1.3109 - acc: 0.7255
- val_loss: 1.3283 - val_acc: 0.7000
Epoch 118/150
30/30 [=====] - 0s 3ms/step - loss: 1.3059 - acc: 0.7259
- val_loss: 1.3231 - val_acc: 0.6990
Epoch 119/150
30/30 [=====] - 0s 2ms/step - loss: 1.3010 - acc: 0.7277
- val_loss: 1.3260 - val_acc: 0.7030
Epoch 120/150
30/30 [=====] - 0s 3ms/step - loss: 1.2974 - acc: 0.7261
- val_loss: 1.3151 - val_acc: 0.6990
Epoch 121/150
30/30 [=====] - 0s 3ms/step - loss: 1.2921 - acc: 0.7263
- val_loss: 1.3103 - val_acc: 0.7020
Epoch 122/150
30/30 [=====] - 0s 3ms/step - loss: 1.2871 - acc: 0.7261
- val_loss: 1.3067 - val_acc: 0.7010
Epoch 123/150
30/30 [=====] - 0s 3ms/step - loss: 1.2829 - acc: 0.7267
- val_loss: 1.3010 - val_acc: 0.7010
Epoch 124/150
30/30 [=====] - 0s 2ms/step - loss: 1.2784 - acc: 0.7271
- val_loss: 1.2994 - val_acc: 0.7070
Epoch 125/150
30/30 [=====] - 0s 3ms/step - loss: 1.2739 - acc: 0.7263
- val_loss: 1.2939 - val_acc: 0.7020
Epoch 126/150
30/30 [=====] - 0s 3ms/step - loss: 1.2695 - acc: 0.7283
```

```
- val_loss: 1.2884 - val_acc: 0.7030
Epoch 127/150
30/30 [=====] - 0s 2ms/step - loss: 1.2653 - acc: 0.7277
- val_loss: 1.2896 - val_acc: 0.7000
Epoch 128/150
30/30 [=====] - 0s 3ms/step - loss: 1.2611 - acc: 0.7297
- val_loss: 1.2823 - val_acc: 0.7040
Epoch 129/150
30/30 [=====] - 0s 3ms/step - loss: 1.2574 - acc: 0.7281
- val_loss: 1.2772 - val_acc: 0.7040
Epoch 130/150
30/30 [=====] - 0s 2ms/step - loss: 1.2529 - acc: 0.7297
- val_loss: 1.2744 - val_acc: 0.7070
Epoch 131/150
30/30 [=====] - 0s 3ms/step - loss: 1.2485 - acc: 0.7305
- val_loss: 1.2689 - val_acc: 0.7040
Epoch 132/150
30/30 [=====] - 0s 3ms/step - loss: 1.2444 - acc: 0.7304
- val_loss: 1.2644 - val_acc: 0.7020
Epoch 133/150
30/30 [=====] - 0s 2ms/step - loss: 1.2403 - acc: 0.7297
- val_loss: 1.2708 - val_acc: 0.7050
Epoch 134/150
30/30 [=====] - 0s 3ms/step - loss: 1.2368 - acc: 0.7308
- val_loss: 1.2579 - val_acc: 0.7060
Epoch 135/150
30/30 [=====] - 0s 2ms/step - loss: 1.2326 - acc: 0.7308
- val_loss: 1.2542 - val_acc: 0.7080
Epoch 136/150
30/30 [=====] - 0s 3ms/step - loss: 1.2286 - acc: 0.7315
- val_loss: 1.2552 - val_acc: 0.7060
Epoch 137/150
30/30 [=====] - 0s 3ms/step - loss: 1.2257 - acc: 0.7317
- val_loss: 1.2508 - val_acc: 0.7080
Epoch 138/150
30/30 [=====] - 0s 3ms/step - loss: 1.2216 - acc: 0.7315
- val_loss: 1.2449 - val_acc: 0.7060
Epoch 139/150
30/30 [=====] - 0s 3ms/step - loss: 1.2179 - acc: 0.7315
- val_loss: 1.2448 - val_acc: 0.7040
Epoch 140/150
30/30 [=====] - 0s 3ms/step - loss: 1.2141 - acc: 0.7329
- val_loss: 1.2380 - val_acc: 0.7080
Epoch 141/150
30/30 [=====] - 0s 3ms/step - loss: 1.2109 - acc: 0.7323
- val_loss: 1.2397 - val_acc: 0.7090
Epoch 142/150
30/30 [=====] - 0s 3ms/step - loss: 1.2070 - acc: 0.7336
- val_loss: 1.2316 - val_acc: 0.7100
```

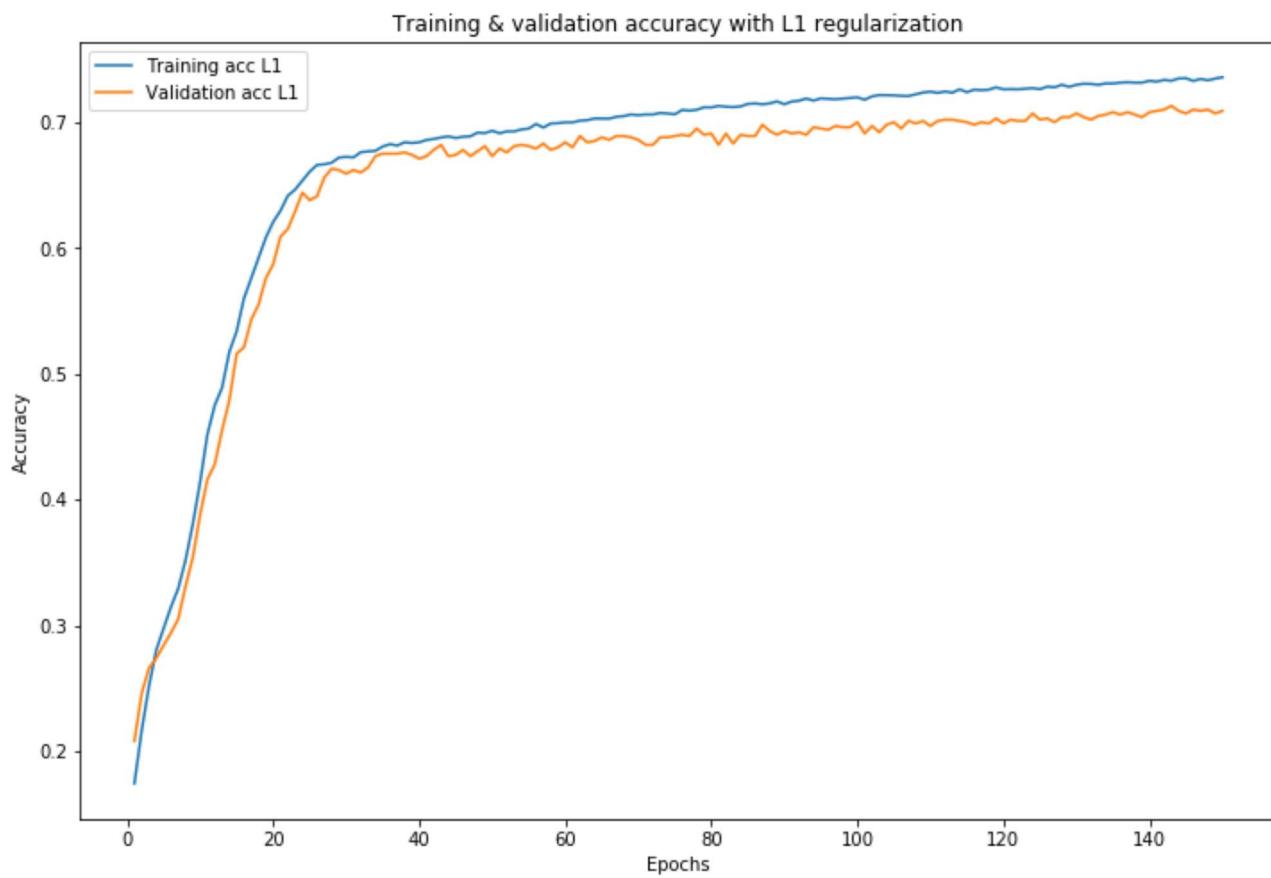
```
Epoch 143/150
30/30 [=====] - 0s 3ms/step - loss: 1.2037 - acc: 0.7329
- val_loss: 1.2307 - val_acc: 0.7130
Epoch 144/150
30/30 [=====] - 0s 3ms/step - loss: 1.2002 - acc: 0.7348
- val_loss: 1.2246 - val_acc: 0.7090
Epoch 145/150
30/30 [=====] - 0s 3ms/step - loss: 1.1967 - acc: 0.7349
- val_loss: 1.2198 - val_acc: 0.7070
Epoch 146/150
30/30 [=====] - 0s 3ms/step - loss: 1.1933 - acc: 0.7328
- val_loss: 1.2160 - val_acc: 0.7100
Epoch 147/150
30/30 [=====] - 0s 3ms/step - loss: 1.1898 - acc: 0.7344
- val_loss: 1.2129 - val_acc: 0.7090
Epoch 148/150
30/30 [=====] - 0s 3ms/step - loss: 1.1868 - acc: 0.7333
- val_loss: 1.2115 - val_acc: 0.7100
Epoch 149/150
30/30 [=====] - 0s 3ms/step - loss: 1.1835 - acc: 0.7347
- val_loss: 1.2086 - val_acc: 0.7070
Epoch 150/150
30/30 [=====] - 0s 3ms/step - loss: 1.1804 - acc: 0.7357
- val_loss: 1.2086 - val_acc: 0.7090
```

```
fig, ax = plt.subplots(figsize=(12, 8))

L1_model_dict = L1_model_val.history

acc_values = L1_model_dict['acc']
val_acc_values = L1_model_dict['val_acc']

epochs = range(1, len(acc_values) + 1)
ax.plot(epochs, acc_values, label='Training acc L1')
ax.plot(epochs, val_acc_values, label='Validation acc L1')
ax.set_title('Training & validation accuracy with L1 regularization')
ax.set_xlabel('Epochs')
ax.set_ylabel('Accuracy')
ax.legend();
```



Notice how the training and validation accuracy don't diverge as much as before. Unfortunately, the validation accuracy isn't still that good. Next, experiment with dropout regularization to see if it offers any advantages.

## Dropout Regularization

It's time to try another technique: applying dropout to layers. As discussed in the earlier lesson, this involves setting a certain proportion of units in each layer to zero. In the following cell:

- Apply a dropout rate of 30% to the input layer
- Add a first hidden layer with 50 units and 'relu' activation
- Apply a dropout rate of 30% to the first hidden layer
- Add a second hidden layer with 25 units and 'relu' activation
- Apply a dropout rate of 30% to the second hidden layer

```
# ⏳ This cell may take about a minute to run
random.seed(123)
dropout_model = models.Sequential()

# Implement dropout to the input layer
```

```
# NOTE: This is where you define the number of units in the input layer
dropout_model.add(layers.Dropout(0.3, input_shape=(2000,)))

# Add the first hidden layer
dropout_model.add(layers.Dense(50, activation='relu'))

# Implement dropout to the first hidden layer
dropout_model.add(layers.Dropout(0.3))

# Add the second hidden layer
dropout_model.add(layers.Dense(25, activation='relu'))

# Implement dropout to the second hidden layer
dropout_model.add(layers.Dropout(0.3))

# Add the output layer
dropout_model.add(layers.Dense(7, activation='softmax'))

# Compile the model
dropout_model.compile(optimizer='SGD',
                      loss='categorical_crossentropy',
                      metrics=['acc'])

# Train the model
dropout_model_val = dropout_model.fit(X_train_tokens,
                                       y_train_lb,
                                       epochs=150,
                                       batch_size=256,
                                       validation_data=(X_val_tokens, y_val_lb))
```

```
Epoch 1/150
30/30 [=====] - 0s 7ms/step - loss: 1.9679 - acc: 0.1705
- val_loss: 1.9422 - val_acc: 0.1890
Epoch 2/150
30/30 [=====] - 0s 4ms/step - loss: 1.9556 - acc: 0.1745
- val_loss: 1.9351 - val_acc: 0.1940
Epoch 3/150
30/30 [=====] - 0s 4ms/step - loss: 1.9452 - acc: 0.1832
- val_loss: 1.9296 - val_acc: 0.2030
Epoch 4/150
30/30 [=====] - 0s 4ms/step - loss: 1.9398 - acc: 0.1871
- val_loss: 1.9239 - val_acc: 0.2110
Epoch 5/150
30/30 [=====] - 0s 4ms/step - loss: 1.9314 - acc: 0.1987
- val_loss: 1.9179 - val_acc: 0.2160
Epoch 6/150
30/30 [=====] - 0s 4ms/step - loss: 1.9263 - acc: 0.2027
```

```
- val_loss: 1.9121 - val_acc: 0.2170
Epoch 7/150
30/30 [=====] - 0s 4ms/step - loss: 1.9160 - acc: 0.2084
- val_loss: 1.9046 - val_acc: 0.2170
Epoch 8/150
30/30 [=====] - 0s 4ms/step - loss: 1.9100 - acc: 0.2147
- val_loss: 1.8969 - val_acc: 0.2290
Epoch 9/150
30/30 [=====] - 0s 4ms/step - loss: 1.9057 - acc: 0.2195
- val_loss: 1.8895 - val_acc: 0.2340
Epoch 10/150
30/30 [=====] - 0s 4ms/step - loss: 1.8991 - acc: 0.2179
- val_loss: 1.8810 - val_acc: 0.2300
Epoch 11/150
30/30 [=====] - 0s 4ms/step - loss: 1.8875 - acc: 0.2305
- val_loss: 1.8718 - val_acc: 0.2380
Epoch 12/150
30/30 [=====] - 0s 4ms/step - loss: 1.8803 - acc: 0.2377
- val_loss: 1.8609 - val_acc: 0.2520
Epoch 13/150
30/30 [=====] - 0s 3ms/step - loss: 1.8713 - acc: 0.2367
- val_loss: 1.8500 - val_acc: 0.2700
Epoch 14/150
30/30 [=====] - 0s 4ms/step - loss: 1.8570 - acc: 0.2469
- val_loss: 1.8371 - val_acc: 0.2740
Epoch 15/150
30/30 [=====] - 0s 4ms/step - loss: 1.8515 - acc: 0.2540
- val_loss: 1.8237 - val_acc: 0.2810
Epoch 16/150
30/30 [=====] - 0s 4ms/step - loss: 1.8390 - acc: 0.2601
- val_loss: 1.8085 - val_acc: 0.2910
Epoch 17/150
30/30 [=====] - 0s 4ms/step - loss: 1.8348 - acc: 0.2572
- val_loss: 1.7927 - val_acc: 0.3100
Epoch 18/150
30/30 [=====] - 0s 4ms/step - loss: 1.8174 - acc: 0.2713
- val_loss: 1.7765 - val_acc: 0.3230
Epoch 19/150
30/30 [=====] - 0s 4ms/step - loss: 1.7987 - acc: 0.2800
- val_loss: 1.7573 - val_acc: 0.3370
Epoch 20/150
30/30 [=====] - 0s 4ms/step - loss: 1.7899 - acc: 0.2879
- val_loss: 1.7381 - val_acc: 0.3500
Epoch 21/150
30/30 [=====] - 0s 4ms/step - loss: 1.7742 - acc: 0.3021
- val_loss: 1.7179 - val_acc: 0.3600
Epoch 22/150
30/30 [=====] - 0s 4ms/step - loss: 1.7686 - acc: 0.2983
- val_loss: 1.7003 - val_acc: 0.3760
```

```
Epoch 23/150
30/30 [=====] - 0s 4ms/step - loss: 1.7437 - acc: 0.3084
- val_loss: 1.6803 - val_acc: 0.3990
Epoch 24/150
30/30 [=====] - 0s 4ms/step - loss: 1.7397 - acc: 0.3091
- val_loss: 1.6595 - val_acc: 0.4020
Epoch 25/150
30/30 [=====] - 0s 4ms/step - loss: 1.7203 - acc: 0.3148
- val_loss: 1.6396 - val_acc: 0.4200
Epoch 26/150
30/30 [=====] - 0s 4ms/step - loss: 1.7082 - acc: 0.3324
- val_loss: 1.6199 - val_acc: 0.4400
Epoch 27/150
30/30 [=====] - 0s 4ms/step - loss: 1.6884 - acc: 0.3379
- val_loss: 1.5981 - val_acc: 0.4480
Epoch 28/150
30/30 [=====] - 0s 4ms/step - loss: 1.6761 - acc: 0.3400
- val_loss: 1.5781 - val_acc: 0.4610
Epoch 29/150
30/30 [=====] - 0s 4ms/step - loss: 1.6505 - acc: 0.3537
- val_loss: 1.5535 - val_acc: 0.4660
Epoch 30/150
30/30 [=====] - 0s 4ms/step - loss: 1.6375 - acc: 0.3596
- val_loss: 1.5345 - val_acc: 0.4700
Epoch 31/150
30/30 [=====] - 0s 4ms/step - loss: 1.6271 - acc: 0.3693
- val_loss: 1.5141 - val_acc: 0.4770
Epoch 32/150
30/30 [=====] - 0s 4ms/step - loss: 1.6110 - acc: 0.3760
- val_loss: 1.4942 - val_acc: 0.4860
Epoch 33/150
30/30 [=====] - 0s 4ms/step - loss: 1.6031 - acc: 0.3783
- val_loss: 1.4763 - val_acc: 0.4980
Epoch 34/150
30/30 [=====] - 0s 4ms/step - loss: 1.5931 - acc: 0.3867
- val_loss: 1.4554 - val_acc: 0.5050
Epoch 35/150
30/30 [=====] - 0s 4ms/step - loss: 1.5656 - acc: 0.3952
- val_loss: 1.4332 - val_acc: 0.5130
Epoch 36/150
30/30 [=====] - 0s 4ms/step - loss: 1.5501 - acc: 0.3992
- val_loss: 1.4106 - val_acc: 0.5280
Epoch 37/150
30/30 [=====] - 0s 4ms/step - loss: 1.5422 - acc: 0.4057
- val_loss: 1.3946 - val_acc: 0.5310
Epoch 38/150
30/30 [=====] - 0s 4ms/step - loss: 1.5293 - acc: 0.4025
- val_loss: 1.3768 - val_acc: 0.5430
Epoch 39/150
```

```
30/30 [=====] - 0s 4ms/step - loss: 1.5167 - acc: 0.4123
- val_loss: 1.3606 - val_acc: 0.5530
Epoch 40/150
30/30 [=====] - 0s 4ms/step - loss: 1.5023 - acc: 0.4232
- val_loss: 1.3424 - val_acc: 0.5680
Epoch 41/150
30/30 [=====] - 0s 4ms/step - loss: 1.4867 - acc: 0.4276
- val_loss: 1.3251 - val_acc: 0.5730
Epoch 42/150
30/30 [=====] - 0s 4ms/step - loss: 1.4747 - acc: 0.4393
- val_loss: 1.3087 - val_acc: 0.5780
Epoch 43/150
30/30 [=====] - 0s 4ms/step - loss: 1.4562 - acc: 0.4413
- val_loss: 1.2895 - val_acc: 0.5860
Epoch 44/150
30/30 [=====] - 0s 4ms/step - loss: 1.4502 - acc: 0.4413
- val_loss: 1.2749 - val_acc: 0.5870
Epoch 45/150
30/30 [=====] - 0s 3ms/step - loss: 1.4416 - acc: 0.4481
- val_loss: 1.2609 - val_acc: 0.5900
Epoch 46/150
30/30 [=====] - 0s 3ms/step - loss: 1.4190 - acc: 0.4581
- val_loss: 1.2428 - val_acc: 0.5970
Epoch 47/150
30/30 [=====] - 0s 4ms/step - loss: 1.4155 - acc: 0.4560
- val_loss: 1.2293 - val_acc: 0.6080
Epoch 48/150
30/30 [=====] - 0s 4ms/step - loss: 1.4048 - acc: 0.4645
- val_loss: 1.2150 - val_acc: 0.6110
Epoch 49/150
30/30 [=====] - 0s 4ms/step - loss: 1.3943 - acc: 0.4649
- val_loss: 1.2020 - val_acc: 0.6110
Epoch 50/150
30/30 [=====] - 0s 4ms/step - loss: 1.3817 - acc: 0.4720
- val_loss: 1.1939 - val_acc: 0.6080
Epoch 51/150
30/30 [=====] - 0s 4ms/step - loss: 1.3694 - acc: 0.4739
- val_loss: 1.1775 - val_acc: 0.6310
Epoch 52/150
30/30 [=====] - 0s 4ms/step - loss: 1.3658 - acc: 0.4831
- val_loss: 1.1647 - val_acc: 0.6290
Epoch 53/150
30/30 [=====] - 0s 4ms/step - loss: 1.3658 - acc: 0.4731
- val_loss: 1.1552 - val_acc: 0.6390
Epoch 54/150
30/30 [=====] - 0s 4ms/step - loss: 1.3381 - acc: 0.4937
- val_loss: 1.1428 - val_acc: 0.6410
Epoch 55/150
30/30 [=====] - 0s 4ms/step - loss: 1.3349 - acc: 0.4919
```

```
- val_loss: 1.1300 - val_acc: 0.6440
Epoch 56/150
30/30 [=====] - 0s 4ms/step - loss: 1.3302 - acc: 0.4975
- val_loss: 1.1212 - val_acc: 0.6490
Epoch 57/150
30/30 [=====] - 0s 4ms/step - loss: 1.3068 - acc: 0.4993
- val_loss: 1.1062 - val_acc: 0.6550
Epoch 58/150
30/30 [=====] - 0s 4ms/step - loss: 1.2975 - acc: 0.5059
- val_loss: 1.0992 - val_acc: 0.6550
Epoch 59/150
30/30 [=====] - 0s 4ms/step - loss: 1.2951 - acc: 0.5080
- val_loss: 1.0876 - val_acc: 0.6540
Epoch 60/150
30/30 [=====] - 0s 4ms/step - loss: 1.2905 - acc: 0.5135
- val_loss: 1.0738 - val_acc: 0.6600
Epoch 61/150
30/30 [=====] - 0s 4ms/step - loss: 1.2933 - acc: 0.5052
- val_loss: 1.0667 - val_acc: 0.6650
Epoch 62/150
30/30 [=====] - 0s 4ms/step - loss: 1.2654 - acc: 0.5240
- val_loss: 1.0536 - val_acc: 0.6660
Epoch 63/150
30/30 [=====] - 0s 4ms/step - loss: 1.2720 - acc: 0.5260
- val_loss: 1.0461 - val_acc: 0.6670
Epoch 64/150
30/30 [=====] - 0s 4ms/step - loss: 1.2522 - acc: 0.5212
- val_loss: 1.0362 - val_acc: 0.6650
Epoch 65/150
30/30 [=====] - 0s 4ms/step - loss: 1.2465 - acc: 0.5317
- val_loss: 1.0295 - val_acc: 0.6680
Epoch 66/150
30/30 [=====] - 0s 4ms/step - loss: 1.2314 - acc: 0.5429
- val_loss: 1.0154 - val_acc: 0.6690
Epoch 67/150
30/30 [=====] - 0s 4ms/step - loss: 1.2322 - acc: 0.5360
- val_loss: 1.0104 - val_acc: 0.6730
Epoch 68/150
30/30 [=====] - 0s 4ms/step - loss: 1.2142 - acc: 0.5475
- val_loss: 0.9996 - val_acc: 0.6750
Epoch 69/150
30/30 [=====] - 0s 4ms/step - loss: 1.2233 - acc: 0.5420
- val_loss: 0.9902 - val_acc: 0.6760
Epoch 70/150
30/30 [=====] - 0s 4ms/step - loss: 1.2116 - acc: 0.5503
- val_loss: 0.9847 - val_acc: 0.6760
Epoch 71/150
30/30 [=====] - 0s 4ms/step - loss: 1.2161 - acc: 0.5485
- val_loss: 0.9826 - val_acc: 0.6840
```

```
Epoch 72/150
30/30 [=====] - 0s 4ms/step - loss: 1.1893 - acc: 0.5535
- val_loss: 0.9718 - val_acc: 0.6810
Epoch 73/150
30/30 [=====] - 0s 4ms/step - loss: 1.1942 - acc: 0.5471
- val_loss: 0.9611 - val_acc: 0.6860
Epoch 74/150
30/30 [=====] - 0s 4ms/step - loss: 1.1724 - acc: 0.5585
- val_loss: 0.9549 - val_acc: 0.6880
Epoch 75/150
30/30 [=====] - 0s 4ms/step - loss: 1.1872 - acc: 0.5507
- val_loss: 0.9460 - val_acc: 0.6930
Epoch 76/150
30/30 [=====] - 0s 3ms/step - loss: 1.1637 - acc: 0.5675
- val_loss: 0.9376 - val_acc: 0.6950
Epoch 77/150
30/30 [=====] - 0s 4ms/step - loss: 1.1635 - acc: 0.5660
- val_loss: 0.9294 - val_acc: 0.6920
Epoch 78/150
30/30 [=====] - 0s 4ms/step - loss: 1.1568 - acc: 0.5624
- val_loss: 0.9224 - val_acc: 0.7000
Epoch 79/150
30/30 [=====] - 0s 4ms/step - loss: 1.1410 - acc: 0.5837
- val_loss: 0.9130 - val_acc: 0.7030
Epoch 80/150
30/30 [=====] - 0s 4ms/step - loss: 1.1469 - acc: 0.5704
- val_loss: 0.9135 - val_acc: 0.6980
Epoch 81/150
30/30 [=====] - 0s 4ms/step - loss: 1.1388 - acc: 0.5804
- val_loss: 0.9085 - val_acc: 0.6990
Epoch 82/150
30/30 [=====] - 0s 4ms/step - loss: 1.1271 - acc: 0.5852
- val_loss: 0.8976 - val_acc: 0.7040
Epoch 83/150
30/30 [=====] - 0s 4ms/step - loss: 1.1121 - acc: 0.5885
- val_loss: 0.8918 - val_acc: 0.7010
Epoch 84/150
30/30 [=====] - 0s 4ms/step - loss: 1.1229 - acc: 0.5831
- val_loss: 0.8885 - val_acc: 0.7060
Epoch 85/150
30/30 [=====] - 0s 4ms/step - loss: 1.1006 - acc: 0.5909
- val_loss: 0.8758 - val_acc: 0.7110
Epoch 86/150
30/30 [=====] - 0s 4ms/step - loss: 1.1000 - acc: 0.5949
- val_loss: 0.8713 - val_acc: 0.7140
Epoch 87/150
30/30 [=====] - 0s 4ms/step - loss: 1.1123 - acc: 0.5835
- val_loss: 0.8668 - val_acc: 0.7130
Epoch 88/150
```

```
30/30 [=====] - 0s 4ms/step - loss: 1.1024 - acc: 0.5895
- val_loss: 0.8627 - val_acc: 0.7120
Epoch 89/150
30/30 [=====] - 0s 4ms/step - loss: 1.1010 - acc: 0.5932
- val_loss: 0.8608 - val_acc: 0.7090
Epoch 90/150
30/30 [=====] - 0s 4ms/step - loss: 1.0728 - acc: 0.6008
- val_loss: 0.8536 - val_acc: 0.7090
Epoch 91/150
30/30 [=====] - 0s 4ms/step - loss: 1.0826 - acc: 0.6049
- val_loss: 0.8487 - val_acc: 0.7130
Epoch 92/150
30/30 [=====] - 0s 4ms/step - loss: 1.0857 - acc: 0.5981
- val_loss: 0.8449 - val_acc: 0.7130
Epoch 93/150
30/30 [=====] - 0s 4ms/step - loss: 1.0719 - acc: 0.6015
- val_loss: 0.8380 - val_acc: 0.7120
Epoch 94/150
30/30 [=====] - 0s 4ms/step - loss: 1.0634 - acc: 0.6075
- val_loss: 0.8347 - val_acc: 0.7140
Epoch 95/150
30/30 [=====] - 0s 4ms/step - loss: 1.0670 - acc: 0.6105
- val_loss: 0.8309 - val_acc: 0.7170
Epoch 96/150
30/30 [=====] - 0s 4ms/step - loss: 1.0529 - acc: 0.6135
- val_loss: 0.8224 - val_acc: 0.7110
Epoch 97/150
30/30 [=====] - 0s 4ms/step - loss: 1.0489 - acc: 0.6097
- val_loss: 0.8205 - val_acc: 0.7140
Epoch 98/150
30/30 [=====] - 0s 4ms/step - loss: 1.0521 - acc: 0.6032
- val_loss: 0.8175 - val_acc: 0.7190
Epoch 99/150
30/30 [=====] - 0s 4ms/step - loss: 1.0263 - acc: 0.6253
- val_loss: 0.8104 - val_acc: 0.7200
Epoch 100/150
30/30 [=====] - 0s 4ms/step - loss: 1.0418 - acc: 0.6173
- val_loss: 0.8055 - val_acc: 0.7250
Epoch 101/150
30/30 [=====] - 0s 4ms/step - loss: 1.0353 - acc: 0.6207
- val_loss: 0.8022 - val_acc: 0.7260
Epoch 102/150
30/30 [=====] - 0s 4ms/step - loss: 1.0334 - acc: 0.6147
- val_loss: 0.8014 - val_acc: 0.7270
Epoch 103/150
30/30 [=====] - 0s 4ms/step - loss: 1.0332 - acc: 0.6223
- val_loss: 0.7988 - val_acc: 0.7250
Epoch 104/150
30/30 [=====] - 0s 4ms/step - loss: 1.0301 - acc: 0.6215
```

```
- val_loss: 0.7955 - val_acc: 0.7280
Epoch 105/150
30/30 [=====] - 0s 4ms/step - loss: 1.0231 - acc: 0.6235
- val_loss: 0.7939 - val_acc: 0.7320
Epoch 106/150
30/30 [=====] - 0s 4ms/step - loss: 1.0157 - acc: 0.6263
- val_loss: 0.7895 - val_acc: 0.7280
Epoch 107/150
30/30 [=====] - 0s 4ms/step - loss: 0.9984 - acc: 0.6357
- val_loss: 0.7829 - val_acc: 0.7300
Epoch 108/150
30/30 [=====] - 0s 4ms/step - loss: 1.0086 - acc: 0.6292
- val_loss: 0.7803 - val_acc: 0.7340
Epoch 109/150
30/30 [=====] - 0s 4ms/step - loss: 1.0101 - acc: 0.6253
- val_loss: 0.7810 - val_acc: 0.7330
Epoch 110/150
30/30 [=====] - 0s 4ms/step - loss: 0.9872 - acc: 0.6419
- val_loss: 0.7718 - val_acc: 0.7390
Epoch 111/150
30/30 [=====] - 0s 4ms/step - loss: 0.9755 - acc: 0.6421
- val_loss: 0.7669 - val_acc: 0.7360
Epoch 112/150
30/30 [=====] - 0s 4ms/step - loss: 0.9904 - acc: 0.6367
- val_loss: 0.7657 - val_acc: 0.7360
Epoch 113/150
30/30 [=====] - 0s 4ms/step - loss: 0.9899 - acc: 0.6368
- val_loss: 0.7680 - val_acc: 0.7350
Epoch 114/150
30/30 [=====] - 0s 3ms/step - loss: 0.9966 - acc: 0.6403
- val_loss: 0.7627 - val_acc: 0.7380
Epoch 115/150
30/30 [=====] - 0s 3ms/step - loss: 0.9878 - acc: 0.6387
- val_loss: 0.7612 - val_acc: 0.7340
Epoch 116/150
30/30 [=====] - 0s 4ms/step - loss: 0.9850 - acc: 0.6367
- val_loss: 0.7564 - val_acc: 0.7370
Epoch 117/150
30/30 [=====] - 0s 4ms/step - loss: 0.9687 - acc: 0.6480
- val_loss: 0.7536 - val_acc: 0.7410
Epoch 118/150
30/30 [=====] - 0s 4ms/step - loss: 0.9773 - acc: 0.6421
- val_loss: 0.7496 - val_acc: 0.7380
Epoch 119/150
30/30 [=====] - 0s 4ms/step - loss: 0.9747 - acc: 0.6347
- val_loss: 0.7468 - val_acc: 0.7410
Epoch 120/150
30/30 [=====] - 0s 4ms/step - loss: 0.9764 - acc: 0.6411
- val_loss: 0.7485 - val_acc: 0.7400
```

```
Epoch 121/150
30/30 [=====] - 0s 4ms/step - loss: 0.9680 - acc: 0.6424
- val_loss: 0.7420 - val_acc: 0.7380
Epoch 122/150
30/30 [=====] - 0s 4ms/step - loss: 0.9625 - acc: 0.6505
- val_loss: 0.7449 - val_acc: 0.7400
Epoch 123/150
30/30 [=====] - 0s 4ms/step - loss: 0.9521 - acc: 0.6484
- val_loss: 0.7390 - val_acc: 0.7400
Epoch 124/150
30/30 [=====] - 0s 4ms/step - loss: 0.9415 - acc: 0.6644
- val_loss: 0.7350 - val_acc: 0.7440
Epoch 125/150
30/30 [=====] - 0s 4ms/step - loss: 0.9321 - acc: 0.6640
- val_loss: 0.7328 - val_acc: 0.7440
Epoch 126/150
30/30 [=====] - 0s 4ms/step - loss: 0.9430 - acc: 0.6525
- val_loss: 0.7315 - val_acc: 0.7430
Epoch 127/150
30/30 [=====] - 0s 4ms/step - loss: 0.9385 - acc: 0.6512
- val_loss: 0.7270 - val_acc: 0.7440
Epoch 128/150
30/30 [=====] - 0s 4ms/step - loss: 0.9205 - acc: 0.6604
- val_loss: 0.7214 - val_acc: 0.7450
Epoch 129/150
30/30 [=====] - 0s 4ms/step - loss: 0.9263 - acc: 0.6624
- val_loss: 0.7190 - val_acc: 0.7460
Epoch 130/150
30/30 [=====] - 0s 4ms/step - loss: 0.9290 - acc: 0.6592
- val_loss: 0.7202 - val_acc: 0.7440
Epoch 131/150
30/30 [=====] - 0s 4ms/step - loss: 0.9226 - acc: 0.6587
- val_loss: 0.7210 - val_acc: 0.7440
Epoch 132/150
30/30 [=====] - 0s 4ms/step - loss: 0.9195 - acc: 0.6636
- val_loss: 0.7150 - val_acc: 0.7430
Epoch 133/150
30/30 [=====] - 0s 4ms/step - loss: 0.9166 - acc: 0.6605
- val_loss: 0.7138 - val_acc: 0.7430
Epoch 134/150
30/30 [=====] - 0s 3ms/step - loss: 0.9144 - acc: 0.6647
- val_loss: 0.7114 - val_acc: 0.7450
Epoch 135/150
30/30 [=====] - 0s 4ms/step - loss: 0.9193 - acc: 0.6564
- val_loss: 0.7115 - val_acc: 0.7470
Epoch 136/150
30/30 [=====] - 0s 4ms/step - loss: 0.8944 - acc: 0.6739
- val_loss: 0.7090 - val_acc: 0.7470
Epoch 137/150
```

```
30/30 [=====] - 0s 4ms/step - loss: 0.9181 - acc: 0.6583
- val_loss: 0.7066 - val_acc: 0.7430
Epoch 138/150
30/30 [=====] - 0s 4ms/step - loss: 0.9219 - acc: 0.6619
- val_loss: 0.7067 - val_acc: 0.7440
Epoch 139/150
30/30 [=====] - 0s 4ms/step - loss: 0.9090 - acc: 0.6684
- val_loss: 0.7030 - val_acc: 0.7500
Epoch 140/150
30/30 [=====] - 0s 4ms/step - loss: 0.9014 - acc: 0.6731
- val_loss: 0.7020 - val_acc: 0.7490
Epoch 141/150
30/30 [=====] - 0s 4ms/step - loss: 0.8881 - acc: 0.6727
- val_loss: 0.7034 - val_acc: 0.7460
Epoch 142/150
30/30 [=====] - 0s 4ms/step - loss: 0.8918 - acc: 0.6719
- val_loss: 0.6972 - val_acc: 0.7460
Epoch 143/150
30/30 [=====] - 0s 4ms/step - loss: 0.8875 - acc: 0.6695
- val_loss: 0.6965 - val_acc: 0.7480
Epoch 144/150
30/30 [=====] - 0s 4ms/step - loss: 0.8888 - acc: 0.6772
- val_loss: 0.6933 - val_acc: 0.7480
Epoch 145/150
30/30 [=====] - 0s 4ms/step - loss: 0.8787 - acc: 0.6756
- val_loss: 0.6894 - val_acc: 0.7490
Epoch 146/150
30/30 [=====] - 0s 4ms/step - loss: 0.8786 - acc: 0.6823
- val_loss: 0.6894 - val_acc: 0.7520
Epoch 147/150
30/30 [=====] - 0s 4ms/step - loss: 0.8749 - acc: 0.6791
- val_loss: 0.6855 - val_acc: 0.7510
Epoch 148/150
30/30 [=====] - 0s 4ms/step - loss: 0.8741 - acc: 0.6800
- val_loss: 0.6870 - val_acc: 0.7510
Epoch 149/150
30/30 [=====] - 0s 4ms/step - loss: 0.8774 - acc: 0.6760
- val_loss: 0.6865 - val_acc: 0.7470
Epoch 150/150
30/30 [=====] - 0s 4ms/step - loss: 0.8750 - acc: 0.6792
- val_loss: 0.6873 - val_acc: 0.7510
```

```
results_train = dropout_model.evaluate(X_train_tokens, y_train_lb)
print(f'Training Loss: {results_train[0]:.3} \nTraining Accuracy: {results_train[1]}:\n\nprint('-----')
```

```
results_test = dropout_model.evaluate(X_test_tokens, y_test_lb)
print(f'Test Loss: {results_test[0]:.3} \nTest Accuracy: {results_test[1]:.3}')
```

```
235/235 [=====] - 0s 572us/step - loss: 0.5797 - acc: 0.8067
Training Loss: 0.58
Training Accuracy: 0.807
-----
47/47 [=====] - 0s 681us/step - loss: 0.6421 - acc: 0.7827
Test Loss: 0.642
Test Accuracy: 0.783
```

You can see here that the validation performance has improved again, and the training and test accuracy are very close!

## Bigger Data?

---

Finally, let's examine if we can improve the model's performance just by adding more data. We've quadrupled the sample dataset from 10,000 to 40,000 observations, and all you need to do is run the code!

```
df_bigger_sample = df.sample(40000, random_state=123)

X = df['Consumer complaint narrative']
y = df['Product']

# Train-test split
X_train_bigger, X_test_bigger, y_train_bigger, y_test_bigger = train_test_split(X,
                                                                           y,
                                                                           test_size=0.2,
                                                                           random_state=123)

# Validation set
X_train_final_bigger, X_val_bigger, y_train_final_bigger, y_val_bigger = train_test_split(
    X_train_bigger, y_train_bigger, test_size=0.2, random_state=123)

# One-hot encoding of the complaints
tokenizer = Tokenizer(num_words=2000)
tokenizer.fit_on_texts(X_train_final_bigger)
```

```
X_train_tokens_bigger = tokenizer.texts_to_matrix(X_train_final_bigger, mode='binary')
X_val_tokens_bigger = tokenizer.texts_to_matrix(X_val_bigger, mode='binary')
X_test_tokens_bigger = tokenizer.texts_to_matrix(X_test_bigger, mode='binary')

# One-hot encoding of products
lb = LabelBinarizer()
lb.fit(y_train_final_bigger)

y_train_lb_bigger = to_categorical(lb.transform(y_train_final_bigger))[:, :, 1]
y_val_lb_bigger = to_categorical(lb.transform(y_val_bigger))[:, :, 1]
y_test_lb_bigger = to_categorical(lb.transform(y_test_bigger))[:, :, 1]
```



```
# ⚠ This cell may take several minutes to run
random.seed(123)
bigger_data_model = models.Sequential()
bigger_data_model.add(layers.Dense(50, activation='relu', input_shape=(2000,)))
bigger_data_model.add(layers.Dense(25, activation='relu'))
bigger_data_model.add(layers.Dense(7, activation='softmax'))

bigger_data_model.compile(optimizer='SGD',
                          loss='categorical_crossentropy',
                          metrics=['acc'])

bigger_data_model_val = bigger_data_model.fit(X_train_tokens_bigger,
                                              y_train_lb_bigger,
                                              epochs=150,
                                              batch_size=256,
                                              validation_data=(X_val_tokens_bigger,
```



```
Epoch 1/150
196/196 [=====] - 0s 2ms/step - loss: 1.8998 - acc: 0.2192 - val_loss: 1.8212 - val_acc: 0.3203
Epoch 2/150
196/196 [=====] - 0s 2ms/step - loss: 1.6876 - acc: 0.4225 - val_loss: 1.5375 - val_acc: 0.5107
Epoch 3/150
196/196 [=====] - 0s 2ms/step - loss: 1.3562 - acc: 0.5901 - val_loss: 1.2113 - val_acc: 0.6348
Epoch 4/150
196/196 [=====] - 0s 2ms/step - loss: 1.0712 - acc: 0.6742 - val_loss: 0.9885 - val_acc: 0.6898
Epoch 5/150
196/196 [=====] - 0s 2ms/step - loss: 0.8954 - acc: 0.7098 - val_loss: 0.8598 - val_acc: 0.7122
Epoch 6/150
```

```
196/196 [=====] - 0s 2ms/step - loss: 0.7929 - acc: 0.7307 - val_loss: 0.7852 - val_acc: 0.7318
Epoch 7/150
196/196 [=====] - 0s 2ms/step - loss: 0.7301 - acc: 0.7443 - val_loss: 0.7382 - val_acc: 0.7427
Epoch 8/150
196/196 [=====] - 0s 2ms/step - loss: 0.6874 - acc: 0.7557 - val_loss: 0.7052 - val_acc: 0.7492
Epoch 9/150
196/196 [=====] - 0s 2ms/step - loss: 0.6566 - acc: 0.7646 - val_loss: 0.6788 - val_acc: 0.7573
Epoch 10/150
196/196 [=====] - 0s 2ms/step - loss: 0.6328 - acc: 0.7711 - val_loss: 0.6607 - val_acc: 0.7660
Epoch 11/150
196/196 [=====] - 0s 2ms/step - loss: 0.6136 - acc: 0.7782 - val_loss: 0.6465 - val_acc: 0.7688
Epoch 12/150
196/196 [=====] - 0s 2ms/step - loss: 0.5973 - acc: 0.7842 - val_loss: 0.6337 - val_acc: 0.7682
Epoch 13/150
196/196 [=====] - 0s 2ms/step - loss: 0.5834 - acc: 0.7891 - val_loss: 0.6224 - val_acc: 0.7768
Epoch 14/150
196/196 [=====] - 0s 2ms/step - loss: 0.5711 - acc: 0.7928 - val_loss: 0.6121 - val_acc: 0.7780
Epoch 15/150
196/196 [=====] - 0s 2ms/step - loss: 0.5600 - acc: 0.7975 - val_loss: 0.6041 - val_acc: 0.7847
Epoch 16/150
196/196 [=====] - 0s 2ms/step - loss: 0.5503 - acc: 0.8008 - val_loss: 0.5986 - val_acc: 0.7885
Epoch 17/150
196/196 [=====] - 0s 2ms/step - loss: 0.5412 - acc: 0.8041 - val_loss: 0.5939 - val_acc: 0.7847
Epoch 18/150
196/196 [=====] - 0s 2ms/step - loss: 0.5329 - acc: 0.8085 - val_loss: 0.5863 - val_acc: 0.7933
Epoch 19/150
196/196 [=====] - 0s 2ms/step - loss: 0.5252 - acc: 0.8105 - val_loss: 0.5843 - val_acc: 0.7887
Epoch 20/150
196/196 [=====] - 0s 2ms/step - loss: 0.5180 - acc: 0.8133 - val_loss: 0.5766 - val_acc: 0.7947
Epoch 21/150
196/196 [=====] - 0s 2ms/step - loss: 0.5117 - acc: 0.8152 - val_loss: 0.5811 - val_acc: 0.7878
Epoch 22/150
196/196 [=====] - 0s 2ms/step - loss: 0.5052 - acc:
```

```
0.8184 - val_loss: 0.5694 - val_acc: 0.7995
Epoch 23/150
196/196 [=====] - 0s 2ms/step - loss: 0.4993 - acc:
0.8203 - val_loss: 0.5676 - val_acc: 0.7922
Epoch 24/150
196/196 [=====] - 0s 2ms/step - loss: 0.4935 - acc:
0.8232 - val_loss: 0.5616 - val_acc: 0.7983
Epoch 25/150
196/196 [=====] - 0s 2ms/step - loss: 0.4884 - acc:
0.8254 - val_loss: 0.5599 - val_acc: 0.8035
Epoch 26/150
196/196 [=====] - 0s 2ms/step - loss: 0.4834 - acc:
0.8280 - val_loss: 0.5579 - val_acc: 0.8010
Epoch 27/150
196/196 [=====] - 0s 2ms/step - loss: 0.4789 - acc:
0.8290 - val_loss: 0.5559 - val_acc: 0.8058
Epoch 28/150
196/196 [=====] - 0s 2ms/step - loss: 0.4743 - acc:
0.8306 - val_loss: 0.5560 - val_acc: 0.7995
Epoch 29/150
196/196 [=====] - 0s 2ms/step - loss: 0.4700 - acc:
0.8332 - val_loss: 0.5495 - val_acc: 0.8035
Epoch 30/150
196/196 [=====] - 0s 2ms/step - loss: 0.4659 - acc:
0.8340 - val_loss: 0.5497 - val_acc: 0.8085
Epoch 31/150
196/196 [=====] - 0s 2ms/step - loss: 0.4619 - acc:
0.8357 - val_loss: 0.5469 - val_acc: 0.8043
Epoch 32/150
196/196 [=====] - 0s 2ms/step - loss: 0.4584 - acc:
0.8366 - val_loss: 0.5480 - val_acc: 0.8075
Epoch 33/150
196/196 [=====] - 0s 2ms/step - loss: 0.4547 - acc:
0.8378 - val_loss: 0.5483 - val_acc: 0.8023
Epoch 34/150
196/196 [=====] - 0s 2ms/step - loss: 0.4514 - acc:
0.8400 - val_loss: 0.5455 - val_acc: 0.8043
Epoch 35/150
196/196 [=====] - 0s 2ms/step - loss: 0.4477 - acc:
0.8414 - val_loss: 0.5418 - val_acc: 0.8077
Epoch 36/150
196/196 [=====] - 0s 2ms/step - loss: 0.4448 - acc:
0.8429 - val_loss: 0.5421 - val_acc: 0.8077
Epoch 37/150
196/196 [=====] - 0s 2ms/step - loss: 0.4418 - acc:
0.8444 - val_loss: 0.5413 - val_acc: 0.8045
Epoch 38/150
196/196 [=====] - 0s 2ms/step - loss: 0.4386 - acc:
0.8455 - val_loss: 0.5415 - val_acc: 0.8100
```

```
Epoch 39/150
196/196 [=====] - 0s 2ms/step - loss: 0.4357 - acc: 0.8456 - val_loss: 0.5411 - val_acc: 0.8030
Epoch 40/150
196/196 [=====] - 0s 2ms/step - loss: 0.4330 - acc: 0.8469 - val_loss: 0.5412 - val_acc: 0.8085
Epoch 41/150
196/196 [=====] - 0s 2ms/step - loss: 0.4305 - acc: 0.8481 - val_loss: 0.5383 - val_acc: 0.8115
Epoch 42/150
196/196 [=====] - 0s 2ms/step - loss: 0.4277 - acc: 0.8497 - val_loss: 0.5378 - val_acc: 0.8098
Epoch 43/150
196/196 [=====] - 0s 2ms/step - loss: 0.4252 - acc: 0.8505 - val_loss: 0.5395 - val_acc: 0.8095
Epoch 44/150
196/196 [=====] - 0s 2ms/step - loss: 0.4231 - acc: 0.8508 - val_loss: 0.5407 - val_acc: 0.8110
Epoch 45/150
196/196 [=====] - 0s 2ms/step - loss: 0.4203 - acc: 0.8518 - val_loss: 0.5389 - val_acc: 0.8058
Epoch 46/150
196/196 [=====] - 0s 2ms/step - loss: 0.4181 - acc: 0.8531 - val_loss: 0.5367 - val_acc: 0.8105
Epoch 47/150
196/196 [=====] - 0s 2ms/step - loss: 0.4156 - acc: 0.8536 - val_loss: 0.5402 - val_acc: 0.8058
Epoch 48/150
196/196 [=====] - 0s 2ms/step - loss: 0.4135 - acc: 0.8544 - val_loss: 0.5371 - val_acc: 0.8073
Epoch 49/150
196/196 [=====] - 0s 2ms/step - loss: 0.4112 - acc: 0.8550 - val_loss: 0.5392 - val_acc: 0.8105
Epoch 50/150
196/196 [=====] - 0s 2ms/step - loss: 0.4091 - acc: 0.8561 - val_loss: 0.5446 - val_acc: 0.8030
Epoch 51/150
196/196 [=====] - 0s 2ms/step - loss: 0.4072 - acc: 0.8567 - val_loss: 0.5356 - val_acc: 0.8133
Epoch 52/150
196/196 [=====] - 0s 2ms/step - loss: 0.4051 - acc: 0.8576 - val_loss: 0.5362 - val_acc: 0.8083
Epoch 53/150
196/196 [=====] - 0s 2ms/step - loss: 0.4034 - acc: 0.8582 - val_loss: 0.5379 - val_acc: 0.8067
Epoch 54/150
196/196 [=====] - 0s 2ms/step - loss: 0.4015 - acc: 0.8582 - val_loss: 0.5364 - val_acc: 0.8095
Epoch 55/150
```

```
196/196 [=====] - 0s 2ms/step - loss: 0.3995 - acc: 0.8600 - val_loss: 0.5373 - val_acc: 0.8105
Epoch 56/150
196/196 [=====] - 0s 2ms/step - loss: 0.3977 - acc: 0.8602 - val_loss: 0.5358 - val_acc: 0.8055
Epoch 57/150
196/196 [=====] - 0s 2ms/step - loss: 0.3960 - acc: 0.8606 - val_loss: 0.5382 - val_acc: 0.8110
Epoch 58/150
196/196 [=====] - 0s 2ms/step - loss: 0.3942 - acc: 0.8611 - val_loss: 0.5406 - val_acc: 0.8055
Epoch 59/150
196/196 [=====] - 0s 2ms/step - loss: 0.3925 - acc: 0.8621 - val_loss: 0.5381 - val_acc: 0.8098
Epoch 60/150
196/196 [=====] - 0s 2ms/step - loss: 0.3908 - acc: 0.8622 - val_loss: 0.5363 - val_acc: 0.8092
Epoch 61/150
196/196 [=====] - 0s 2ms/step - loss: 0.3893 - acc: 0.8625 - val_loss: 0.5386 - val_acc: 0.8083
Epoch 62/150
196/196 [=====] - 0s 2ms/step - loss: 0.3877 - acc: 0.8634 - val_loss: 0.5379 - val_acc: 0.8090
Epoch 63/150
196/196 [=====] - 0s 2ms/step - loss: 0.3858 - acc: 0.8647 - val_loss: 0.5380 - val_acc: 0.8115
Epoch 64/150
196/196 [=====] - 0s 2ms/step - loss: 0.3843 - acc: 0.8645 - val_loss: 0.5382 - val_acc: 0.8112
Epoch 65/150
196/196 [=====] - 0s 2ms/step - loss: 0.3828 - acc: 0.8659 - val_loss: 0.5401 - val_acc: 0.8092
Epoch 66/150
196/196 [=====] - 0s 2ms/step - loss: 0.3811 - acc: 0.8660 - val_loss: 0.5369 - val_acc: 0.8105
Epoch 67/150
196/196 [=====] - 0s 2ms/step - loss: 0.3796 - acc: 0.8670 - val_loss: 0.5399 - val_acc: 0.8075
Epoch 68/150
196/196 [=====] - 0s 2ms/step - loss: 0.3782 - acc: 0.8671 - val_loss: 0.5390 - val_acc: 0.8108
Epoch 69/150
196/196 [=====] - 0s 2ms/step - loss: 0.3769 - acc: 0.8674 - val_loss: 0.5403 - val_acc: 0.8087
Epoch 70/150
196/196 [=====] - 0s 2ms/step - loss: 0.3755 - acc: 0.8685 - val_loss: 0.5412 - val_acc: 0.8112
Epoch 71/150
196/196 [=====] - 0s 2ms/step - loss: 0.3738 - acc:
```

```
0.8678 - val_loss: 0.5417 - val_acc: 0.8120
Epoch 72/150
196/196 [=====] - 0s 2ms/step - loss: 0.3728 - acc:
0.8691 - val_loss: 0.5419 - val_acc: 0.8077
Epoch 73/150
196/196 [=====] - 0s 2ms/step - loss: 0.3714 - acc:
0.8698 - val_loss: 0.5435 - val_acc: 0.8075
Epoch 74/150
196/196 [=====] - 0s 2ms/step - loss: 0.3702 - acc:
0.8701 - val_loss: 0.5438 - val_acc: 0.8073
Epoch 75/150
196/196 [=====] - 0s 2ms/step - loss: 0.3688 - acc:
0.8704 - val_loss: 0.5423 - val_acc: 0.8095
Epoch 76/150
196/196 [=====] - 0s 2ms/step - loss: 0.3676 - acc:
0.8709 - val_loss: 0.5413 - val_acc: 0.8090
Epoch 77/150
196/196 [=====] - 0s 2ms/step - loss: 0.3659 - acc:
0.8713 - val_loss: 0.5437 - val_acc: 0.8085
Epoch 78/150
196/196 [=====] - 0s 2ms/step - loss: 0.3648 - acc:
0.8722 - val_loss: 0.5466 - val_acc: 0.8058
Epoch 79/150
196/196 [=====] - 0s 2ms/step - loss: 0.3634 - acc:
0.8728 - val_loss: 0.5475 - val_acc: 0.8020
Epoch 80/150
196/196 [=====] - 0s 2ms/step - loss: 0.3624 - acc:
0.8721 - val_loss: 0.5484 - val_acc: 0.8045
Epoch 81/150
196/196 [=====] - 0s 2ms/step - loss: 0.3610 - acc:
0.8730 - val_loss: 0.5492 - val_acc: 0.8062
Epoch 82/150
196/196 [=====] - 0s 2ms/step - loss: 0.3600 - acc:
0.8738 - val_loss: 0.5508 - val_acc: 0.8085
Epoch 83/150
196/196 [=====] - 0s 2ms/step - loss: 0.3588 - acc:
0.8736 - val_loss: 0.5463 - val_acc: 0.8090
Epoch 84/150
196/196 [=====] - 0s 2ms/step - loss: 0.3574 - acc:
0.8750 - val_loss: 0.5458 - val_acc: 0.8065
Epoch 85/150
196/196 [=====] - 0s 2ms/step - loss: 0.3565 - acc:
0.8748 - val_loss: 0.5487 - val_acc: 0.8073
Epoch 86/150
196/196 [=====] - 0s 2ms/step - loss: 0.3551 - acc:
0.8753 - val_loss: 0.5491 - val_acc: 0.8095
Epoch 87/150
196/196 [=====] - 0s 2ms/step - loss: 0.3541 - acc:
0.8757 - val_loss: 0.5565 - val_acc: 0.8002
```

```
Epoch 88/150
196/196 [=====] - 0s 2ms/step - loss: 0.3528 - acc: 0.8763 - val_loss: 0.5490 - val_acc: 0.8077
Epoch 89/150
196/196 [=====] - 0s 2ms/step - loss: 0.3518 - acc: 0.8772 - val_loss: 0.5517 - val_acc: 0.8045
Epoch 90/150
196/196 [=====] - 0s 2ms/step - loss: 0.3505 - acc: 0.8761 - val_loss: 0.5536 - val_acc: 0.8030
Epoch 91/150
196/196 [=====] - 0s 2ms/step - loss: 0.3496 - acc: 0.8775 - val_loss: 0.5554 - val_acc: 0.8027
Epoch 92/150
196/196 [=====] - 0s 2ms/step - loss: 0.3488 - acc: 0.8776 - val_loss: 0.5531 - val_acc: 0.8075
Epoch 93/150
196/196 [=====] - 0s 2ms/step - loss: 0.3474 - acc: 0.8781 - val_loss: 0.5565 - val_acc: 0.8037
Epoch 94/150
196/196 [=====] - 0s 2ms/step - loss: 0.3466 - acc: 0.8776 - val_loss: 0.5517 - val_acc: 0.8075
Epoch 95/150
196/196 [=====] - 0s 2ms/step - loss: 0.3454 - acc: 0.8783 - val_loss: 0.5571 - val_acc: 0.8052
Epoch 96/150
196/196 [=====] - 0s 2ms/step - loss: 0.3440 - acc: 0.8792 - val_loss: 0.5562 - val_acc: 0.8045
Epoch 97/150
196/196 [=====] - 0s 2ms/step - loss: 0.3433 - acc: 0.8800 - val_loss: 0.5520 - val_acc: 0.8087
Epoch 98/150
196/196 [=====] - 0s 2ms/step - loss: 0.3420 - acc: 0.8795 - val_loss: 0.5575 - val_acc: 0.8043
Epoch 99/150
196/196 [=====] - 0s 2ms/step - loss: 0.3412 - acc: 0.8808 - val_loss: 0.5593 - val_acc: 0.8060
Epoch 100/150
196/196 [=====] - 0s 2ms/step - loss: 0.3403 - acc: 0.8808 - val_loss: 0.5591 - val_acc: 0.8033
Epoch 101/150
196/196 [=====] - 0s 2ms/step - loss: 0.3389 - acc: 0.8812 - val_loss: 0.5546 - val_acc: 0.8100
Epoch 102/150
196/196 [=====] - 0s 2ms/step - loss: 0.3386 - acc: 0.8812 - val_loss: 0.5558 - val_acc: 0.8105
Epoch 103/150
196/196 [=====] - 0s 2ms/step - loss: 0.3371 - acc: 0.8820 - val_loss: 0.5614 - val_acc: 0.8055
Epoch 104/150
```

```
196/196 [=====] - 0s 2ms/step - loss: 0.3364 - acc: 0.8820 - val_loss: 0.5615 - val_acc: 0.8070
Epoch 105/150
196/196 [=====] - 0s 2ms/step - loss: 0.3350 - acc: 0.8822 - val_loss: 0.5602 - val_acc: 0.8030
Epoch 106/150
196/196 [=====] - 0s 2ms/step - loss: 0.3342 - acc: 0.8833 - val_loss: 0.5608 - val_acc: 0.8067
Epoch 107/150
196/196 [=====] - 0s 2ms/step - loss: 0.3334 - acc: 0.8835 - val_loss: 0.5609 - val_acc: 0.8058
Epoch 108/150
196/196 [=====] - 0s 2ms/step - loss: 0.3320 - acc: 0.8830 - val_loss: 0.5593 - val_acc: 0.8075
Epoch 109/150
196/196 [=====] - 0s 2ms/step - loss: 0.3315 - acc: 0.8835 - val_loss: 0.5724 - val_acc: 0.8018
Epoch 110/150
196/196 [=====] - 0s 2ms/step - loss: 0.3302 - acc: 0.8847 - val_loss: 0.5700 - val_acc: 0.8012
Epoch 111/150
196/196 [=====] - 0s 2ms/step - loss: 0.3295 - acc: 0.8843 - val_loss: 0.5670 - val_acc: 0.8058
Epoch 112/150
196/196 [=====] - 0s 2ms/step - loss: 0.3284 - acc: 0.8837 - val_loss: 0.5662 - val_acc: 0.8055
Epoch 113/150
196/196 [=====] - 0s 2ms/step - loss: 0.3275 - acc: 0.8856 - val_loss: 0.5633 - val_acc: 0.8102
Epoch 114/150
196/196 [=====] - 0s 2ms/step - loss: 0.3265 - acc: 0.8850 - val_loss: 0.5692 - val_acc: 0.8012
Epoch 115/150
196/196 [=====] - 0s 2ms/step - loss: 0.3251 - acc: 0.8860 - val_loss: 0.5671 - val_acc: 0.8023
Epoch 116/150
196/196 [=====] - 0s 2ms/step - loss: 0.3245 - acc: 0.8863 - val_loss: 0.5731 - val_acc: 0.8027
Epoch 117/150
196/196 [=====] - 0s 2ms/step - loss: 0.3239 - acc: 0.8865 - val_loss: 0.5724 - val_acc: 0.8023
Epoch 118/150
196/196 [=====] - 0s 2ms/step - loss: 0.3223 - acc: 0.8875 - val_loss: 0.5715 - val_acc: 0.8015
Epoch 119/150
196/196 [=====] - 0s 2ms/step - loss: 0.3215 - acc: 0.8872 - val_loss: 0.5689 - val_acc: 0.8070
Epoch 120/150
196/196 [=====] - 0s 2ms/step - loss: 0.3204 - acc:
```

```
0.8883 - val_loss: 0.5696 - val_acc: 0.8050
Epoch 121/150
196/196 [=====] - 0s 2ms/step - loss: 0.3199 - acc:
0.8881 - val_loss: 0.5734 - val_acc: 0.8055
Epoch 122/150
196/196 [=====] - 0s 2ms/step - loss: 0.3191 - acc:
0.8886 - val_loss: 0.5779 - val_acc: 0.8045
Epoch 123/150
196/196 [=====] - 0s 2ms/step - loss: 0.3179 - acc:
0.8889 - val_loss: 0.5770 - val_acc: 0.7987
Epoch 124/150
196/196 [=====] - 0s 2ms/step - loss: 0.3169 - acc:
0.8894 - val_loss: 0.5767 - val_acc: 0.8010
Epoch 125/150
196/196 [=====] - 0s 2ms/step - loss: 0.3160 - acc:
0.8900 - val_loss: 0.5744 - val_acc: 0.8018
Epoch 126/150
196/196 [=====] - 0s 2ms/step - loss: 0.3151 - acc:
0.8898 - val_loss: 0.5920 - val_acc: 0.7962
Epoch 127/150
196/196 [=====] - 0s 2ms/step - loss: 0.3140 - acc:
0.8902 - val_loss: 0.5820 - val_acc: 0.8018
Epoch 128/150
196/196 [=====] - 0s 2ms/step - loss: 0.3133 - acc:
0.8905 - val_loss: 0.5797 - val_acc: 0.8055
Epoch 129/150
196/196 [=====] - 0s 2ms/step - loss: 0.3120 - acc:
0.8914 - val_loss: 0.5861 - val_acc: 0.8035
Epoch 130/150
196/196 [=====] - 0s 2ms/step - loss: 0.3111 - acc:
0.8919 - val_loss: 0.5835 - val_acc: 0.7990
Epoch 131/150
196/196 [=====] - 0s 2ms/step - loss: 0.3102 - acc:
0.8928 - val_loss: 0.5777 - val_acc: 0.8010
Epoch 132/150
196/196 [=====] - 0s 2ms/step - loss: 0.3092 - acc:
0.8922 - val_loss: 0.5835 - val_acc: 0.8087
Epoch 133/150
196/196 [=====] - 0s 2ms/step - loss: 0.3086 - acc:
0.8921 - val_loss: 0.5869 - val_acc: 0.8000
Epoch 134/150
196/196 [=====] - 0s 2ms/step - loss: 0.3074 - acc:
0.8932 - val_loss: 0.5801 - val_acc: 0.8055
Epoch 135/150
196/196 [=====] - 0s 2ms/step - loss: 0.3064 - acc:
0.8932 - val_loss: 0.5829 - val_acc: 0.8012
Epoch 136/150
196/196 [=====] - 0s 2ms/step - loss: 0.3059 - acc:
0.8931 - val_loss: 0.5826 - val_acc: 0.8058
```

```
Epoch 137/150
196/196 [=====] - 0s 2ms/step - loss: 0.3044 - acc: 0.8950 - val_loss: 0.5835 - val_acc: 0.8040
Epoch 138/150
196/196 [=====] - 0s 2ms/step - loss: 0.3036 - acc: 0.8945 - val_loss: 0.5925 - val_acc: 0.7958
Epoch 139/150
196/196 [=====] - 0s 2ms/step - loss: 0.3028 - acc: 0.8945 - val_loss: 0.5849 - val_acc: 0.8058
Epoch 140/150
196/196 [=====] - 0s 2ms/step - loss: 0.3016 - acc: 0.8959 - val_loss: 0.5870 - val_acc: 0.8033
Epoch 141/150
196/196 [=====] - 0s 2ms/step - loss: 0.3007 - acc: 0.8964 - val_loss: 0.5896 - val_acc: 0.8020
Epoch 142/150
196/196 [=====] - 0s 2ms/step - loss: 0.2998 - acc: 0.8964 - val_loss: 0.5976 - val_acc: 0.7987
Epoch 143/150
196/196 [=====] - 0s 2ms/step - loss: 0.2986 - acc: 0.8971 - val_loss: 0.5875 - val_acc: 0.8075
Epoch 144/150
196/196 [=====] - 0s 2ms/step - loss: 0.2974 - acc: 0.8970 - val_loss: 0.5898 - val_acc: 0.8023
Epoch 145/150
196/196 [=====] - 0s 2ms/step - loss: 0.2965 - acc: 0.8978 - val_loss: 0.5957 - val_acc: 0.7997
Epoch 146/150
196/196 [=====] - 0s 2ms/step - loss: 0.2958 - acc: 0.8975 - val_loss: 0.5994 - val_acc: 0.7980
Epoch 147/150
196/196 [=====] - 0s 2ms/step - loss: 0.2946 - acc: 0.8988 - val_loss: 0.6050 - val_acc: 0.7922
Epoch 148/150
196/196 [=====] - 0s 2ms/step - loss: 0.2938 - acc: 0.8985 - val_loss: 0.5948 - val_acc: 0.8033
Epoch 149/150
196/196 [=====] - 0s 2ms/step - loss: 0.2928 - acc: 0.8993 - val_loss: 0.5952 - val_acc: 0.8037
Epoch 150/150
196/196 [=====] - 0s 2ms/step - loss: 0.2918 - acc: 0.8994 - val_loss: 0.6006 - val_acc: 0.8040
```

```
results_train = bigger_data_model.evaluate(X_train_tokens_bigger, y_train_lb_bigger)
print(f'Training Loss: {results_train[0]:.3} \nTraining Accuracy: {results_train[1]:.3}
print('-----')
```

```
results_test = bigger_data_model.evaluate(X_val_tokens_bigger, y_val_lb_bigger)
print(f'Test Loss: {results_test[0]:.3} \nTest Accuracy: {results_test[1]:.3}')
```

```
1563/1563 [=====] - 1s 519us/step - loss: 0.2886 - acc: 0.9006
Training Loss: 0.289
Training Accuracy: 0.901
-----
125/125 [=====] - 0s 996us/step - loss: 0.6006 - acc: 0.8040
Test Loss: 0.601
Test Accuracy: 0.804
```

With the same amount of epochs and no regularization technique, you were able to get both better test accuracy and loss. You can still consider early stopping, L1, L2 and dropout here. It's clear that having more data has a strong impact on model performance!

## Additional Resources

- [https://github.com/susanli2016/Machine-Learning-with-Python/blob/master/Consumer\\_complaints.ipynb](https://github.com/susanli2016/Machine-Learning-with-Python/blob/master/Consumer_complaints.ipynb)
- <https://machinelearningmastery.com/dropout-regularization-deep-learning-models-keras/>
- <https://catalog.data.gov/dataset/consumer-complaint-database>

## Summary

In this lesson, you built deep learning models using a validation set and used several techniques such as L2 and L1 regularization, dropout regularization, and early stopping to improve the accuracy of your models.

## Releases

No releases published

## Packages

No packages published

---

## Contributors 7



## Languages

- Jupyter Notebook 100.0%