### **Movie Reviews**

#### 5.1

```
In [2]: import keras
        keras.__version__
Out[2]: '2.4.3'
In [3]: #import data
        from keras.datasets import imdb
        #split data
        (train data, train labels), (test data, test labels) = imdb.load data(num words=1
        Downloading data from https://storage.googleapis.com/tensorflow/tf-keras-datase
        ts/imdb.npz (https://storage.googleapis.com/tensorflow/tf-keras-datasets/imdb.n
        <__array_function__ internals>:5: VisibleDeprecationWarning: Creating an ndarra
        y from ragged nested sequences (which is a list-or-tuple of lists-or-tuples-or
        ndarrays with different lengths or shapes) is deprecated. If you meant to do th
        is, you must specify 'dtype=object' when creating the ndarray
        /opt/conda/lib/python3.8/site-packages/tensorflow/python/keras/datasets/imdb.p
        y:159: VisibleDeprecationWarning: Creating an ndarray from ragged nested sequen
        ces (which is a list-or-tuple of lists-or-tuples-or ndarrays with different len
        gths or shapes) is deprecated. If you meant to do this, you must specify 'dtype
        =object' when creating the ndarray
          x train, y train = np.array(xs[:idx]), np.array(labels[:idx])
        /opt/conda/lib/python3.8/site-packages/tensorflow/python/keras/datasets/imdb.p
        y:160: VisibleDeprecationWarning: Creating an ndarray from ragged nested sequen
        ces (which is a list-or-tuple of lists-or-tuples-or ndarrays with different len
        gths or shapes) is deprecated. If you meant to do this, you must specify 'dtype
        =object' when creating the ndarray
          x test, y test = np.array(xs[idx:]), np.array(labels[idx:])
In [4]: import numpy as np
        def vectorize_sequence(sequences, dimension=10000):
            #create all-zero matrix of shape (len(seq), dim)
            results = np.zeros((len(sequences), dimension))
            for i, seq in enumerate(sequences):
                results[i,seq]= 1.
            return results
        x train = vectorize sequence(train data)
        x_test = vectorize_sequence(test_data)
In [5]: #vectorized labels
        y_train = np.asarray(train_labels).astype('float32')
        y_test = np.asarray(test_labels).astype('float32')
```

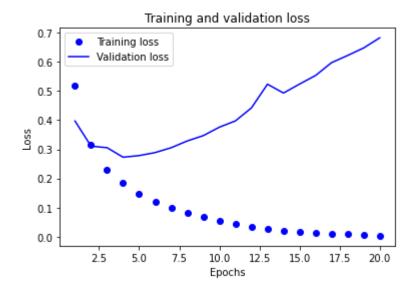
```
In [6]: from keras import models
        from keras import layers
        #building model
        model = models.Sequential()
        model.add(layers.Dense(16, activation='relu', input_shape=(10000,)))
        model.add(layers.Dense(16, activation='relu'))
        model.add(layers.Dense(1, activation='sigmoid'))
In [7]: model.compile(optimizer='rmsprop',
                      loss='binary_crossentropy',
                      metrics=['accuracy'])
In [8]: from keras import optimizers
        model.compile(optimizer=optimizers.RMSprop(lr=0.001),
                      loss='binary crossentropy',
                      metrics=['accuracy'])
In [9]: from keras import losses
        from keras import metrics
        model.compile(optimizer=optimizers.RMSprop(lr=0.001),
                      loss=losses.binary_crossentropy,
                      metrics=[metrics.binary accuracy])
```

```
Epoch 1/20
ccuracy: 0.7177 - val loss: 0.3968 - val binary accuracy: 0.8697
Epoch 2/20
30/30 [============== ] - 0s 13ms/step - loss: 0.3330 - binary_a
ccuracy: 0.9030 - val_loss: 0.3107 - val_binary_accuracy: 0.8849
ccuracy: 0.9264 - val_loss: 0.3054 - val_binary_accuracy: 0.8760
Epoch 4/20
ccuracy: 0.9410 - val_loss: 0.2728 - val_binary_accuracy: 0.8907
Epoch 5/20
ccuracy: 0.9540 - val_loss: 0.2784 - val_binary_accuracy: 0.8891
Epoch 6/20
ccuracy: 0.9671 - val_loss: 0.2888 - val_binary_accuracy: 0.8894
Epoch 7/20
ccuracy: 0.9706 - val_loss: 0.3054 - val_binary_accuracy: 0.8863
Epoch 8/20
ccuracy: 0.9796 - val_loss: 0.3285 - val_binary_accuracy: 0.8791
ccuracy: 0.9830 - val_loss: 0.3468 - val_binary_accuracy: 0.8792
Epoch 10/20
ccuracy: 0.9891 - val_loss: 0.3757 - val_binary_accuracy: 0.8764
Epoch 11/20
ccuracy: 0.9912 - val_loss: 0.3971 - val_binary_accuracy: 0.8766
Epoch 12/20
ccuracy: 0.9949 - val_loss: 0.4418 - val_binary_accuracy: 0.8686
Epoch 13/20
ccuracy: 0.9956 - val_loss: 0.5229 - val_binary_accuracy: 0.8554
Epoch 14/20
ccuracy: 0.9974 - val_loss: 0.4927 - val_binary_accuracy: 0.8690
Epoch 15/20
ccuracy: 0.9982 - val_loss: 0.5233 - val_binary_accuracy: 0.8688
```

```
In [11]: history_dict = history.history
history_dict.keys()
```

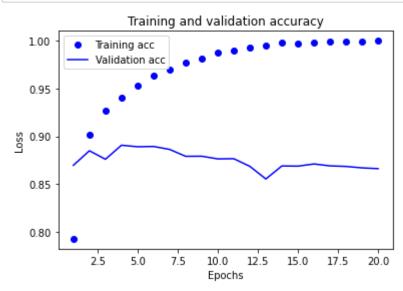
Out[11]: dict\_keys(['loss', 'binary\_accuracy', 'val\_loss', 'val\_binary\_accuracy'])

```
In [12]:
         import matplotlib.pyplot as plt
         acc = history.history['binary accuracy']
         val_acc = history.history['val_binary_accuracy']
         loss = history.history['loss']
         val loss = history.history['val loss']
         epochs = range(1, len(acc) + 1)
         # "bo" is for "blue dot"
         plt.plot(epochs, loss, 'bo', label='Training loss')
         # b is for "solid blue line"
         plt.plot(epochs, val_loss, 'b', label='Validation loss')
         plt.title('Training and validation loss')
         plt.xlabel('Epochs')
         plt.ylabel('Loss')
         plt.legend()
         plt.show()
```



```
In [13]: plt.clf() # clear figure
    acc_values = history_dict['binary_accuracy']
    val_acc_values = history_dict['val_binary_accuracy']

plt.plot(epochs, acc, 'bo', label='Training acc')
    plt.plot(epochs, val_acc, 'b', label='Validation acc')
    plt.title('Training and validation accuracy')
    plt.xlabel('Epochs')
    plt.ylabel('Loss')
    plt.legend()
```



```
In [14]: #new model using 4 epochs
       model = models.Sequential()
       model.add(layers.Dense(16, activation='relu', input shape=(10000,)))
       model.add(layers.Dense(16, activation='relu'))
       model.add(layers.Dense(1, activation='sigmoid'))
       model.compile(optimizer='rmsprop',
                  loss='binary crossentropy',
                  metrics=['accuracy'])
       model.fit(x_train, y_train, epochs=4, batch_size=512)
       results = model.evaluate(x_test, y_test)
       Epoch 1/4
       0.7375
       Epoch 2/4
       0.9062
       Epoch 3/4
       49/49 [============= ] - 0s 7ms/step - loss: 0.2009 - accuracy:
       0.9310
       Epoch 4/4
       0.9441
       782/782 [============= ] - 2s 2ms/step - loss: 0.3215 - accurac
       y: 0.8730
In [15]: #results of naive approach
       results
Out[15]: [0.32153692841529846, 0.8729599714279175]
In [16]: #using trained network to generate predictions on new data
       model.predict(x test)
Out[16]: array([[0.27416345],
             [0.9997012],
             [0.97655076],
             [0.18231839],
             [0.15986568],
             [0.7800609 ]], dtype=float32)
```

# **News classifier**

5.2

```
In [17]: #importing data set
         from keras.datasets import reuters
         #splitting data
         (train data, train labels), (test data, test labels) = reuters.load data(num word
         Downloading data from https://storage.googleapis.com/tensorflow/tf-keras-datase
         ts/reuters.npz (https://storage.googleapis.com/tensorflow/tf-keras-datasets/reu
         ters.npz)
         2113536/2110848 [============= ] - 0s Ous/step
         /opt/conda/lib/python3.8/site-packages/tensorflow/python/keras/datasets/reuter
         s.py:148: VisibleDeprecationWarning: Creating an ndarray from ragged nested seq
         uences (which is a list-or-tuple of lists-or-tuples-or ndarrays with different
         lengths or shapes) is deprecated. If you meant to do this, you must specify 'dt
         ype=object' when creating the ndarray
           x_train, y_train = np.array(xs[:idx]), np.array(labels[:idx])
         /opt/conda/lib/python3.8/site-packages/tensorflow/python/keras/datasets/reuter
         s.py:149: VisibleDeprecationWarning: Creating an ndarray from ragged nested seq
         uences (which is a list-or-tuple of lists-or-tuples-or ndarrays with different
         lengths or shapes) is deprecated. If you meant to do this, you must specify 'dt
         ype=object' when creating the ndarray
           x_test, y_test = np.array(xs[idx:]), np.array(labels[idx:])
In [18]: print(len(train data))
         print(len(test_data))
         8982
         2246
In [20]:
         #data prep using same vectorizer from 5.1
         # Our vectorized training data
         x train = vectorize sequence(train data)
         # Our vectorized test data
         x test = vectorize sequence(test data)
In [21]: def to one hot(labels, dimension=46):
             results = np.zeros((len(labels), dimension))
             for i, label in enumerate(labels):
                 results[i, label] = 1.
             return results
         # Our vectorized training labels
         one hot train labels = to one hot(train labels)
         # Our vectorized test labels
         one_hot_test_labels = to_one_hot(test_labels)
In [22]: from keras.utils.np_utils import to_categorical
         one_hot_train_labels = to_categorical(train_labels)
         one hot test labels = to categorical(test labels)
```

```
Epoch 1/20
16/16 [============= ] - 1s 32ms/step - loss: 3.1424 - accurac
y: 0.3662 - val loss: 1.7362 - val accuracy: 0.6420
Epoch 2/20
16/16 [============] - 0s 18ms/step - loss: 1.5376 - accurac
y: 0.6828 - val_loss: 1.2827 - val_accuracy: 0.7150
16/16 [================= ] - 0s 20ms/step - loss: 1.0852 - accurac
y: 0.7680 - val_loss: 1.1113 - val_accuracy: 0.7670
Epoch 4/20
16/16 [=================== ] - 0s 21ms/step - loss: 0.8161 - accurac
y: 0.8267 - val_loss: 1.0133 - val_accuracy: 0.7840
Epoch 5/20
16/16 [================= ] - 0s 19ms/step - loss: 0.6506 - accurac
y: 0.8619 - val_loss: 0.9426 - val_accuracy: 0.8110
Epoch 6/20
16/16 [============== ] - 0s 19ms/step - loss: 0.5177 - accurac
y: 0.8932 - val_loss: 0.9051 - val_accuracy: 0.8190
Epoch 7/20
y: 0.9205 - val_loss: 0.8954 - val_accuracy: 0.8220
Epoch 8/20
16/16 [============== ] - 0s 24ms/step - loss: 0.3262 - accurac
y: 0.9335 - val_loss: 0.8733 - val_accuracy: 0.8180
16/16 [================= ] - 0s 21ms/step - loss: 0.2703 - accurac
y: 0.9427 - val_loss: 0.8584 - val_accuracy: 0.8300
Epoch 10/20
16/16 [============== ] - 0s 19ms/step - loss: 0.2304 - accurac
y: 0.9495 - val_loss: 0.9349 - val_accuracy: 0.8060
Epoch 11/20
16/16 [=============== ] - 0s 18ms/step - loss: 0.1884 - accurac
y: 0.9548 - val_loss: 0.9046 - val_accuracy: 0.8180
Epoch 12/20
16/16 [============== ] - 0s 15ms/step - loss: 0.1621 - accurac
y: 0.9593 - val_loss: 0.9102 - val_accuracy: 0.8240
Epoch 13/20
16/16 [================= ] - 0s 16ms/step - loss: 0.1494 - accurac
y: 0.9594 - val_loss: 0.9272 - val_accuracy: 0.8200
Epoch 14/20
16/16 [============== ] - 0s 18ms/step - loss: 0.1373 - accurac
y: 0.9581 - val_loss: 0.9647 - val_accuracy: 0.8180
Epoch 15/20
16/16 [============== ] - 0s 16ms/step - loss: 0.1300 - accurac
y: 0.9571 - val_loss: 0.9837 - val_accuracy: 0.8130
```

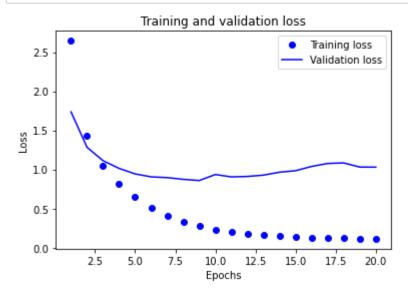
```
Epoch 16/20
16/16 [=============] - 0s 16ms/step - loss: 0.1187 - accuracy: 0.9592 - val_loss: 1.0383 - val_accuracy: 0.8050
Epoch 17/20
16/16 [=============] - 0s 15ms/step - loss: 0.1307 - accuracy: 0.9563 - val_loss: 1.0765 - val_accuracy: 0.8080
Epoch 18/20
16/16 [===============] - 0s 15ms/step - loss: 0.1142 - accuracy: 0.9598 - val_loss: 1.0835 - val_accuracy: 0.8070
Epoch 19/20
16/16 [=================] - 0s 15ms/step - loss: 0.1025 - accuracy: 0.9622 - val_loss: 1.0304 - val_accuracy: 0.8080
Epoch 20/20
16/16 [===================] - 0s 17ms/step - loss: 0.1044 - accuracy: 0.9612 - val_loss: 1.0292 - val_accuracy: 0.8100
```

```
In [26]: loss = history.history['loss']
    val_loss = history.history['val_loss']

    epochs = range(1, len(loss) + 1)

    plt.plot(epochs, loss, 'bo', label='Training loss')
    plt.plot(epochs, val_loss, 'b', label='Validation loss')
    plt.title('Training and validation loss')
    plt.xlabel('Epochs')
    plt.ylabel('Loss')
    plt.legend()

plt.show()
```



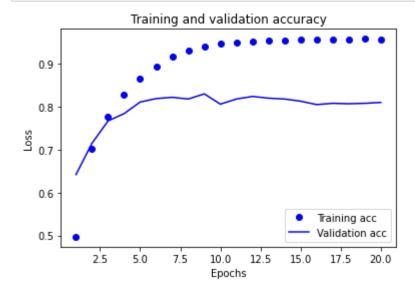
```
In [31]: plt.clf() # clear figure

acc = history.history['accuracy']

val_acc = history.history['val_accuracy']

plt.plot(epochs, acc, 'bo', label='Training acc')
plt.plot(epochs, val_acc, 'b', label='Validation acc')
plt.title('Training and validation accuracy')
plt.xlabel('Epochs')
plt.ylabel('Loss')
plt.legend()

plt.show()
```



```
In [32]: model = models.Sequential()
        model.add(layers.Dense(64, activation='relu', input shape=(10000,)))
        model.add(layers.Dense(64, activation='relu'))
        model.add(layers.Dense(46, activation='softmax'))
        model.compile(optimizer='rmsprop',
                     loss='categorical crossentropy',
                     metrics=['accuracy'])
        model.fit(partial x train,
                 partial_y_train,
                 epochs=8,
                 batch size=512,
                 validation_data=(x_val, y_val))
        results = model.evaluate(x test, one hot test labels)
        Epoch 1/8
        16/16 [=============== ] - 1s 29ms/step - loss: 3.2018 - accurac
        y: 0.3955 - val_loss: 1.7802 - val_accuracy: 0.6520
        16/16 [============] - 0s 15ms/step - loss: 1.5378 - accurac
        y: 0.7019 - val_loss: 1.3209 - val_accuracy: 0.7190
        Epoch 3/8
        16/16 [================ ] - 0s 15ms/step - loss: 1.0653 - accurac
        y: 0.7768 - val_loss: 1.1394 - val_accuracy: 0.7580
        Epoch 4/8
        16/16 [=============== ] - 0s 15ms/step - loss: 0.7995 - accurac
        y: 0.8326 - val_loss: 1.0440 - val_accuracy: 0.7740
        Epoch 5/8
        16/16 [============== ] - 0s 15ms/step - loss: 0.6331 - accurac
        y: 0.8654 - val_loss: 0.9673 - val_accuracy: 0.7870
        Epoch 6/8
        16/16 [============== ] - 0s 14ms/step - loss: 0.5182 - accurac
        y: 0.8915 - val_loss: 0.9269 - val_accuracy: 0.8030
        Epoch 7/8
        16/16 [============== ] - 0s 15ms/step - loss: 0.4055 - accurac
        y: 0.9180 - val loss: 0.9358 - val accuracy: 0.7990
        Epoch 8/8
        16/16 [=============== ] - 0s 17ms/step - loss: 0.3284 - accurac
        y: 0.9321 - val loss: 0.9160 - val accuracy: 0.8070
        0.7823
In [33]: |#results of naive model
        print(results)
        import copy
        test_labels_copy = copy.copy(test_labels)
        np.random.shuffle(test labels copy)
        float(np.sum(np.array(test_labels) == np.array(test_labels_copy))) / len(test_labels)
        [0.9781224727630615, 0.7822796106338501]
Out[33]: 0.188780053428317
```

# **Housing Prices**

#### 5.3

```
In [34]: #import data
         from keras.datasets import boston housing
         #split data
         (train_data, train_targets), (test_data, test_targets) = boston_housing.load_dat
         Downloading data from https://storage.googleapis.com/tensorflow/tf-keras-datase
         ts/boston housing.npz (https://storage.googleapis.com/tensorflow/tf-keras-datas
         ets/boston housing.npz)
         57344/57026 [=========== ] - Os Ous/step
In [35]: #data prep
         mean = train data.mean(axis=0)
         train data -= mean
         std = train_data.std(axis=0)
         train data /= std
         test data -= mean
         test data /= std
In [36]: #building model
         def build model():
             # Because we will need to instantiate
             # the same model multiple times,
             # we use a function to construct it.
             model = models.Sequential()
             model.add(layers.Dense(64, activation='relu',
                                    input shape=(train data.shape[1],)))
             model.add(layers.Dense(64, activation='relu'))
             model.add(layers.Dense(1))
             model.compile(optimizer='rmsprop', loss='mse', metrics=['mae'])
             return model
```

```
In [37]: #validation of modal
         k = 4
         num val samples = len(train data) // k
         num epochs = 100
         all scores = []
         for i in range(k):
             print('processing fold #', i)
             # Prepare the validation data: data from partition # k
             val_data = train_data[i * num_val_samples: (i + 1) * num_val_samples]
             val_targets = train_targets[i * num_val_samples: (i + 1) * num_val_samples]
             # Prepare the training data: data from all other partitions
             partial_train_data = np.concatenate(
                  [train_data[:i * num_val_samples],
                  train data[(i + 1) * num val samples:]],
                 axis=0)
             partial train targets = np.concatenate(
                  [train_targets[:i * num_val_samples],
                  train_targets[(i + 1) * num_val_samples:]],
                 axis=0)
             # Build the Keras model (already compiled)
             model = build model()
             # Train the model (in silent mode, verbose=0)
             model.fit(partial_train_data, partial_train_targets,
                       epochs=num epochs, batch size=1, verbose=0)
             # Evaluate the model on the validation data
             val mse, val mae = model.evaluate(val data, val targets, verbose=0)
             all scores.append(val mae)
         processing fold # 0
         processing fold # 1
         processing fold # 2
         processing fold # 3
In [38]: print(all scores)
         print(np.mean(all scores))
         [2.058684825897217, 2.555415391921997, 2.7275359630584717, 2.690248489379883]
         2.507971167564392
In [39]: from keras import backend as K
         # Some memory clean-up
         K.clear session()
```

```
In [43]: num epochs = 500
         all mae histories = []
         for i in range(k):
             print('processing fold #', i)
             # Prepare the validation data: data from partition # k
             val_data = train_data[i * num_val_samples: (i + 1) * num_val_samples]
             val_targets = train_targets[i * num_val_samples: (i + 1) * num_val_samples]
             # Prepare the training data: data from all other partitions
             partial_train_data = np.concatenate(
                 [train_data[:i * num_val_samples],
                  train_data[(i + 1) * num_val_samples:]],
                 axis=0)
             partial train targets = np.concatenate(
                 [train targets[:i * num val samples],
                  train_targets[(i + 1) * num_val_samples:]],
                 axis=0)
             # Build the Keras model (already compiled)
             model = build model()
             # Train the model (in silent mode, verbose=0)
             history = model.fit(partial_train_data, partial_train_targets,
                                 validation data=(val data, val targets),
                                 epochs=num_epochs, batch_size=1, verbose=0)
             #history dict = history.history
             #print(history dict.keys())
             mae history = history.history['mae']
             all_mae_histories.append(mae_history)
         processing fold # 0
         processing fold # 1
         processing fold # 2
         processing fold # 3
In [44]: average mae history = [
             np.mean([x[i]] for x in all mae histories]) for i in range(num epochs)]
```

```
In [45]: plt.plot(range(1, len(average_mae_history) + 1), average_mae_history)
    plt.xlabel('Epochs')
    plt.ylabel('Validation MAE')
    plt.show()
```

