```
In [1]:
        from tensorflow.keras.models import Sequential
         from tensorflow.keras.layers import Dense, Embedding, GRU, LSTM, Dropout
         from tensorflow.keras import utils
         from tensorflow.keras.preprocessing.sequence import pad_sequences
         from tensorflow.keras.preprocessing.text import Tokenizer
         from tensorflow.keras.callbacks import ModelCheckpoint, ReduceLROnPlateau
         import pandas as pd
         import numpy as np
         import matplotlib.pyplot as plt
In [2]: | from google.colab import drive
         drive.mount('/content/drive')
         Mounted at /content/drive
In [3]: num words = 10000
         max_review_len = 200
In [4]: | train = pd.read csv('/content/drive/MyDrive/A6/train.csv',
                               header=None,
                               names=['Label', 'Review'])
         test = pd.read_csv('/content/drive/MyDrive/A6/test.csv',
                               header=None,
                               names=['Label', 'Review'])
In [5]:
         train
Out[5]:
                  Label
                                                            Review
               0
                      1
                             Unfortunately, the frustration of being Dr. Go...
               1
                      2
                           Been going to Dr. Goldberg for over 10 years. ...
               2
                      1
                           I don't know what Dr. Goldberg was like before...
               3
                      1
                             I'm writing this review to give you a heads up...
                      2
                              All the food is great here. But the best thing...
          559995
                      2
                        Ryan was as good as everyone on yelp has claim...
          559996
                      2
                          Professional \nFriendly\nOn time AND affordabl...
          559997
                      1
                         Phone calls always go to voicemail and message...
          559998
                      1
                            Looks like all of the good reviews have gone t...
          559999
                      2
                           Ryan Rocks! I called him this morning for some...
```

```
In [6]: y_train, y_test = train['Label'] - 1, test['Label'] - 1
In [7]: reviews = train['Review']
        reviews[:5]
Out[7]: 0
              Unfortunately, the frustration of being Dr. Go...
              Been going to Dr. Goldberg for over 10 years. ...
        2
              I don't know what Dr. Goldberg was like before...
              I'm writing this review to give you a heads up...
              All the food is great here. But the best thing...
        Name: Review, dtype: object
In [8]: | tokenizer = Tokenizer(num_words=num_words)
        tokenizer.fit_on_texts(reviews)
        tokenizer.word_index
Out[8]: {'the': 1,
          'and': 2,
         'i': 3,
          'to': 4,
          'a': 5,
          'was': 6,
          'of': 7,
          'it': 8,
          'for': 9,
         'in': 10,
         'is': 11,
          'n': 12,
          'that': 13,
          'my': 14,
          'we': 15,
          'this': 16,
          'but': 17,
          'with': 18,
          'they': 19,
In [9]: | sequences = tokenizer.texts_to_sequences(reviews)
```

```
In [10]: index = 42
print(reviews[index])
print(sequences[index])
```

Some of the worst pizza I've ever had. We used a coupon from the paper f or a 2 topping 8 cut Sicilian. First of all the pizza wasn't even cut thr ough, and the sad attempt at cutting was so uneven that 4 of the slices w ere about an inch wide, while the others were about 4\" each. The topping s were scarce, they used mini pepperoni and put maybe 8 on the whole pizz a. The onions were huge chunks and the mushrooms were straight from a ca n. The worst part though was the thick doughy crust that tasted more like a fishy sourdough roll. I'm serious... It was so noticeable that it made me wonder if the dough was bad or if they for some weird reason put fish sauce in it. It was gross. \n\nWe also ordered steak and Italian hoagies. The veggies were old and wilted, and there was no dressing on either. The Italian had deli meat that was clearly bottom of the line and not very ge nerous. The \"steak\" (if you an call it that) was greyish instead of bro wn and looked like it was a processed meat chopped into pieces. No flavor or seasoning and the texture was reminiscent of spam. It was so bad that I only ate 1/4 of it and tossed the rest. \n\nI have ordered from here in the past and always been disappointed. I thought I would give them anothe r try since I'd never ordered a Sicilian pizza from there. What a mistak e. I will never order from them again! [67, 7, 1, 384, 171, 112, 151, 24, 15, 299, 5, 1073, 50, 1, 1090, 9, 5, 1 15, 2403, 479, 501, 7487, 105, 7, 37, 1, 171, 168, 74, 501, 292, 2, 1, 85 5, 1992, 25, 2543, 6, 27, 4979, 13, 202, 7, 1, 1247, 26, 56, 59, 2584, 17 12, 166, 1, 674, 26, 56, 202, 304, 1, 1027, 26, 9651, 19, 299, 1301, 215 0, 2, 271, 273, 479, 21, 1, 375, 171, 1, 949, 26, 323, 2138, 2, 1, 1174, 26, 1193, 50, 5, 75, 1, 384, 377, 195, 6, 1, 1076, 4608, 881, 13, 321, 7 3, 38, 5, 3163, 4442, 527, 93, 1610, 8, 6, 27, 6270, 13, 8, 138, 32, 145 9, 39, 1, 2046, 6, 143, 49, 39, 19, 9, 67, 892, 435, 271, 383, 181, 10, 8, 8, 6, 982, 12, 261, 84, 89, 309, 2, 678, 1, 995, 26, 254, 2, 4716, 2, 33, 6, 57, 961, 21, 405, 1, 678, 24, 1726, 258, 13, 6, 935, 1150, 7, 1, 2 60, 2, 22, 42, 1349, 1, 309, 39, 20, 59, 290, 8, 13, 6, 389, 7, 1285, 2, 269, 38, 8, 6, 5, 3871, 258, 2196, 178, 767, 57, 279, 49, 1789, 2, 1, 128 0, 6, 5415, 7, 5604, 8, 6, 27, 143, 13, 3, 68, 506, 200, 202, 7, 8, 2, 30 00, 1, 615, 12, 85, 23, 89, 50, 45, 10, 1, 546, 2, 128, 69, 337, 3, 252, 3, 53, 163, 76, 157, 121, 167, 364, 88, 89, 5, 7487, 171, 50, 33, 60, 5, 1002, 3, 66, 88, 96, 50, 76, 108]

```
In [11]: print(tokenizer.word_index['some'])
    print(tokenizer.word_index['of'])
    print(tokenizer.word_index['the'])
    print(tokenizer.word_index['worst'])
```

67

7

1

384

```
x_train = pad_sequences(sequences, maxlen=max_review_len)
In [12]:
          x_train[0]
Out[12]: array([
                     0,
                            0,
                                  0,
                                         0,
                                                0,
                                                      0,
                                                             0,
                                                                   0,
                                                                          0,
                                                                                0,
                                                                                       0,
                                  0,
                                         0,
                                                                                0,
                                               0,
                            0,
                                                                                       0,
                     0,
                                                      0,
                                                             0,
                                                                   0,
                                                                          0,
                     0,
                            0,
                                  0,
                                         0,
                                               0,
                                                      0,
                                                             0,
                                                                          0,
                                                                                0,
                                                                                       0,
                                                                   0,
                            0,
                     0,
                                  0,
                                         0,
                                               0,
                                                      0,
                                                             0,
                                                                          0,
                                                                                0,
                                                                                       0,
                                                                   0,
                     0,
                            0,
                                  0,
                                         0,
                                               0,
                                                      0,
                                                             0,
                                                                   0,
                                                                          0,
                                                                                0,
                                                                                       0,
                     0,
                            0,
                                  0,
                                         0,
                                               0,
                                                      0,
                                                             0,
                                                                          0,
                                                                                0,
                                                                   0,
                                                                                       0,
                            0,
                     0,
                                  0,
                                         0,
                                               0,
                                                      0,
                                                             0,
                                                                          0,
                                                                                0,
                                                                   0,
                                                                                       0,
                     0,
                                               0,
                            0,
                                  0,
                                         0,
                                                      0,
                                                             0,
                                                                   0,
                                                                          0,
                                                                                0,
                                                                                       0,
                                707,
                                         1, 4339,
                                                      7,
                                                                 939, 1665,
                                                                               11,
                                                                                       5,
                     0,
                            0,
                                                          173,
                                                                               80, 3128,
                  2161,
                            7,
                                      137,
                                             112,
                                                     24,
                                                            18,
                                                                  27,
                                                                        189,
                                  1,
                                                    129,
                                                                 519,
                    10, 2636,
                                 34, 1416,
                                             462,
                                                             8,
                                                                         13,
                                                                              177,
                                                                                     129,
                           88, 3812,
                   803,
                                             446,
                                                      8,
                                                          392,
                                                                 841,
                                                                        115,
                                                                              360,
                                         1,
                                                                                       7,
                                              59, 1266,
                  4136, 1438,
                                  4,
                                        43,
                                                          118,
                                                                  95,
                                                                         46,
                                                                                9,
                                                                                      13,
                    49, 1687,
                                                                  23,
                                                                        563,
                                  4,
                                      418,
                                              18,
                                                      8,
                                                             3,
                                                                              178,
                                                                                      16,
                   460,
                           18,
                               189,
                                        80, 3128,
                                                      2,
                                                                  40,
                                                                         78,
                                                                               43,
                                                                                       8,
                                                             3,
                    20,
                           23,
                                644, 1343,
                                              20,
                                                     23, 3252,
                                                                  18, 2921,
                                                                              637,
                                                                                     227,
                   430,
                         476, 4312,
                                             446,
                                                                  22,
                                         1,
                                                     63,
                                                             2,
                                                                        215,
                                                                                1,
                                                                                      63,
                    18, 2269,
                                             249,
                                                                  23,
                                 13,
                                         3,
                                                     13,
                                                             3,
                                                                          4,
                                                                              163,
                                                                                     939,
                   115, 237], dtype=int32)
          model = Sequential()
In [13]:
          model.add(Embedding(num_words, 64, input_length=max_review_len))
          model.add(GRU(128))
          model.add(Dense(1, activation='sigmoid'))
In [14]: model.compile(optimizer='adam',
                          loss='binary_crossentropy',
                         metrics=['accuracy'])
In [15]: model save path = '/content/best model.h5'
          checkpoint_callback = ModelCheckpoint(model_save_path,
                                                    monitor='val accuracy',
                                                    save_best_only=True,
                                                    verbose=1)
```

```
In [16]: history = model.fit(x_train,
                       y_train,
                       epochs=5,
                       batch size=128,
                       validation split=0.1,
                       callbacks=[checkpoint_callback])
       Epoch 1/5
       3938/3938 [=========== ] - ETA: 0s - loss: 0.1882 - acc
       uracy: 0.9245
       Epoch 00001: val_accuracy improved from -inf to 0.95968, saving model to
       /content/best model.h5
       3938/3938 [============= ] - 164s 40ms/step - loss: 0.188
       2 - accuracy: 0.9245 - val_loss: 0.1060 - val_accuracy: 0.9597
       Epoch 2/5
       uracy: 0.9546
       Epoch 00002: val_accuracy improved from 0.95968 to 0.96548, saving model
       to /content/best model.h5
       3938/3938 [=========== ] - 158s 40ms/step - loss: 0.118
       5 - accuracy: 0.9546 - val_loss: 0.0918 - val_accuracy: 0.9655
       Epoch 3/5
       uracy: 0.9638
       Epoch 00003: val_accuracy did not improve from 0.96548
       3938/3938 [============ ] - 158s 40ms/step - loss: 0.096
       4 - accuracy: 0.9638 - val loss: 0.0912 - val accuracy: 0.9654
       Epoch 4/5
       uracy: 0.9709
       Epoch 00004: val_accuracy did not improve from 0.96548
       3938/3938 [============ ] - 158s 40ms/step - loss: 0.079
       2 - accuracy: 0.9709 - val_loss: 0.0970 - val_accuracy: 0.9643
       Epoch 5/5
       3938/3938 [============ ] - ETA: 0s - loss: 0.0633 - acc
       uracy: 0.9772
       Epoch 00005: val_accuracy did not improve from 0.96548
       3938/3938 [============== ] - 158s 40ms/step - loss: 0.063
```

3 - accuracy: 0.9772 - val loss: 0.1061 - val accuracy: 0.9641

```
In [17]: model.load weights(model save path)
```

```
test_sequences = tokenizer.texts_to_sequences(test['Review'])
         x_test = pad_sequences(test_sequences, maxlen=max_review_len)
         x_test[0]
Out[18]: array([
                   0,
                          0,
                                0,
                                      0,
                                            0,
                                                  0,
                                                        0,
                                                              0,
                                                                    0,
                                                                          0,
                                                                                 0,
                   0,
                          0,
                                0,
                                      0,
                                            0,
                                                  0,
                                                        0,
                                                              0,
                                                                    0,
                                                                          0,
                                                                                 0,
                   0,
                          0,
                                0,
                                      0,
                                            0,
                                                  0,
                                                        0,
                                                                    0,
                                                                          0,
                                                                                 0,
                                                              0,
                                            0,
                   0,
                          0,
                                0,
                                      0,
                                                  0,
                                                        0,
                                                              0,
                                                                    0,
                                                                          0,
                                                                                 0,
                   0,
                          0,
                                0,
                                      0,
                                            0,
                                                        0,
                                                                    0,
                                                                          0,
                                                  0,
                                                              0,
                                                                                0,
                                                        0,
                                0,
                                                  0,
                                                              0,
                   0,
                          0,
                                      0,
                                            0,
                                                                    0,
                                                                                0,
                   0,
                                            0,
                         0,
                                0,
                                      0,
                                                  0,
                                                        0,
                                                              0,
                                                                    0,
                                                                                0,
                   0,
                         0, 8158,
                                      4,
                                           80,
                                                        3,
                                                             23, 1361, 1547,
                                                336,
                                                                                56,
                   1,
                        52,
                              49,
                                      1,
                                          256,
                                                  3,
                                                       23,
                                                             69,
                                                                  287, 1837,
                                                                                52,
                  45,
                         9,
                                1,
                                    546,
                                          148,
                                                274,
                                                      165,
                                                             2, 1221,
                                                                                14,
                                                      257,
                                                            511,
                  137,
                        18,
                              289,
                                    38, 8346, 2311,
                                                                   28, 1167,
                                                                                2,
                              568,
                                    520, 1426,
                  113,
                        60,
                                                 16,
                                                       11,
                                                             44,
                                                                   31,
                                                                         13,
                                                                                3,
                         22,
                              249,
                                                            694, 2235,
                                                                          7,
                                            3, 144,
                                                      173,
                  82,
                                     38,
                                                                               40,
                         7,
                                     80, 2540, 5951,
                  72,
                              14,
                                                       23,
                                                             69,
                                                                    9,
                                                                                14,
                                                                         21,
                   7, 1558,
                                2,
                                     23, 1696,
                                                 14, 1894, 1775,
                                                                  470,
                                                                         17,
                                                                                45,
                  14,
                                2, 1346, 5923,
                                                 95,
                                                       37,
                                                                   91, 1024,
                                                                                2,
                        52,
                                                             69,
                                            4, 1480, 1037,
                  311,
                                                                   40, 4124,
                        48,
                                4,
                                     32,
                                                             19,
                                                                                1,
                                                             99,
                                                                    8,
                  317,
                        127,
                                    562,
                                            5, 203,
                                                      114,
                                                                         90,
                                8,
                                                                                10,
                 1168,
                        274], dtype=int32)
In [19]:
         scores = model.evaluate(x_test, y_test, verbose=1)
         print("The percent of correct answers:", round(scores[1] * 100, 4))
         - accuracy: 0.9569
         The percent of correct answers: 95.6868
 In [ ]:
```