Income_Prediction_Report

April 15, 2020

1 Income prediction based on census data

1.1 Introduction

The prediction task is to determine whether a person makes over 50K a year.

In this report, I will use deep neural network (DNN model).

1.2 Dataset

This dataset is downloaded from Kaggle.

Data recource: UCI machine learning repository

Now, let's import some libraries and our dataset.

[3]:
```
# Import some libraries
import pandas as pd
import numpy as np
import matplotlib as mlt
import matplotlib.pyplot as plt
import seaborn as sns
```

[4]:
```
# Import datasets
train_data=pd.read_csv('/Users/Stylewsxcde991/Desktop/     /qbs-competition-1/
↪data/train.csv',index_col=0)
X_test=pd.read_csv('/Users/Stylewsxcde991/Desktop/     /qbs-competition-1/data/
↪test.csv',index_col=0)
```

Now let's look some basic information of our dataset.

[5]:
```
# Some basic information of training data
print('The shape of training data: ' + str(train_data.shape))
print('')
print('The shape of training data: ' + str(X_test.shape))
print('')
print('Basic information of our training data: ')
print(train_data.info())
print('')
```

```python
print('Basic information of our testing data: ')
print(X_test.info())
```

```
The shape of training data: (29514, 15)

The shape of training data: (19328, 14)

Basic information of our training data:
<class 'pandas.core.frame.DataFrame'>
Int64Index: 29514 entries, 2 to 48841
Data columns (total 15 columns):
 #   Column          Non-Null Count  Dtype 
---  ------          --------------  ----- 
 0   Age             29514 non-null  int64 
 1   Workclass       27665 non-null  object
 2   fnlwgt          29514 non-null  int64 
 3   Education       29514 non-null  object
 4   Education_Num   29514 non-null  int64 
 5   Martial_Status  29514 non-null  object
 6   Occupation      27657 non-null  object
 7   Relationship    29514 non-null  object
 8   Race            29514 non-null  object
 9   Sex             29514 non-null  object
 10  Capital_Gain    29514 non-null  int64 
 11  Capital_Loss    29514 non-null  int64 
 12  Hours_per_week  29514 non-null  int64 
 13  Country         28988 non-null  object
 14  Target          29514 non-null  int64 
dtypes: int64(7), object(8)
memory usage: 3.6+ MB
None

Basic information of our testing data:
<class 'pandas.core.frame.DataFrame'>
Int64Index: 19328 entries, 1 to 48842
Data columns (total 14 columns):
 #   Column          Non-Null Count  Dtype 
---  ------          --------------  ----- 
 0   Age             19328 non-null  int64 
 1   Workclass       18378 non-null  object
 2   fnlwgt          19328 non-null  int64 
 3   Education       19328 non-null  object
 4   Education_Num   19328 non-null  int64 
 5   Martial_Status  19328 non-null  object
 6   Occupation      18376 non-null  object
 7   Relationship    19328 non-null  object
 8   Race            19328 non-null  object
```

```
 9   Sex             19328 non-null  object
 10  Capital_Gain    19328 non-null  int64
 11  Capital_Loss    19328 non-null  int64
 12  Hours_per_week  19328 non-null  int64
 13  Country         18997 non-null  object
dtypes: int64(6), object(8)
memory usage: 2.2+ MB
None
```

According to above information, the shape of training data is (29514, 15) and the shape of test data is (19328, 14).

Furthermore, notice that we have missing data problem in our training dataset and test dataset (there are Null value in some features).

In particular, we have to deal with the missing data problem of 'Workclass', 'Occupation', 'Country' in our training dataset and testing dataset. We can deal with this problem by replacing all Null value with 'unknown'.

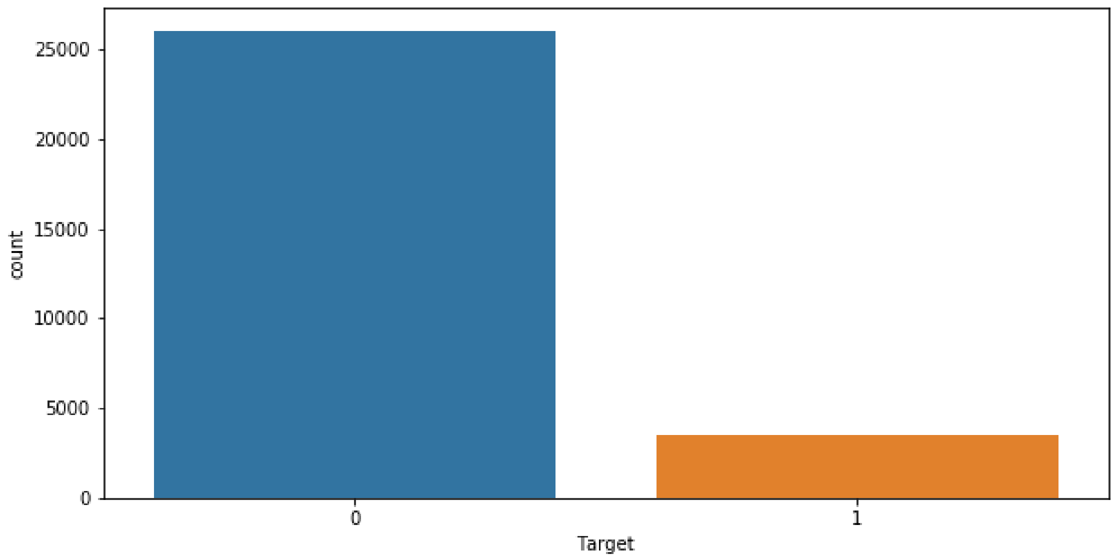
[6]:
```python
# Deal with missing data
train_data.Workclass=train_data.Workclass.fillna('unknown')
train_data.Occupation=train_data.Occupation.fillna('unknown')
train_data.Country=train_data.Country.fillna('unknown')
X_test.Workclass=X_test.Workclass.fillna('unknown')
X_test.Occupation=X_test.Occupation.fillna('unknown')
X_test.Country=X_test.Country.fillna('unknown')
```

Now let's see the proportion of each target class (make over 50k a year or not).

[7]:
```python
# The proportion of each target class
NotOver50k,Over50k = train_data.Target.value_counts()
print(f'NotOver50k {NotOver50k}')
print(f'Over50k {Over50k}')
print(f'Over50k proportion {round((100*Over50k/(Over50k+NotOver50k)),2)}%')
plt.figure(figsize=(10,5))
sns.countplot(train_data['Target'])
```

```
NotOver50k 26008
Over50k 3506
Over50k proportion 11.88%
```

[7]:
```
<matplotlib.axes._subplots.AxesSubplot at 0x11c7ff710>
```



According to above calculation and plot, there are only 11.88% samples makes over 50K a year.

Therefore, our training dataset is quiet imbalanced.

Now, let's do explorative data analysis for numerical features in our training dataset.

As following, we analyze the correlation coefficients between our numerical features.

[8]:
```
# EDA for numerical features
# data.corr()
plt.figure(figsize=(10,8))
sns.heatmap(train_data.corr(),cmap='Accent',annot=True)
plt.title('Heatmap showing correlations between numerical data')
```

[8]:
```
Text(0.5, 1, 'Heatmap showing correlations between numerical data')
```

One thing to note is that the correlation coefficient between 'fnlwgt' and our target is quiet small (which is -0.01).

Therefore, I don't consider 'fnlwgt' in my NN models.

(In fact, I have tried to incorporate 'fnlwgt' in my NN models and got really bad results.)

In addition, because I think 'Education' and 'Education_Num' contain the same information, I only use 'Education_Num' in my NN models.

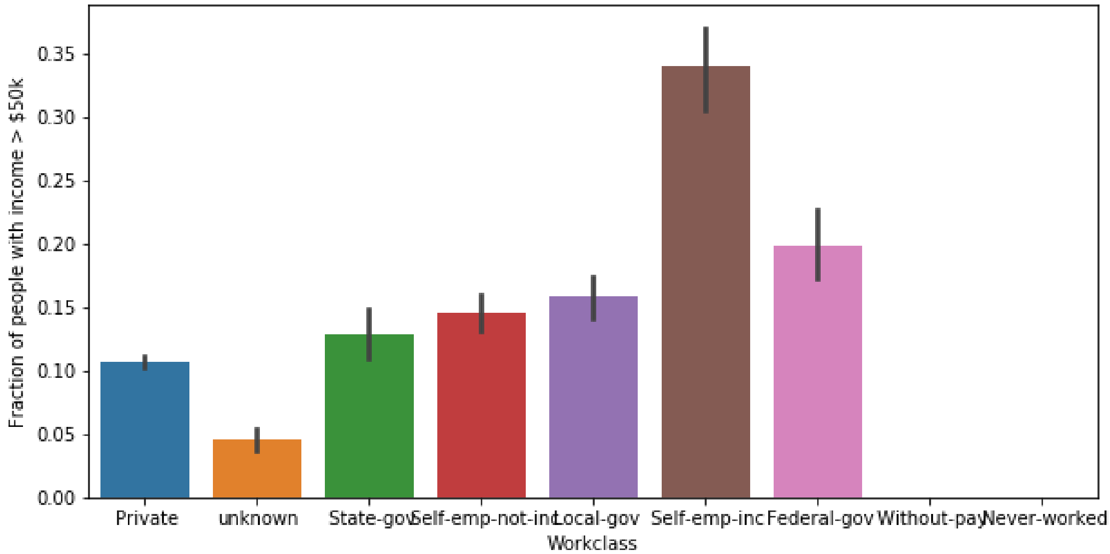
Now, let's do the explorative data analysis for some categorical features.

[9]:
```
# Drop 'fnlwgt' & 'Education'
train_data = train_data.drop(columns=['fnlwgt','Education'])
X_test = X_test.drop(columns=['fnlwgt','Education'])
```

[10]:
```
# EDA for categorical features
plt.figure(figsize=(10,5))
ax = sns.barplot(x='Workclass',y='Target',data=train_data)
```

```
ax.set(ylabel='Fraction of people with income > $50k')
```

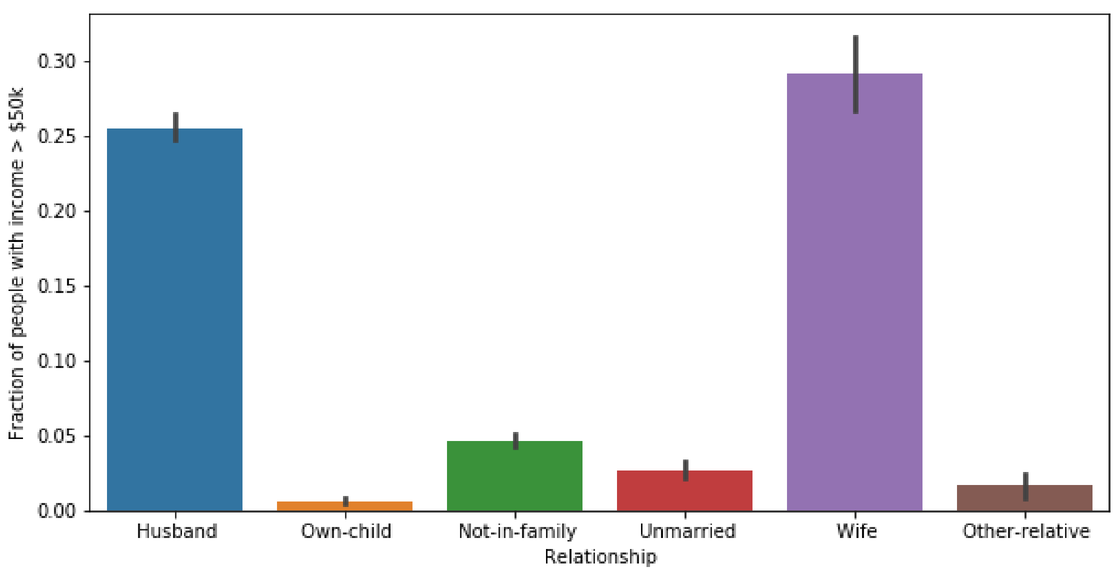
[10]: [Text(0, 0.5, 'Fraction of people with income > $50k')]



As above, people who are 'Self-emp-inc' are more likely makes over 50K a year.

[11]:
```
plt.figure(figsize=(10,5))
ax = sns.barplot(x='Relationship',y='Target',data=train_data)
ax.set(ylabel='Fraction of people with income > $50k')
```

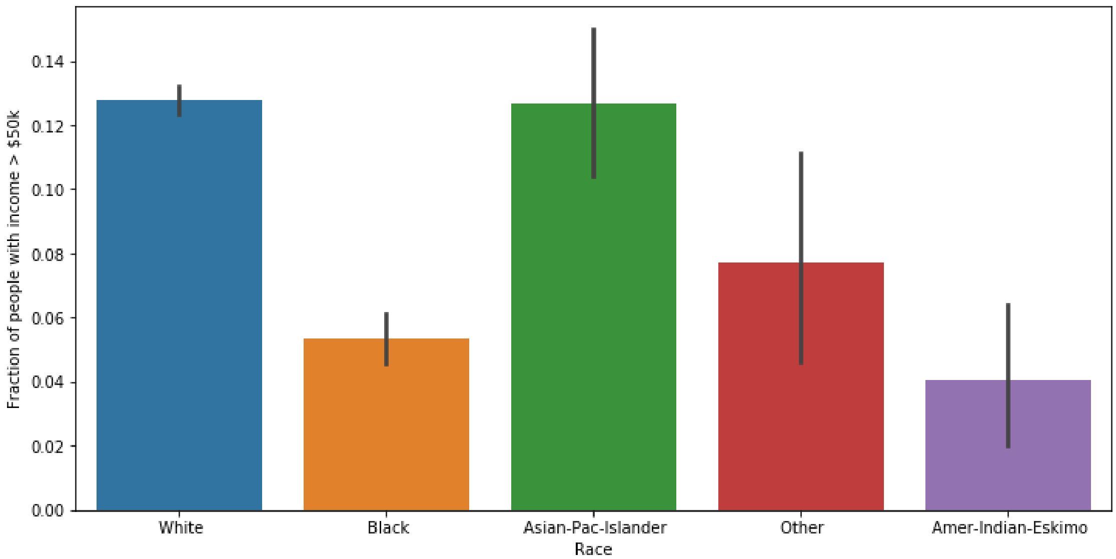
[11]: [Text(0, 0.5, 'Fraction of people with income > $50k')]



As above, 'Husband' and 'Wife' are more likely makes over 50K a year.

[12]:
```
plt.figure(figsize=(12,6))
ax=sns.barplot(x='Race',y='Target',data=train_data)
ax.set(ylabel='Fraction of people with income > $50k')
```

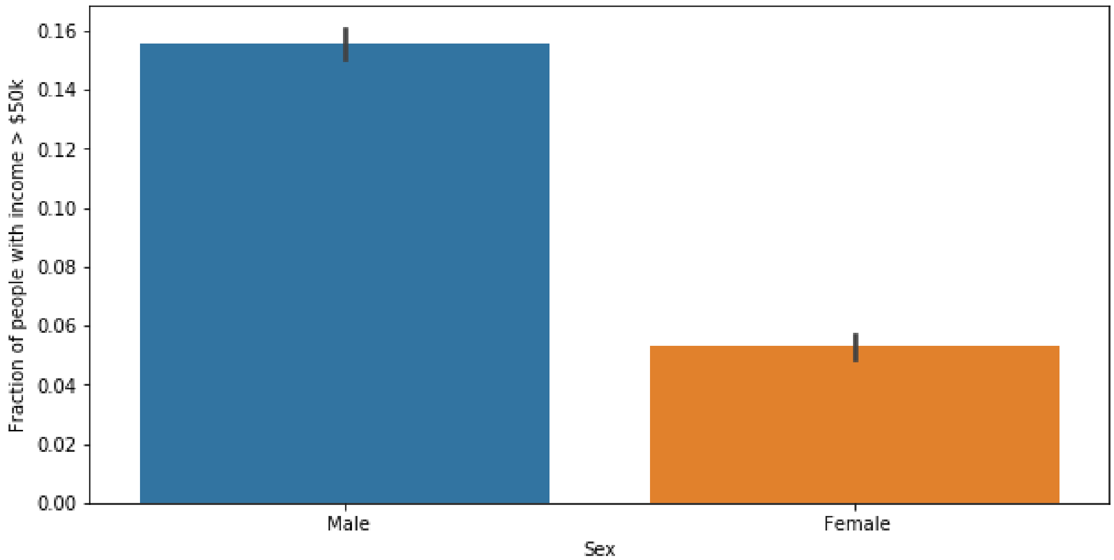
[12]: [Text(0, 0.5, 'Fraction of people with income > $50k')]



As above, 'White' and 'Asian-Pac-Islander Race' have higher proportion of people who make over 50K a year.

[13]:
```
plt.figure(figsize=(10,5))
ax = sns.barplot(x='Sex',y='Target',data=train_data)
ax.set(ylabel='Fraction of people with income > $50k')
```

[13]: [Text(0, 0.5, 'Fraction of people with income > $50k')]



As above, 'Male' group has higher proportion of people who make over 50K a year.

Now, we split our training data into features (X) and label (y).

[14]:
```
# Split our train_data
X_train=train_data.iloc[:,:-1]
y_train=train_data.iloc[:,-1]
```

Now, we have to deal with the issue of categorical features.

In our training data and testing data, we have many categorical features ('Workclass','Martial_Status', 'Education', 'Occupation','Relationship','Race','Sex','Country'). Because our NN models can only deal with numbers, we have to encode these categorical features into numbers.

In fact, there are many different ways to encode categorical features. The method I used is so-called 'One-hot encoding' (as following).

[15]:
```
# Use 'One-hot encoding' to encode categorical features.
X = X_train.append(X_test)
X = pd.get_dummies(X)
X_train = X[:29514]
X_test = X[29514:]
```

Now, we transform our datasets from dataframes to arrays, so we can feed them into NN models.

In addition, we use 7500 samples in our training dataset to be our validation set and use other samples to train our NN models.

[16]:
```
# Change dataframes to arrays
X_train = np.asarray(X_train)
```

```
X_test = np.asarray(X_test)
y_train = np.asarray(y_train).astype('float32')

# validation set
X_valid = X_train[:7500]
partial_X_train = X_train[7500:]
y_valid = y_train[:7500]
partial_y_train = y_train[7500:]
```

So far, we have dealed with all issues of data pre-processing.

Now, we can start to build our NN models.

DL model draft:

3 hidden layers in this model.

The first hidden layer: 16 units with 'relu' activation function.

The second hidden layer: 16 units with 'relu' activation function.

The third hidden layer: 1 unit with 'sigmoid' activation function.

I choose the 'rmsprop' optimizer, 'binary_crossentropy' loss function, and the 'accuracy' metrics.

parameter initialization:

I use 200 epochs to train my model. The batch_size of my model is 512.

parameter tuning:

Because I think this model did well in my training dataset, I didn't tune it's parameters.

To begin with, we build our NN model with very simple structure as following.

[17]:
```
# Construct our model
from keras import models
from keras import layers

model = models.Sequential()
model.add(layers.Dense(16, activation='relu'))
model.add(layers.Dense(16, activation='relu'))
model.add(layers.Dense(1, activation='sigmoid'))

model.compile(optimizer='rmsprop',
              loss='binary_crossentropy',
              metrics=['accuracy'])
```

```
Using TensorFlow backend.
```

Now we can start to fit our NN model and record all information in 'history'.

[18]:
```
# Iterate on your training data by calling the fit() method of your model
history = model.fit(partial_X_train,
```

```
                    partial_y_train,
                    epochs=200,
                    batch_size=512,
                    validation_data=(X_valid, y_valid))
```

```
Train on 22014 samples, validate on 7500 samples
Epoch 1/200
22014/22014 [==============================] - 0s 16us/step - loss: 1.8925 -
accuracy: 0.8715 - val_loss: 0.5837 - val_accuracy: 0.8764
Epoch 2/200
22014/22014 [==============================] - 0s 7us/step - loss: 0.5647 -
accuracy: 0.8739 - val_loss: 1.2489 - val_accuracy: 0.8592
Epoch 3/200
22014/22014 [==============================] - 0s 10us/step - loss: 0.6010 -
accuracy: 0.8739 - val_loss: 0.3940 - val_accuracy: 0.8836
Epoch 4/200
22014/22014 [==============================] - 0s 8us/step - loss: 0.5143 -
accuracy: 0.8804 - val_loss: 0.3838 - val_accuracy: 0.8935
Epoch 5/200
22014/22014 [==============================] - 0s 11us/step - loss: 0.5509 -
accuracy: 0.8828 - val_loss: 0.7664 - val_accuracy: 0.8792
Epoch 6/200
22014/22014 [==============================] - 0s 10us/step - loss: 0.4745 -
accuracy: 0.8846 - val_loss: 0.3215 - val_accuracy: 0.8963
Epoch 7/200
22014/22014 [==============================] - 0s 8us/step - loss: 0.5098 -
accuracy: 0.8823 - val_loss: 0.2977 - val_accuracy: 0.8944
Epoch 8/200
22014/22014 [==============================] - 0s 7us/step - loss: 0.4519 -
accuracy: 0.8816 - val_loss: 0.5431 - val_accuracy: 0.8811
Epoch 9/200
22014/22014 [==============================] - 0s 7us/step - loss: 0.4244 -
accuracy: 0.8857 - val_loss: 0.3005 - val_accuracy: 0.8884
Epoch 10/200
22014/22014 [==============================] - 0s 7us/step - loss: 0.4429 -
accuracy: 0.8853 - val_loss: 0.7937 - val_accuracy: 0.8813
Epoch 11/200
22014/22014 [==============================] - 0s 8us/step - loss: 0.4647 -
accuracy: 0.8815 - val_loss: 0.3102 - val_accuracy: 0.8960
Epoch 12/200
22014/22014 [==============================] - 0s 7us/step - loss: 0.4352 -
accuracy: 0.8853 - val_loss: 0.4118 - val_accuracy: 0.8893
Epoch 13/200
22014/22014 [==============================] - 0s 7us/step - loss: 0.4972 -
accuracy: 0.8850 - val_loss: 0.2706 - val_accuracy: 0.9003
Epoch 14/200
22014/22014 [==============================] - 0s 7us/step - loss: 0.4196 -
```

```
accuracy: 0.8874 - val_loss: 0.2990 - val_accuracy: 0.8953
Epoch 15/200
22014/22014 [==============================] - 0s 7us/step - loss: 0.4542 -
accuracy: 0.8877 - val_loss: 0.4380 - val_accuracy: 0.8868
Epoch 16/200
22014/22014 [==============================] - 0s 7us/step - loss: 0.4061 -
accuracy: 0.8869 - val_loss: 0.3144 - val_accuracy: 0.9009
Epoch 17/200
22014/22014 [==============================] - 0s 7us/step - loss: 0.4080 -
accuracy: 0.8890 - val_loss: 0.2770 - val_accuracy: 0.8965
Epoch 18/200
22014/22014 [==============================] - 0s 7us/step - loss: 0.4145 -
accuracy: 0.8923 - val_loss: 0.3009 - val_accuracy: 0.8992
Epoch 19/200
22014/22014 [==============================] - 0s 6us/step - loss: 0.4015 -
accuracy: 0.8902 - val_loss: 0.5950 - val_accuracy: 0.8859
Epoch 20/200
22014/22014 [==============================] - 0s 7us/step - loss: 0.4108 -
accuracy: 0.8911 - val_loss: 0.2868 - val_accuracy: 0.8899
Epoch 21/200
22014/22014 [==============================] - 0s 10us/step - loss: 0.3739 -
accuracy: 0.8938 - val_loss: 0.2728 - val_accuracy: 0.8999
Epoch 22/200
22014/22014 [==============================] - 0s 12us/step - loss: 0.4680 -
accuracy: 0.8901 - val_loss: 0.2656 - val_accuracy: 0.9069
Epoch 23/200
22014/22014 [==============================] - 0s 7us/step - loss: 0.3571 -
accuracy: 0.8964 - val_loss: 0.5500 - val_accuracy: 0.8708
Epoch 24/200
22014/22014 [==============================] - 0s 8us/step - loss: 0.3791 -
accuracy: 0.8917 - val_loss: 0.5286 - val_accuracy: 0.8760
Epoch 25/200
22014/22014 [==============================] - 0s 11us/step - loss: 0.3957 -
accuracy: 0.8942 - val_loss: 0.3088 - val_accuracy: 0.9031
Epoch 26/200
22014/22014 [==============================] - 0s 8us/step - loss: 0.3555 -
accuracy: 0.8954 - val_loss: 0.4018 - val_accuracy: 0.8788
Epoch 27/200
22014/22014 [==============================] - 0s 10us/step - loss: 0.3866 -
accuracy: 0.8943 - val_loss: 0.4961 - val_accuracy: 0.8903
Epoch 28/200
22014/22014 [==============================] - 0s 9us/step - loss: 0.3884 -
accuracy: 0.8936 - val_loss: 0.6257 - val_accuracy: 0.8745
Epoch 29/200
22014/22014 [==============================] - 0s 8us/step - loss: 0.3785 -
accuracy: 0.8952 - val_loss: 0.3698 - val_accuracy: 0.8947
Epoch 30/200
22014/22014 [==============================] - 0s 8us/step - loss: 0.3797 -
```

```
accuracy: 0.8935 - val_loss: 0.2673 - val_accuracy: 0.9084
Epoch 31/200
22014/22014 [==============================] - 0s 10us/step - loss: 0.3456 -
accuracy: 0.8989 - val_loss: 0.4913 - val_accuracy: 0.8880
Epoch 32/200
22014/22014 [==============================] - 0s 8us/step - loss: 0.3641 -
accuracy: 0.8959 - val_loss: 0.8985 - val_accuracy: 0.8825
Epoch 33/200
22014/22014 [==============================] - 0s 9us/step - loss: 0.3589 -
accuracy: 0.8962 - val_loss: 0.2861 - val_accuracy: 0.9061
Epoch 34/200
22014/22014 [==============================] - 0s 7us/step - loss: 0.3534 -
accuracy: 0.8962 - val_loss: 0.2926 - val_accuracy: 0.9055
Epoch 35/200
22014/22014 [==============================] - 0s 7us/step - loss: 0.3727 -
accuracy: 0.8958 - val_loss: 0.3313 - val_accuracy: 0.9008
Epoch 36/200
22014/22014 [==============================] - 0s 7us/step - loss: 0.3698 -
accuracy: 0.8956 - val_loss: 0.2755 - val_accuracy: 0.9007
Epoch 37/200
22014/22014 [==============================] - 0s 9us/step - loss: 0.3338 -
accuracy: 0.8996 - val_loss: 0.3124 - val_accuracy: 0.8973
Epoch 38/200
22014/22014 [==============================] - 0s 8us/step - loss: 0.3535 -
accuracy: 0.8943 - val_loss: 0.2497 - val_accuracy: 0.9072
Epoch 39/200
22014/22014 [==============================] - 0s 9us/step - loss: 0.3361 -
accuracy: 0.8952 - val_loss: 0.2820 - val_accuracy: 0.9021
Epoch 40/200
22014/22014 [==============================] - 0s 9us/step - loss: 0.3336 -
accuracy: 0.8972 - val_loss: 0.2574 - val_accuracy: 0.9013
Epoch 41/200
22014/22014 [==============================] - 0s 10us/step - loss: 0.3356 -
accuracy: 0.8947 - val_loss: 0.4659 - val_accuracy: 0.8920
Epoch 42/200
22014/22014 [==============================] - 0s 7us/step - loss: 0.3246 -
accuracy: 0.8965 - val_loss: 0.2592 - val_accuracy: 0.9043
Epoch 43/200
22014/22014 [==============================] - 0s 10us/step - loss: 0.3303 -
accuracy: 0.8957 - val_loss: 0.5505 - val_accuracy: 0.8756
Epoch 44/200
22014/22014 [==============================] - 0s 7us/step - loss: 0.3387 -
accuracy: 0.8949 - val_loss: 0.3118 - val_accuracy: 0.8979
Epoch 45/200
22014/22014 [==============================] - 0s 6us/step - loss: 0.3299 -
accuracy: 0.8960 - val_loss: 0.2506 - val_accuracy: 0.9067
Epoch 46/200
22014/22014 [==============================] - 0s 7us/step - loss: 0.3045 -
```

```
accuracy: 0.8980 - val_loss: 0.2530 - val_accuracy: 0.9029
Epoch 47/200
22014/22014 [==============================] - 0s 8us/step - loss: 0.3460 -
accuracy: 0.8974 - val_loss: 0.2758 - val_accuracy: 0.9047
Epoch 48/200
22014/22014 [==============================] - 0s 8us/step - loss: 0.3106 -
accuracy: 0.8968 - val_loss: 0.2594 - val_accuracy: 0.9004
Epoch 49/200
22014/22014 [==============================] - 0s 7us/step - loss: 0.3339 -
accuracy: 0.8955 - val_loss: 0.2723 - val_accuracy: 0.8929
Epoch 50/200
22014/22014 [==============================] - 0s 6us/step - loss: 0.3204 -
accuracy: 0.8980 - val_loss: 0.2723 - val_accuracy: 0.9039
Epoch 51/200
22014/22014 [==============================] - 0s 8us/step - loss: 0.3289 -
accuracy: 0.8982 - val_loss: 0.2717 - val_accuracy: 0.9072
Epoch 52/200
22014/22014 [==============================] - 0s 8us/step - loss: 0.3232 -
accuracy: 0.8987 - val_loss: 0.2795 - val_accuracy: 0.9076
Epoch 53/200
22014/22014 [==============================] - 0s 7us/step - loss: 0.3427 -
accuracy: 0.8986 - val_loss: 0.2555 - val_accuracy: 0.9032
Epoch 54/200
22014/22014 [==============================] - 0s 7us/step - loss: 0.3066 -
accuracy: 0.8980 - val_loss: 0.2852 - val_accuracy: 0.9033
Epoch 55/200
22014/22014 [==============================] - 0s 9us/step - loss: 0.3205 -
accuracy: 0.8972 - val_loss: 0.4088 - val_accuracy: 0.8989
Epoch 56/200
22014/22014 [==============================] - 0s 7us/step - loss: 0.3095 -
accuracy: 0.8977 - val_loss: 0.3364 - val_accuracy: 0.8975
Epoch 57/200
22014/22014 [==============================] - 0s 7us/step - loss: 0.3046 -
accuracy: 0.8972 - val_loss: 0.2809 - val_accuracy: 0.9045
Epoch 58/200
22014/22014 [==============================] - 0s 7us/step - loss: 0.3200 -
accuracy: 0.9001 - val_loss: 0.2776 - val_accuracy: 0.9064
Epoch 59/200
22014/22014 [==============================] - 0s 7us/step - loss: 0.3055 -
accuracy: 0.9007 - val_loss: 0.5609 - val_accuracy: 0.8765
Epoch 60/200
22014/22014 [==============================] - 0s 8us/step - loss: 0.3180 -
accuracy: 0.8953 - val_loss: 0.2713 - val_accuracy: 0.9011
Epoch 61/200
22014/22014 [==============================] - 0s 10us/step - loss: 0.3037 -
accuracy: 0.8975 - val_loss: 0.2860 - val_accuracy: 0.9027
Epoch 62/200
22014/22014 [==============================] - 0s 8us/step - loss: 0.3045 -
```

```
accuracy: 0.9001 - val_loss: 0.5184 - val_accuracy: 0.8900
Epoch 63/200
22014/22014 [==============================] - 0s 8us/step - loss: 0.3202 -
accuracy: 0.8982 - val_loss: 0.6067 - val_accuracy: 0.8824
Epoch 64/200
22014/22014 [==============================] - 0s 7us/step - loss: 0.3112 -
accuracy: 0.8980 - val_loss: 0.3224 - val_accuracy: 0.8980
Epoch 65/200
22014/22014 [==============================] - 0s 7us/step - loss: 0.3220 -
accuracy: 0.8992 - val_loss: 0.2642 - val_accuracy: 0.9024
Epoch 66/200
22014/22014 [==============================] - 0s 7us/step - loss: 0.2779 -
accuracy: 0.9015 - val_loss: 0.2729 - val_accuracy: 0.8944
Epoch 67/200
22014/22014 [==============================] - 0s 8us/step - loss: 0.2981 -
accuracy: 0.8989 - val_loss: 0.4676 - val_accuracy: 0.8871
Epoch 68/200
22014/22014 [==============================] - 0s 7us/step - loss: 0.2959 -
accuracy: 0.9009 - val_loss: 0.3403 - val_accuracy: 0.9011
Epoch 69/200
22014/22014 [==============================] - 0s 6us/step - loss: 0.2963 -
accuracy: 0.8988 - val_loss: 0.3363 - val_accuracy: 0.8980
Epoch 70/200
22014/22014 [==============================] - 0s 7us/step - loss: 0.3107 -
accuracy: 0.9014 - val_loss: 0.2491 - val_accuracy: 0.9072
Epoch 71/200
22014/22014 [==============================] - 0s 7us/step - loss: 0.2875 -
accuracy: 0.9000 - val_loss: 0.2439 - val_accuracy: 0.9105
Epoch 72/200
22014/22014 [==============================] - 0s 7us/step - loss: 0.3251 -
accuracy: 0.8990 - val_loss: 0.2637 - val_accuracy: 0.9052
Epoch 73/200
22014/22014 [==============================] - 0s 10us/step - loss: 0.2686 -
accuracy: 0.9020 - val_loss: 0.2484 - val_accuracy: 0.9087
Epoch 74/200
22014/22014 [==============================] - 0s 9us/step - loss: 0.2952 -
accuracy: 0.8999 - val_loss: 0.2578 - val_accuracy: 0.9037
Epoch 75/200
22014/22014 [==============================] - 0s 9us/step - loss: 0.3049 -
accuracy: 0.9030 - val_loss: 0.2527 - val_accuracy: 0.9064
Epoch 76/200
22014/22014 [==============================] - 0s 9us/step - loss: 0.2846 -
accuracy: 0.8997 - val_loss: 0.2478 - val_accuracy: 0.9061
Epoch 77/200
22014/22014 [==============================] - 0s 9us/step - loss: 0.3234 -
accuracy: 0.9036 - val_loss: 0.3321 - val_accuracy: 0.9011
Epoch 78/200
22014/22014 [==============================] - 0s 7us/step - loss: 0.2876 -
```

```
accuracy: 0.9045 - val_loss: 0.3266 - val_accuracy: 0.9007
Epoch 79/200
22014/22014 [==============================] - 0s 7us/step - loss: 0.2938 -
accuracy: 0.9001 - val_loss: 0.2888 - val_accuracy: 0.9044
Epoch 80/200
22014/22014 [==============================] - 0s 6us/step - loss: 0.2962 -
accuracy: 0.9024 - val_loss: 0.2567 - val_accuracy: 0.9064
Epoch 81/200
22014/22014 [==============================] - 0s 7us/step - loss: 0.2795 -
accuracy: 0.9020 - val_loss: 0.3028 - val_accuracy: 0.9029
Epoch 82/200
22014/22014 [==============================] - 0s 10us/step - loss: 0.2849 -
accuracy: 0.9016 - val_loss: 0.4257 - val_accuracy: 0.8949
Epoch 83/200
22014/22014 [==============================] - 0s 7us/step - loss: 0.2845 -
accuracy: 0.9004 - val_loss: 0.2462 - val_accuracy: 0.9064
Epoch 84/200
22014/22014 [==============================] - 0s 8us/step - loss: 0.2866 -
accuracy: 0.9035 - val_loss: 0.2510 - val_accuracy: 0.9073
Epoch 85/200
22014/22014 [==============================] - 0s 7us/step - loss: 0.2834 -
accuracy: 0.9034 - val_loss: 0.5833 - val_accuracy: 0.8715
Epoch 86/200
22014/22014 [==============================] - 0s 8us/step - loss: 0.2909 -
accuracy: 0.9024 - val_loss: 0.2388 - val_accuracy: 0.9091
Epoch 87/200
22014/22014 [==============================] - 0s 8us/step - loss: 0.2957 -
accuracy: 0.9016 - val_loss: 0.2585 - val_accuracy: 0.9045
Epoch 88/200
22014/22014 [==============================] - 0s 7us/step - loss: 0.2753 -
accuracy: 0.9058 - val_loss: 0.2596 - val_accuracy: 0.9024
Epoch 89/200
22014/22014 [==============================] - 0s 7us/step - loss: 0.2770 -
accuracy: 0.9040 - val_loss: 0.2880 - val_accuracy: 0.8992
Epoch 90/200
22014/22014 [==============================] - 0s 7us/step - loss: 0.2803 -
accuracy: 0.9028 - val_loss: 0.2472 - val_accuracy: 0.9049
Epoch 91/200
22014/22014 [==============================] - 0s 8us/step - loss: 0.2873 -
accuracy: 0.9007 - val_loss: 0.2867 - val_accuracy: 0.8977
Epoch 92/200
22014/22014 [==============================] - 0s 9us/step - loss: 0.3039 -
accuracy: 0.9031 - val_loss: 0.2424 - val_accuracy: 0.9073
Epoch 93/200
22014/22014 [==============================] - 0s 7us/step - loss: 0.2829 -
accuracy: 0.9023 - val_loss: 0.2497 - val_accuracy: 0.9019
Epoch 94/200
22014/22014 [==============================] - 0s 7us/step - loss: 0.2757 -
```

```
accuracy: 0.9058 - val_loss: 0.2401 - val_accuracy: 0.9087
Epoch 95/200
22014/22014 [==============================] - 0s 8us/step - loss: 0.2997 -
accuracy: 0.9029 - val_loss: 0.2674 - val_accuracy: 0.8936
Epoch 96/200
22014/22014 [==============================] - 0s 8us/step - loss: 0.2725 -
accuracy: 0.9052 - val_loss: 0.6735 - val_accuracy: 0.8860
Epoch 97/200
22014/22014 [==============================] - 0s 7us/step - loss: 0.2863 -
accuracy: 0.9032 - val_loss: 0.2736 - val_accuracy: 0.9047
Epoch 98/200
22014/22014 [==============================] - 0s 6us/step - loss: 0.2721 -
accuracy: 0.9052 - val_loss: 0.3668 - val_accuracy: 0.8947
Epoch 99/200
22014/22014 [==============================] - 0s 8us/step - loss: 0.2868 -
accuracy: 0.9027 - val_loss: 0.2390 - val_accuracy: 0.9099
Epoch 100/200
22014/22014 [==============================] - 0s 8us/step - loss: 0.2859 -
accuracy: 0.9058 - val_loss: 0.2408 - val_accuracy: 0.9125
Epoch 101/200
22014/22014 [==============================] - 0s 10us/step - loss: 0.2800 -
accuracy: 0.9034 - val_loss: 0.2487 - val_accuracy: 0.9041
Epoch 102/200
22014/22014 [==============================] - 0s 9us/step - loss: 0.2794 -
accuracy: 0.9051 - val_loss: 0.2390 - val_accuracy: 0.9096
Epoch 103/200
22014/22014 [==============================] - 0s 7us/step - loss: 0.2744 -
accuracy: 0.9052 - val_loss: 0.2478 - val_accuracy: 0.9075
Epoch 104/200
22014/22014 [==============================] - 0s 7us/step - loss: 0.2821 -
accuracy: 0.9019 - val_loss: 0.4266 - val_accuracy: 0.8843
Epoch 105/200
22014/22014 [==============================] - 0s 9us/step - loss: 0.2977 -
accuracy: 0.9030 - val_loss: 0.2892 - val_accuracy: 0.9040
Epoch 106/200
22014/22014 [==============================] - 0s 10us/step - loss: 0.2897 -
accuracy: 0.9058 - val_loss: 0.2693 - val_accuracy: 0.9065
Epoch 107/200
22014/22014 [==============================] - 0s 10us/step - loss: 0.2803 -
accuracy: 0.9038 - val_loss: 0.2897 - val_accuracy: 0.8984
Epoch 108/200
22014/22014 [==============================] - 0s 8us/step - loss: 0.2736 -
accuracy: 0.9057 - val_loss: 0.2388 - val_accuracy: 0.9091
Epoch 109/200
22014/22014 [==============================] - 0s 7us/step - loss: 0.2862 -
accuracy: 0.9018 - val_loss: 0.3905 - val_accuracy: 0.8820
Epoch 110/200
22014/22014 [==============================] - 0s 7us/step - loss: 0.2560 -
```

```
accuracy: 0.9075 - val_loss: 0.4085 - val_accuracy: 0.8969
Epoch 111/200
22014/22014 [==============================] - 0s 7us/step - loss: 0.2720 -
accuracy: 0.9039 - val_loss: 0.3350 - val_accuracy: 0.9001
Epoch 112/200
22014/22014 [==============================] - 0s 9us/step - loss: 0.2773 -
accuracy: 0.9060 - val_loss: 0.2696 - val_accuracy: 0.9071
Epoch 113/200
22014/22014 [==============================] - 0s 10us/step - loss: 0.2816 -
accuracy: 0.9043 - val_loss: 0.2424 - val_accuracy: 0.9056
Epoch 114/200
22014/22014 [==============================] - 0s 8us/step - loss: 0.2575 -
accuracy: 0.9076 - val_loss: 0.2609 - val_accuracy: 0.9029
Epoch 115/200
22014/22014 [==============================] - 0s 9us/step - loss: 0.2762 -
accuracy: 0.9022 - val_loss: 0.2673 - val_accuracy: 0.9015
Epoch 116/200
22014/22014 [==============================] - 0s 7us/step - loss: 0.2797 -
accuracy: 0.9029 - val_loss: 0.2374 - val_accuracy: 0.9109
Epoch 117/200
22014/22014 [==============================] - 0s 8us/step - loss: 0.2550 -
accuracy: 0.9059 - val_loss: 0.2407 - val_accuracy: 0.9048
Epoch 118/200
22014/22014 [==============================] - 0s 10us/step - loss: 0.2710 -
accuracy: 0.9067 - val_loss: 0.2496 - val_accuracy: 0.9080
Epoch 119/200
22014/22014 [==============================] - 0s 9us/step - loss: 0.2736 -
accuracy: 0.9039 - val_loss: 0.3230 - val_accuracy: 0.8983
Epoch 120/200
22014/22014 [==============================] - 0s 8us/step - loss: 0.2667 -
accuracy: 0.9055 - val_loss: 0.2415 - val_accuracy: 0.9105
Epoch 121/200
22014/22014 [==============================] - 0s 8us/step - loss: 0.2703 -
accuracy: 0.9058 - val_loss: 0.2389 - val_accuracy: 0.9104
Epoch 122/200
22014/22014 [==============================] - 0s 9us/step - loss: 0.2600 -
accuracy: 0.9048 - val_loss: 0.2376 - val_accuracy: 0.9077
Epoch 123/200
22014/22014 [==============================] - 0s 8us/step - loss: 0.2668 -
accuracy: 0.9047 - val_loss: 0.2402 - val_accuracy: 0.9072
Epoch 124/200
22014/22014 [==============================] - 0s 8us/step - loss: 0.2608 -
accuracy: 0.9055 - val_loss: 0.2526 - val_accuracy: 0.9051
Epoch 125/200
22014/22014 [==============================] - 0s 7us/step - loss: 0.2845 -
accuracy: 0.9050 - val_loss: 0.5259 - val_accuracy: 0.8917
Epoch 126/200
22014/22014 [==============================] - 0s 6us/step - loss: 0.2787 -
```

```
accuracy: 0.9079 - val_loss: 0.3295 - val_accuracy: 0.8967
Epoch 127/200
22014/22014 [==============================] - 0s 6us/step - loss: 0.2754 -
accuracy: 0.9061 - val_loss: 0.4342 - val_accuracy: 0.8911
Epoch 128/200
22014/22014 [==============================] - 0s 6us/step - loss: 0.2847 -
accuracy: 0.9042 - val_loss: 0.2526 - val_accuracy: 0.9092
Epoch 129/200
22014/22014 [==============================] - 0s 6us/step - loss: 0.2768 -
accuracy: 0.9070 - val_loss: 0.2527 - val_accuracy: 0.9071
Epoch 130/200
22014/22014 [==============================] - 0s 6us/step - loss: 0.2724 -
accuracy: 0.9062 - val_loss: 0.2461 - val_accuracy: 0.9095
Epoch 131/200
22014/22014 [==============================] - 0s 6us/step - loss: 0.3029 -
accuracy: 0.9070 - val_loss: 0.2489 - val_accuracy: 0.9109
Epoch 132/200
22014/22014 [==============================] - 0s 7us/step - loss: 0.2733 -
accuracy: 0.9071 - val_loss: 0.2478 - val_accuracy: 0.9081
Epoch 133/200
22014/22014 [==============================] - 0s 7us/step - loss: 0.2775 -
accuracy: 0.9072 - val_loss: 0.2441 - val_accuracy: 0.9095
Epoch 134/200
22014/22014 [==============================] - 0s 7us/step - loss: 0.2584 -
accuracy: 0.9073 - val_loss: 0.3219 - val_accuracy: 0.8988
Epoch 135/200
22014/22014 [==============================] - 0s 7us/step - loss: 0.2764 -
accuracy: 0.9070 - val_loss: 0.2480 - val_accuracy: 0.9068
Epoch 136/200
22014/22014 [==============================] - 0s 6us/step - loss: 0.2586 -
accuracy: 0.9067 - val_loss: 0.2828 - val_accuracy: 0.8995
Epoch 137/200
22014/22014 [==============================] - 0s 6us/step - loss: 0.2737 -
accuracy: 0.9059 - val_loss: 0.2389 - val_accuracy: 0.9113
Epoch 138/200
22014/22014 [==============================] - 0s 6us/step - loss: 0.2523 -
accuracy: 0.9083 - val_loss: 1.1833 - val_accuracy: 0.8871
Epoch 139/200
22014/22014 [==============================] - 0s 6us/step - loss: 0.3032 -
accuracy: 0.9068 - val_loss: 0.2438 - val_accuracy: 0.9060
Epoch 140/200
22014/22014 [==============================] - 0s 6us/step - loss: 0.2509 -
accuracy: 0.9068 - val_loss: 0.2753 - val_accuracy: 0.9031
Epoch 141/200
22014/22014 [==============================] - 0s 7us/step - loss: 0.2608 -
accuracy: 0.9083 - val_loss: 0.2778 - val_accuracy: 0.9051
Epoch 142/200
22014/22014 [==============================] - 0s 7us/step - loss: 0.2583 -
```

```
accuracy: 0.9071 - val_loss: 0.2424 - val_accuracy: 0.9069
Epoch 143/200
22014/22014 [==============================] - 0s 7us/step - loss: 0.2762 -
accuracy: 0.9085 - val_loss: 0.2450 - val_accuracy: 0.9067
Epoch 144/200
22014/22014 [==============================] - 0s 6us/step - loss: 0.2637 -
accuracy: 0.9062 - val_loss: 0.2374 - val_accuracy: 0.9123
Epoch 145/200
22014/22014 [==============================] - 0s 6us/step - loss: 0.2557 -
accuracy: 0.9079 - val_loss: 0.2373 - val_accuracy: 0.9108
Epoch 146/200
22014/22014 [==============================] - 0s 7us/step - loss: 0.2520 -
accuracy: 0.9068 - val_loss: 0.2782 - val_accuracy: 0.9039
Epoch 147/200
22014/22014 [==============================] - 0s 7us/step - loss: 0.2599 -
accuracy: 0.9047 - val_loss: 0.2299 - val_accuracy: 0.9128
Epoch 148/200
22014/22014 [==============================] - 0s 7us/step - loss: 0.2535 -
accuracy: 0.9049 - val_loss: 0.2312 - val_accuracy: 0.9099
Epoch 149/200
22014/22014 [==============================] - 0s 6us/step - loss: 0.2580 -
accuracy: 0.9068 - val_loss: 0.2339 - val_accuracy: 0.9117
Epoch 150/200
22014/22014 [==============================] - 0s 6us/step - loss: 0.2496 -
accuracy: 0.9071 - val_loss: 0.2386 - val_accuracy: 0.9068
Epoch 151/200
22014/22014 [==============================] - 0s 6us/step - loss: 0.2629 -
accuracy: 0.9061 - val_loss: 0.2429 - val_accuracy: 0.9093
Epoch 152/200
22014/22014 [==============================] - 0s 7us/step - loss: 0.2718 -
accuracy: 0.9057 - val_loss: 0.2327 - val_accuracy: 0.9117
Epoch 153/200
22014/22014 [==============================] - 0s 7us/step - loss: 0.2435 -
accuracy: 0.9098 - val_loss: 0.2398 - val_accuracy: 0.9052
Epoch 154/200
22014/22014 [==============================] - 0s 7us/step - loss: 0.2647 -
accuracy: 0.9070 - val_loss: 0.2858 - val_accuracy: 0.9017
Epoch 155/200
22014/22014 [==============================] - 0s 6us/step - loss: 0.2443 -
accuracy: 0.9096 - val_loss: 0.2433 - val_accuracy: 0.9089
Epoch 156/200
22014/22014 [==============================] - 0s 6us/step - loss: 0.2430 -
accuracy: 0.9076 - val_loss: 0.2578 - val_accuracy: 0.9031
Epoch 157/200
22014/22014 [==============================] - 0s 6us/step - loss: 0.2421 -
accuracy: 0.9068 - val_loss: 0.2445 - val_accuracy: 0.8992
Epoch 158/200
22014/22014 [==============================] - 0s 6us/step - loss: 0.2429 -
```

```
accuracy: 0.9061 - val_loss: 0.2331 - val_accuracy: 0.9064
Epoch 159/200
22014/22014 [==============================] - 0s 6us/step - loss: 0.2552 -
accuracy: 0.9082 - val_loss: 0.2303 - val_accuracy: 0.9117
Epoch 160/200
22014/22014 [==============================] - 0s 6us/step - loss: 0.2462 -
accuracy: 0.9060 - val_loss: 0.2467 - val_accuracy: 0.9043
Epoch 161/200
22014/22014 [==============================] - 0s 7us/step - loss: 0.2482 -
accuracy: 0.9065 - val_loss: 0.2579 - val_accuracy: 0.9025
Epoch 162/200
22014/22014 [==============================] - 0s 6us/step - loss: 0.2352 -
accuracy: 0.9077 - val_loss: 0.2759 - val_accuracy: 0.9047
Epoch 163/200
22014/22014 [==============================] - 0s 6us/step - loss: 0.2454 -
accuracy: 0.9075 - val_loss: 0.2241 - val_accuracy: 0.9133
Epoch 164/200
22014/22014 [==============================] - 0s 7us/step - loss: 0.2346 -
accuracy: 0.9042 - val_loss: 0.2395 - val_accuracy: 0.9001
Epoch 165/200
22014/22014 [==============================] - 0s 6us/step - loss: 0.2366 -
accuracy: 0.9085 - val_loss: 0.2388 - val_accuracy: 0.9059
Epoch 166/200
22014/22014 [==============================] - 0s 6us/step - loss: 0.2381 -
accuracy: 0.9052 - val_loss: 0.3031 - val_accuracy: 0.8915
Epoch 167/200
22014/22014 [==============================] - 0s 7us/step - loss: 0.2271 -
accuracy: 0.9080 - val_loss: 0.2539 - val_accuracy: 0.8991
Epoch 168/200
22014/22014 [==============================] - 0s 7us/step - loss: 0.2348 -
accuracy: 0.9077 - val_loss: 0.2279 - val_accuracy: 0.9069
Epoch 169/200
22014/22014 [==============================] - 0s 7us/step - loss: 0.2285 -
accuracy: 0.9076 - val_loss: 0.2417 - val_accuracy: 0.9041
Epoch 170/200
22014/22014 [==============================] - 0s 7us/step - loss: 0.2349 -
accuracy: 0.9066 - val_loss: 0.2654 - val_accuracy: 0.9029
Epoch 171/200
22014/22014 [==============================] - 0s 6us/step - loss: 0.2313 -
accuracy: 0.9082 - val_loss: 0.2420 - val_accuracy: 0.9005
Epoch 172/200
22014/22014 [==============================] - 0s 6us/step - loss: 0.2325 -
accuracy: 0.9053 - val_loss: 0.2275 - val_accuracy: 0.9097
Epoch 173/200
22014/22014 [==============================] - 0s 6us/step - loss: 0.2262 -
accuracy: 0.9063 - val_loss: 0.2270 - val_accuracy: 0.9077
Epoch 174/200
22014/22014 [==============================] - 0s 6us/step - loss: 0.2319 -
```

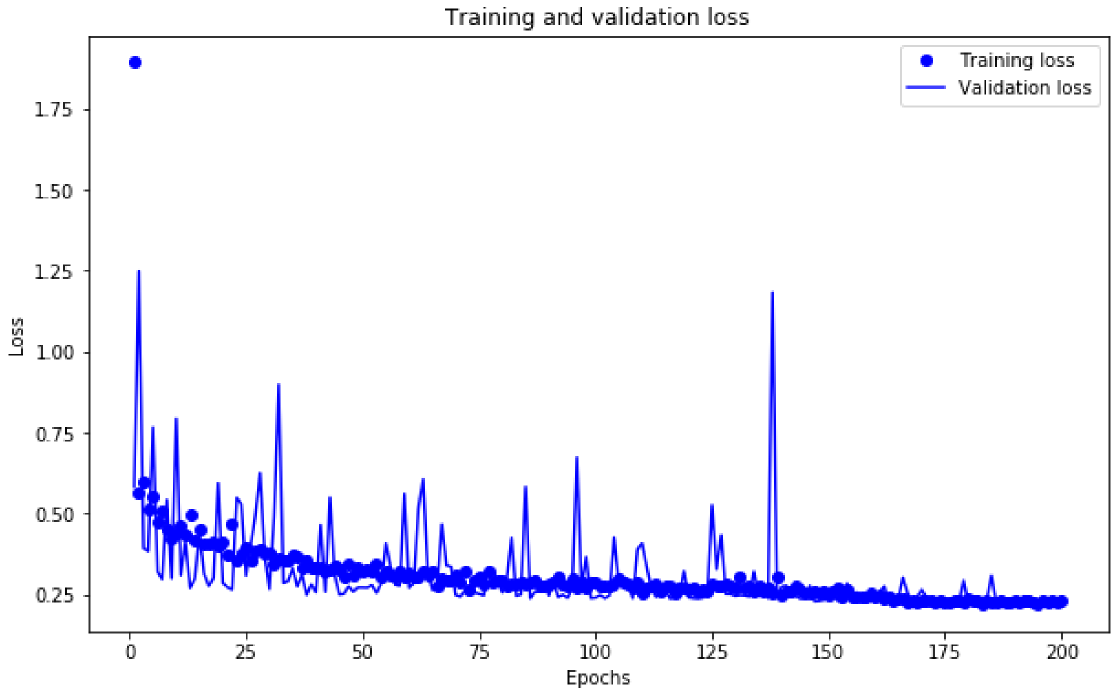
```
accuracy: 0.9067 - val_loss: 0.2255 - val_accuracy: 0.9099
Epoch 175/200
22014/22014 [==============================] - 0s 6us/step - loss: 0.2290 -
accuracy: 0.9095 - val_loss: 0.2334 - val_accuracy: 0.9060
Epoch 176/200
22014/22014 [==============================] - 0s 6us/step - loss: 0.2267 -
accuracy: 0.9081 - val_loss: 0.2304 - val_accuracy: 0.9081
Epoch 177/200
22014/22014 [==============================] - 0s 6us/step - loss: 0.2298 -
accuracy: 0.9061 - val_loss: 0.2295 - val_accuracy: 0.9085
Epoch 178/200
22014/22014 [==============================] - 0s 6us/step - loss: 0.2328 -
accuracy: 0.9077 - val_loss: 0.2254 - val_accuracy: 0.9135
Epoch 179/200
22014/22014 [==============================] - 0s 6us/step - loss: 0.2283 -
accuracy: 0.9093 - val_loss: 0.2940 - val_accuracy: 0.8959
Epoch 180/200
22014/22014 [==============================] - 0s 6us/step - loss: 0.2369 -
accuracy: 0.9085 - val_loss: 0.2264 - val_accuracy: 0.9101
Epoch 181/200
22014/22014 [==============================] - 0s 7us/step - loss: 0.2319 -
accuracy: 0.9097 - val_loss: 0.2268 - val_accuracy: 0.9117
Epoch 182/200
22014/22014 [==============================] - 0s 6us/step - loss: 0.2335 -
accuracy: 0.9075 - val_loss: 0.2401 - val_accuracy: 0.9040
Epoch 183/200
22014/22014 [==============================] - 0s 6us/step - loss: 0.2233 -
accuracy: 0.9106 - val_loss: 0.2331 - val_accuracy: 0.9053
Epoch 184/200
22014/22014 [==============================] - 0s 6us/step - loss: 0.2322 -
accuracy: 0.9085 - val_loss: 0.2335 - val_accuracy: 0.9068
Epoch 185/200
22014/22014 [==============================] - 0s 6us/step - loss: 0.2291 -
accuracy: 0.9098 - val_loss: 0.3091 - val_accuracy: 0.8933
Epoch 186/200
22014/22014 [==============================] - 0s 6us/step - loss: 0.2281 -
accuracy: 0.9078 - val_loss: 0.2277 - val_accuracy: 0.9131
Epoch 187/200
22014/22014 [==============================] - 0s 7us/step - loss: 0.2309 -
accuracy: 0.9075 - val_loss: 0.2251 - val_accuracy: 0.9124
Epoch 188/200
22014/22014 [==============================] - 0s 6us/step - loss: 0.2267 -
accuracy: 0.9099 - val_loss: 0.2296 - val_accuracy: 0.9088
Epoch 189/200
22014/22014 [==============================] - 0s 6us/step - loss: 0.2270 -
accuracy: 0.9083 - val_loss: 0.2282 - val_accuracy: 0.9121
Epoch 190/200
22014/22014 [==============================] - 0s 6us/step - loss: 0.2298 -
```

```
accuracy: 0.9074 - val_loss: 0.2275 - val_accuracy: 0.9111
Epoch 191/200
22014/22014 [==============================] - 0s 6us/step - loss: 0.2256 -
accuracy: 0.9097 - val_loss: 0.2237 - val_accuracy: 0.9132
Epoch 192/200
22014/22014 [==============================] - 0s 6us/step - loss: 0.2296 -
accuracy: 0.9086 - val_loss: 0.2454 - val_accuracy: 0.9033
Epoch 193/200
22014/22014 [==============================] - 0s 6us/step - loss: 0.2329 -
accuracy: 0.9085 - val_loss: 0.2256 - val_accuracy: 0.9127
Epoch 194/200
22014/22014 [==============================] - 0s 6us/step - loss: 0.2272 -
accuracy: 0.9078 - val_loss: 0.2285 - val_accuracy: 0.9085
Epoch 195/200
22014/22014 [==============================] - 0s 6us/step - loss: 0.2210 -
accuracy: 0.9093 - val_loss: 0.2265 - val_accuracy: 0.9104
Epoch 196/200
22014/22014 [==============================] - 0s 6us/step - loss: 0.2318 -
accuracy: 0.9109 - val_loss: 0.2276 - val_accuracy: 0.9113
Epoch 197/200
22014/22014 [==============================] - 0s 6us/step - loss: 0.2249 -
accuracy: 0.9102 - val_loss: 0.2275 - val_accuracy: 0.9112
Epoch 198/200
22014/22014 [==============================] - 0s 6us/step - loss: 0.2309 -
accuracy: 0.9082 - val_loss: 0.2325 - val_accuracy: 0.9100
Epoch 199/200
22014/22014 [==============================] - 0s 6us/step - loss: 0.2237 -
accuracy: 0.9099 - val_loss: 0.2280 - val_accuracy: 0.9119
Epoch 200/200
22014/22014 [==============================] - 0s 6us/step - loss: 0.2298 -
accuracy: 0.9081 - val_loss: 0.2292 - val_accuracy: 0.9093
```

Now, we can plot the results of loss values from the training and validation set.

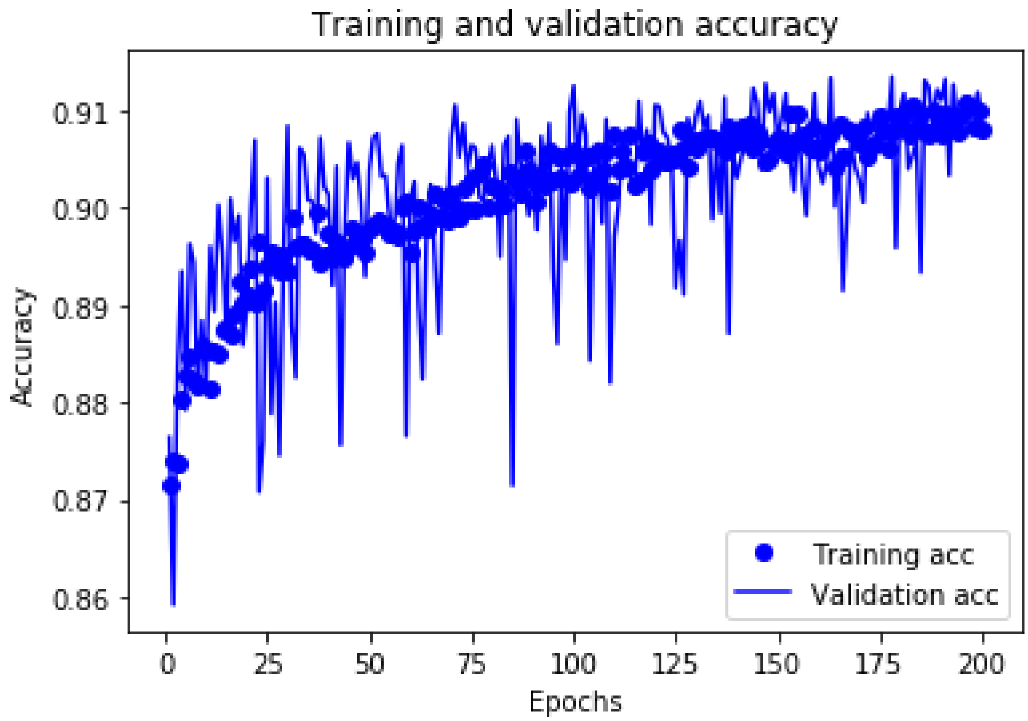
[19]:
```python
# plot the results of loss values from the training set and validtion set
history_dict = history.history
loss_values = history_dict['loss']
val_loss_values = history_dict['val_loss']

epochs = range(1, len(history_dict['accuracy']) + 1)

plt.figure(figsize=(10,6))
plt.plot(epochs, loss_values, 'bo', label='Training loss')
plt.plot(epochs, val_loss_values, 'b', label='Validation loss')
plt.title('Training and validation loss')
plt.xlabel('Epochs')
plt.ylabel('Loss')
plt.legend()
```

```
plt.show()
```



Now, we can plot the results of accuracy from the training and validation set.

[20]:
```
# plot the results of accuracy from the training set and validtion set
acc = history_dict['accuracy']
val_acc = history_dict['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('Accuracy')
plt.legend()
plt.show()
```



This simple works well in our training set and validation set(with accuracy about 0.9).

However, this model didn't perform well on our test set. When I used all the training data to train this model, the public score of this model is about 0.62. Why this happened? Well, I think the reason is about our imbalanced data, there are too many 0s in 'Target'. As a result, our model prefer to give us many 0s, which leads to our fail on testing dataset.

Now, in order to deal with the issue of imbalanced dataset, I use a weighted model.

To begin with, let's calculate class_weights.

[21]:
```
# Calculate class weight
NotOver50k, Over50k = np.bincount(train_data.Target)
total_count = len(train_data.Target)

weight_no_over50k = (1/NotOver50k)*(total_count)/2.0
weight_over50k = (1/Over50k)*(total_count)/2.0

class_weights = {0:weight_no_over50k, 1:weight_over50k}
```

Now, we can use class_weights as an argument when we construct our second NN model.

DL model draft:

3 hidden layers in this model.

The first hidden layer: 16 units with 'relu' activation function.

The second hidden layer: 16 units with ‘relu’ activation function.

The third hidden layer: 1 unit with ‘sigmoid’ activation function.

I choose the ‘rmsprop’ optimizer, ‘binary_crossentropy’ loss function, and the ‘accuracy’ metrics.

parameter initialization:

I use 200 epochs to train my model. The batch_size of my model is 512.

parameter tuning:

Because I think this model did well in my training dataset, I didn’t tune it’s parameters.

Now, let’s construct our second NN model.

[22]:
```
# Use weighted model!
model = models.Sequential()
model.add(layers.Dense(16, activation='relu'))
model.add(layers.Dense(16, activation='relu'))
model.add(layers.Dense(1, activation='sigmoid'))

model.compile(optimizer='rmsprop',
              loss='binary_crossentropy',
              metrics=['accuracy'])

history = model.fit(partial_X_train,
                    partial_y_train,
                    epochs=200,
                    batch_size=512,
                    validation_data=(X_valid, y_valid),
                    class_weight=class_weights)
```

```
Train on 22014 samples, validate on 7500 samples
Epoch 1/200
22014/22014 [==============================] - 0s 19us/step - loss: 36.1078 -
accuracy: 0.6971 - val_loss: 0.7639 - val_accuracy: 0.7001
Epoch 2/200
22014/22014 [==============================] - 0s 12us/step - loss: 0.7924 -
accuracy: 0.7053 - val_loss: 1.1875 - val_accuracy: 0.7517
Epoch 3/200
22014/22014 [==============================] - 0s 11us/step - loss: 0.7911 -
accuracy: 0.7337 - val_loss: 1.5703 - val_accuracy: 0.7423
Epoch 4/200
22014/22014 [==============================] - 0s 9us/step - loss: 0.8578 -
accuracy: 0.7393 - val_loss: 0.8123 - val_accuracy: 0.7447
Epoch 5/200
22014/22014 [==============================] - 0s 15us/step - loss: 0.6984 -
accuracy: 0.7435 - val_loss: 1.2160 - val_accuracy: 0.7577
Epoch 6/200
22014/22014 [==============================] - 0s 9us/step - loss: 0.6525 -
```

```
accuracy: 0.7446 - val_loss: 1.0948 - val_accuracy: 0.7339
Epoch 7/200
22014/22014 [==============================] - 0s 9us/step - loss: 0.7435 -
accuracy: 0.7474 - val_loss: 0.3850 - val_accuracy: 0.8023
Epoch 8/200
22014/22014 [==============================] - 0s 9us/step - loss: 0.7130 -
accuracy: 0.7453 - val_loss: 0.7885 - val_accuracy: 0.7639
Epoch 9/200
22014/22014 [==============================] - 0s 9us/step - loss: 0.6277 -
accuracy: 0.7557 - val_loss: 0.8590 - val_accuracy: 0.7180
Epoch 10/200
22014/22014 [==============================] - 0s 11us/step - loss: 0.6450 -
accuracy: 0.7523 - val_loss: 1.1902 - val_accuracy: 0.7676
Epoch 11/200
22014/22014 [==============================] - 0s 8us/step - loss: 0.6350 -
accuracy: 0.7629 - val_loss: 1.0369 - val_accuracy: 0.7180
Epoch 12/200
22014/22014 [==============================] - 0s 11us/step - loss: 0.5879 -
accuracy: 0.7659 - val_loss: 1.3399 - val_accuracy: 0.6761
Epoch 13/200
22014/22014 [==============================] - 0s 8us/step - loss: 0.6461 -
accuracy: 0.7650 - val_loss: 0.6391 - val_accuracy: 0.7660
Epoch 14/200
22014/22014 [==============================] - 0s 9us/step - loss: 0.5861 -
accuracy: 0.7646 - val_loss: 0.3685 - val_accuracy: 0.8239
Epoch 15/200
22014/22014 [==============================] - 0s 6us/step - loss: 0.6452 -
accuracy: 0.7669 - val_loss: 0.5489 - val_accuracy: 0.7464
Epoch 16/200
22014/22014 [==============================] - 0s 8us/step - loss: 0.6269 -
accuracy: 0.7751 - val_loss: 0.3917 - val_accuracy: 0.8027
Epoch 17/200
22014/22014 [==============================] - 0s 7us/step - loss: 0.7118 -
accuracy: 0.7686 - val_loss: 0.5356 - val_accuracy: 0.7715
Epoch 18/200
22014/22014 [==============================] - 0s 8us/step - loss: 0.5345 -
accuracy: 0.7794 - val_loss: 0.5262 - val_accuracy: 0.7431
Epoch 19/200
22014/22014 [==============================] - 0s 8us/step - loss: 0.7159 -
accuracy: 0.7744 - val_loss: 1.3546 - val_accuracy: 0.7645
Epoch 20/200
22014/22014 [==============================] - 0s 8us/step - loss: 0.5761 -
accuracy: 0.7787 - val_loss: 0.7012 - val_accuracy: 0.7767
Epoch 21/200
22014/22014 [==============================] - 0s 7us/step - loss: 0.6401 -
accuracy: 0.7786 - val_loss: 1.0219 - val_accuracy: 0.7356
Epoch 22/200
22014/22014 [==============================] - 0s 7us/step - loss: 0.5628 -
```

```
accuracy: 0.7769 - val_loss: 0.7507 - val_accuracy: 0.7900
Epoch 23/200
22014/22014 [==============================] - 0s 7us/step - loss: 0.5583 -
accuracy: 0.7764 - val_loss: 0.7045 - val_accuracy: 0.7352
Epoch 24/200
22014/22014 [==============================] - 0s 7us/step - loss: 0.6000 -
accuracy: 0.7760 - val_loss: 0.4992 - val_accuracy: 0.7589
Epoch 25/200
22014/22014 [==============================] - 0s 8us/step - loss: 0.5321 -
accuracy: 0.7807 - val_loss: 0.6666 - val_accuracy: 0.7581
Epoch 26/200
22014/22014 [==============================] - 0s 7us/step - loss: 0.5386 -
accuracy: 0.7830 - val_loss: 1.0790 - val_accuracy: 0.7152
Epoch 27/200
22014/22014 [==============================] - 0s 8us/step - loss: 0.5742 -
accuracy: 0.7772 - val_loss: 1.2616 - val_accuracy: 0.7416
Epoch 28/200
22014/22014 [==============================] - 0s 8us/step - loss: 0.6174 -
accuracy: 0.7784 - val_loss: 0.5936 - val_accuracy: 0.7565
Epoch 29/200
22014/22014 [==============================] - 0s 9us/step - loss: 0.4868 -
accuracy: 0.7852 - val_loss: 0.3948 - val_accuracy: 0.8027
Epoch 30/200
22014/22014 [==============================] - 0s 10us/step - loss: 0.6053 -
accuracy: 0.7795 - val_loss: 1.2338 - val_accuracy: 0.7763
Epoch 31/200
22014/22014 [==============================] - 0s 7us/step - loss: 0.5252 -
accuracy: 0.7831 - val_loss: 0.9792 - val_accuracy: 0.7901
Epoch 32/200
22014/22014 [==============================] - 0s 7us/step - loss: 0.6533 -
accuracy: 0.7776 - val_loss: 0.4808 - val_accuracy: 0.7732
Epoch 33/200
22014/22014 [==============================] - 0s 8us/step - loss: 0.4827 -
accuracy: 0.7817 - val_loss: 0.4736 - val_accuracy: 0.8037
Epoch 34/200
22014/22014 [==============================] - 0s 7us/step - loss: 0.5740 -
accuracy: 0.7768 - val_loss: 0.7182 - val_accuracy: 0.7964
Epoch 35/200
22014/22014 [==============================] - 0s 8us/step - loss: 0.5182 -
accuracy: 0.7791 - val_loss: 0.4619 - val_accuracy: 0.8137
Epoch 36/200
22014/22014 [==============================] - 0s 7us/step - loss: 0.5400 -
accuracy: 0.7826 - val_loss: 0.4653 - val_accuracy: 0.7752
Epoch 37/200
22014/22014 [==============================] - 0s 8us/step - loss: 0.5185 -
accuracy: 0.7851 - val_loss: 0.6590 - val_accuracy: 0.7561
Epoch 38/200
22014/22014 [==============================] - 0s 7us/step - loss: 0.5253 -
```

```
accuracy: 0.7841 - val_loss: 0.3891 - val_accuracy: 0.8305
Epoch 39/200
22014/22014 [==============================] - 0s 8us/step - loss: 0.5710 -
accuracy: 0.7838 - val_loss: 0.6012 - val_accuracy: 0.7597
Epoch 40/200
22014/22014 [==============================] - 0s 7us/step - loss: 0.5971 -
accuracy: 0.7825 - val_loss: 0.4319 - val_accuracy: 0.8105
Epoch 41/200
22014/22014 [==============================] - 0s 7us/step - loss: 0.5614 -
accuracy: 0.7865 - val_loss: 0.4034 - val_accuracy: 0.7923
Epoch 42/200
22014/22014 [==============================] - 0s 7us/step - loss: 0.5584 -
accuracy: 0.7834 - val_loss: 1.2954 - val_accuracy: 0.7129
Epoch 43/200
22014/22014 [==============================] - 0s 7us/step - loss: 0.5169 -
accuracy: 0.7819 - val_loss: 0.4781 - val_accuracy: 0.7685
Epoch 44/200
22014/22014 [==============================] - 0s 7us/step - loss: 0.5311 -
accuracy: 0.7835 - val_loss: 0.4114 - val_accuracy: 0.7901
Epoch 45/200
22014/22014 [==============================] - 0s 7us/step - loss: 0.5404 -
accuracy: 0.7852 - val_loss: 0.4643 - val_accuracy: 0.7653
Epoch 46/200
22014/22014 [==============================] - 0s 7us/step - loss: 0.5343 -
accuracy: 0.7844 - val_loss: 0.4548 - val_accuracy: 0.7996
Epoch 47/200
22014/22014 [==============================] - 0s 6us/step - loss: 0.5660 -
accuracy: 0.7831 - val_loss: 0.7202 - val_accuracy: 0.7607
Epoch 48/200
22014/22014 [==============================] - 0s 6us/step - loss: 0.5478 -
accuracy: 0.7839 - val_loss: 0.6085 - val_accuracy: 0.7705
Epoch 49/200
22014/22014 [==============================] - 0s 6us/step - loss: 0.5251 -
accuracy: 0.7884 - val_loss: 0.5945 - val_accuracy: 0.7929
Epoch 50/200
22014/22014 [==============================] - 0s 6us/step - loss: 0.5195 -
accuracy: 0.7885 - val_loss: 0.8170 - val_accuracy: 0.7752
Epoch 51/200
22014/22014 [==============================] - 0s 6us/step - loss: 0.5044 -
accuracy: 0.7869 - val_loss: 0.4523 - val_accuracy: 0.8345
Epoch 52/200
22014/22014 [==============================] - 0s 6us/step - loss: 0.5538 -
accuracy: 0.7891 - val_loss: 1.1294 - val_accuracy: 0.7251
Epoch 53/200
22014/22014 [==============================] - 0s 6us/step - loss: 0.5661 -
accuracy: 0.7890 - val_loss: 0.4668 - val_accuracy: 0.7683
Epoch 54/200
22014/22014 [==============================] - 0s 6us/step - loss: 0.7619 -
```

```
accuracy: 0.7846 - val_loss: 0.8684 - val_accuracy: 0.7732
Epoch 55/200
22014/22014 [==============================] - 0s 6us/step - loss: 0.5338 -
accuracy: 0.7908 - val_loss: 0.4339 - val_accuracy: 0.7816
Epoch 56/200
22014/22014 [==============================] - 0s 6us/step - loss: 0.6918 -
accuracy: 0.7834 - val_loss: 0.4406 - val_accuracy: 0.7947
Epoch 57/200
22014/22014 [==============================] - 0s 9us/step - loss: 0.5060 -
accuracy: 0.7917 - val_loss: 0.4267 - val_accuracy: 0.8015
Epoch 58/200
22014/22014 [==============================] - 0s 7us/step - loss: 0.5778 -
accuracy: 0.7863 - val_loss: 0.7752 - val_accuracy: 0.7635
Epoch 59/200
22014/22014 [==============================] - 0s 6us/step - loss: 0.6016 -
accuracy: 0.7850 - val_loss: 0.3880 - val_accuracy: 0.8033
Epoch 60/200
22014/22014 [==============================] - 0s 6us/step - loss: 0.6227 -
accuracy: 0.7837 - val_loss: 1.2242 - val_accuracy: 0.7949
Epoch 61/200
22014/22014 [==============================] - 0s 7us/step - loss: 0.5602 -
accuracy: 0.7919 - val_loss: 0.4420 - val_accuracy: 0.7772
Epoch 62/200
22014/22014 [==============================] - 0s 7us/step - loss: 0.6086 -
accuracy: 0.7855 - val_loss: 0.7287 - val_accuracy: 0.7616
Epoch 63/200
22014/22014 [==============================] - 0s 6us/step - loss: 0.6883 -
accuracy: 0.7834 - val_loss: 0.9766 - val_accuracy: 0.7641
Epoch 64/200
22014/22014 [==============================] - 0s 6us/step - loss: 0.5687 -
accuracy: 0.7883 - val_loss: 0.4355 - val_accuracy: 0.7889
Epoch 65/200
22014/22014 [==============================] - 0s 6us/step - loss: 0.6346 -
accuracy: 0.7901 - val_loss: 0.5009 - val_accuracy: 0.7839
Epoch 66/200
22014/22014 [==============================] - 0s 7us/step - loss: 0.5939 -
accuracy: 0.7879 - val_loss: 0.4922 - val_accuracy: 0.7812
Epoch 67/200
22014/22014 [==============================] - 0s 7us/step - loss: 0.5730 -
accuracy: 0.7876 - val_loss: 0.4239 - val_accuracy: 0.7875
Epoch 68/200
22014/22014 [==============================] - 0s 7us/step - loss: 0.5396 -
accuracy: 0.7912 - val_loss: 0.9779 - val_accuracy: 0.7419
Epoch 69/200
22014/22014 [==============================] - 0s 7us/step - loss: 0.5617 -
accuracy: 0.7889 - val_loss: 0.4442 - val_accuracy: 0.7867
Epoch 70/200
22014/22014 [==============================] - 0s 7us/step - loss: 0.6240 -
```

```
accuracy: 0.7910 - val_loss: 0.4306 - val_accuracy: 0.7797
Epoch 71/200
22014/22014 [==============================] - 0s 7us/step - loss: 0.5322 -
accuracy: 0.7886 - val_loss: 0.4252 - val_accuracy: 0.7904
Epoch 72/200
22014/22014 [==============================] - 0s 7us/step - loss: 0.5557 -
accuracy: 0.7910 - val_loss: 0.4362 - val_accuracy: 0.7821
Epoch 73/200
22014/22014 [==============================] - 0s 7us/step - loss: 0.5601 -
accuracy: 0.7914 - val_loss: 0.4282 - val_accuracy: 0.7920
Epoch 74/200
22014/22014 [==============================] - 0s 6us/step - loss: 0.5633 -
accuracy: 0.7902 - val_loss: 0.3796 - val_accuracy: 0.8191
Epoch 75/200
22014/22014 [==============================] - 0s 7us/step - loss: 0.5120 -
accuracy: 0.7930 - val_loss: 0.3823 - val_accuracy: 0.8044
Epoch 76/200
22014/22014 [==============================] - 0s 6us/step - loss: 0.5292 -
accuracy: 0.7900 - val_loss: 0.4660 - val_accuracy: 0.8101
Epoch 77/200
22014/22014 [==============================] - 0s 6us/step - loss: 0.5567 -
accuracy: 0.7911 - val_loss: 0.5078 - val_accuracy: 0.7660
Epoch 78/200
22014/22014 [==============================] - 0s 6us/step - loss: 0.6405 -
accuracy: 0.7926 - val_loss: 0.4600 - val_accuracy: 0.7905
Epoch 79/200
22014/22014 [==============================] - 0s 7us/step - loss: 0.6053 -
accuracy: 0.7895 - val_loss: 0.5916 - val_accuracy: 0.7839
Epoch 80/200
22014/22014 [==============================] - 0s 6us/step - loss: 0.5027 -
accuracy: 0.7957 - val_loss: 1.4119 - val_accuracy: 0.7553
Epoch 81/200
22014/22014 [==============================] - 0s 7us/step - loss: 0.6152 -
accuracy: 0.7860 - val_loss: 0.4689 - val_accuracy: 0.7913
Epoch 82/200
22014/22014 [==============================] - 0s 7us/step - loss: 0.4965 -
accuracy: 0.7935 - val_loss: 0.5763 - val_accuracy: 0.7827
Epoch 83/200
22014/22014 [==============================] - 0s 8us/step - loss: 0.5585 -
accuracy: 0.7927 - val_loss: 0.5130 - val_accuracy: 0.7768
Epoch 84/200
22014/22014 [==============================] - 0s 6us/step - loss: 0.5742 -
accuracy: 0.7920 - val_loss: 0.3601 - val_accuracy: 0.8204
Epoch 85/200
22014/22014 [==============================] - 0s 6us/step - loss: 0.4926 -
accuracy: 0.7938 - val_loss: 0.3949 - val_accuracy: 0.8041
Epoch 86/200
22014/22014 [==============================] - 0s 7us/step - loss: 0.5464 -
```

```
accuracy: 0.7911 - val_loss: 1.1301 - val_accuracy: 0.7477
Epoch 87/200
22014/22014 [==============================] - 0s 8us/step - loss: 0.5645 -
accuracy: 0.7917 - val_loss: 0.3736 - val_accuracy: 0.8127
Epoch 88/200
22014/22014 [==============================] - 0s 7us/step - loss: 0.5038 -
accuracy: 0.7976 - val_loss: 0.7597 - val_accuracy: 0.7331
Epoch 89/200
22014/22014 [==============================] - 0s 7us/step - loss: 0.5995 -
accuracy: 0.7898 - val_loss: 1.0028 - val_accuracy: 0.7883
Epoch 90/200
22014/22014 [==============================] - 0s 8us/step - loss: 0.5195 -
accuracy: 0.7916 - val_loss: 0.5263 - val_accuracy: 0.7684
Epoch 91/200
22014/22014 [==============================] - 0s 6us/step - loss: 0.5792 -
accuracy: 0.7919 - val_loss: 0.5973 - val_accuracy: 0.7885
Epoch 92/200
22014/22014 [==============================] - 0s 6us/step - loss: 0.5097 -
accuracy: 0.7910 - val_loss: 0.6712 - val_accuracy: 0.8333
Epoch 93/200
22014/22014 [==============================] - 0s 7us/step - loss: 0.4960 -
accuracy: 0.7929 - val_loss: 1.2770 - val_accuracy: 0.7735
Epoch 94/200
22014/22014 [==============================] - 0s 7us/step - loss: 0.5287 -
accuracy: 0.7931 - val_loss: 0.6996 - val_accuracy: 0.7847
Epoch 95/200
22014/22014 [==============================] - 0s 7us/step - loss: 0.4913 -
accuracy: 0.7921 - val_loss: 0.4209 - val_accuracy: 0.7956
Epoch 96/200
22014/22014 [==============================] - 0s 8us/step - loss: 0.4887 -
accuracy: 0.7918 - val_loss: 0.3997 - val_accuracy: 0.8032
Epoch 97/200
22014/22014 [==============================] - 0s 9us/step - loss: 0.4706 -
accuracy: 0.7950 - val_loss: 0.3541 - val_accuracy: 0.8295
Epoch 98/200
22014/22014 [==============================] - 0s 8us/step - loss: 0.4844 -
accuracy: 0.7931 - val_loss: 0.5378 - val_accuracy: 0.8288
Epoch 99/200
22014/22014 [==============================] - 0s 10us/step - loss: 0.4917 -
accuracy: 0.7968 - val_loss: 0.4776 - val_accuracy: 0.7696
Epoch 100/200
22014/22014 [==============================] - 0s 8us/step - loss: 0.4653 -
accuracy: 0.7924 - val_loss: 0.3373 - val_accuracy: 0.8365
Epoch 101/200
22014/22014 [==============================] - 0s 7us/step - loss: 0.4842 -
accuracy: 0.7919 - val_loss: 0.3767 - val_accuracy: 0.8197
Epoch 102/200
22014/22014 [==============================] - 0s 7us/step - loss: 0.4774 -
```

```
accuracy: 0.7951 - val_loss: 0.4378 - val_accuracy: 0.8001
Epoch 103/200
22014/22014 [==============================] - 0s 7us/step - loss: 0.4770 -
accuracy: 0.7914 - val_loss: 0.6837 - val_accuracy: 0.7971
Epoch 104/200
22014/22014 [==============================] - 0s 7us/step - loss: 0.5005 -
accuracy: 0.7902 - val_loss: 0.4873 - val_accuracy: 0.7999
Epoch 105/200
22014/22014 [==============================] - 0s 7us/step - loss: 0.4567 -
accuracy: 0.7944 - val_loss: 0.7552 - val_accuracy: 0.7353
Epoch 106/200
22014/22014 [==============================] - 0s 6us/step - loss: 0.4705 -
accuracy: 0.7943 - val_loss: 0.6798 - val_accuracy: 0.7373
Epoch 107/200
22014/22014 [==============================] - 0s 7us/step - loss: 0.4710 -
accuracy: 0.7917 - val_loss: 0.4792 - val_accuracy: 0.7796
Epoch 108/200
22014/22014 [==============================] - 0s 7us/step - loss: 0.4890 -
accuracy: 0.7945 - val_loss: 0.4879 - val_accuracy: 0.7552
Epoch 109/200
22014/22014 [==============================] - 0s 7us/step - loss: 0.4506 -
accuracy: 0.7936 - val_loss: 0.8967 - val_accuracy: 0.7539
Epoch 110/200
22014/22014 [==============================] - 0s 7us/step - loss: 0.4784 -
accuracy: 0.7938 - val_loss: 0.5656 - val_accuracy: 0.7757
Epoch 111/200
22014/22014 [==============================] - 0s 7us/step - loss: 0.5046 -
accuracy: 0.7919 - val_loss: 0.4178 - val_accuracy: 0.7905
Epoch 112/200
22014/22014 [==============================] - 0s 7us/step - loss: 0.4489 -
accuracy: 0.7967 - val_loss: 0.5015 - val_accuracy: 0.7984
Epoch 113/200
22014/22014 [==============================] - 0s 8us/step - loss: 0.4896 -
accuracy: 0.7929 - val_loss: 0.4366 - val_accuracy: 0.7843
Epoch 114/200
22014/22014 [==============================] - 0s 7us/step - loss: 0.4719 -
accuracy: 0.7940 - val_loss: 0.3563 - val_accuracy: 0.8257
Epoch 115/200
22014/22014 [==============================] - 0s 8us/step - loss: 0.5172 -
accuracy: 0.7922 - val_loss: 0.4780 - val_accuracy: 0.7584
Epoch 116/200
22014/22014 [==============================] - 0s 8us/step - loss: 0.4467 -
accuracy: 0.7969 - val_loss: 0.6823 - val_accuracy: 0.7568
Epoch 117/200
22014/22014 [==============================] - 0s 8us/step - loss: 0.4631 -
accuracy: 0.7949 - val_loss: 0.4697 - val_accuracy: 0.7583
Epoch 118/200
22014/22014 [==============================] - 0s 8us/step - loss: 0.4892 -
```

```
accuracy: 0.7928 - val_loss: 1.3529 - val_accuracy: 0.6903
Epoch 119/200
22014/22014 [==============================] - 0s 7us/step - loss: 0.4880 -
accuracy: 0.7905 - val_loss: 0.4449 - val_accuracy: 0.7797
Epoch 120/200
22014/22014 [==============================] - 0s 8us/step - loss: 0.4579 -
accuracy: 0.7938 - val_loss: 0.4231 - val_accuracy: 0.7952
Epoch 121/200
22014/22014 [==============================] - 0s 9us/step - loss: 0.4648 -
accuracy: 0.7917 - val_loss: 0.4076 - val_accuracy: 0.7948
Epoch 122/200
22014/22014 [==============================] - 0s 10us/step - loss: 0.4499 -
accuracy: 0.7940 - val_loss: 0.3234 - val_accuracy: 0.8461
Epoch 123/200
22014/22014 [==============================] - 0s 9us/step - loss: 0.4782 -
accuracy: 0.7935 - val_loss: 0.3324 - val_accuracy: 0.8444
Epoch 124/200
22014/22014 [==============================] - 0s 8us/step - loss: 0.4832 -
accuracy: 0.7932 - val_loss: 0.4398 - val_accuracy: 0.7991
Epoch 125/200
22014/22014 [==============================] - 0s 8us/step - loss: 0.4608 -
accuracy: 0.7976 - val_loss: 1.0953 - val_accuracy: 0.7615
Epoch 126/200
22014/22014 [==============================] - 0s 9us/step - loss: 0.4858 -
accuracy: 0.7956 - val_loss: 0.6463 - val_accuracy: 0.7297
Epoch 127/200
22014/22014 [==============================] - 0s 7us/step - loss: 0.4413 -
accuracy: 0.7931 - val_loss: 0.4086 - val_accuracy: 0.7931
Epoch 128/200
22014/22014 [==============================] - 0s 10us/step - loss: 0.4770 -
accuracy: 0.7927 - val_loss: 0.3648 - val_accuracy: 0.8137
Epoch 129/200
22014/22014 [==============================] - 0s 8us/step - loss: 0.4427 -
accuracy: 0.7975 - val_loss: 0.8332 - val_accuracy: 0.7380
Epoch 130/200
22014/22014 [==============================] - 0s 8us/step - loss: 0.4670 -
accuracy: 0.7913 - val_loss: 0.4977 - val_accuracy: 0.7863
Epoch 131/200
22014/22014 [==============================] - 0s 11us/step - loss: 0.5117 -
accuracy: 0.7934 - val_loss: 0.5098 - val_accuracy: 0.7815
Epoch 132/200
22014/22014 [==============================] - 0s 9us/step - loss: 0.4397 -
accuracy: 0.7961 - val_loss: 0.3966 - val_accuracy: 0.7985
Epoch 133/200
22014/22014 [==============================] - 0s 9us/step - loss: 0.4587 -
accuracy: 0.7978 - val_loss: 1.0605 - val_accuracy: 0.7007
Epoch 134/200
22014/22014 [==============================] - 0s 8us/step - loss: 0.4514 -
```

```
accuracy: 0.7916 - val_loss: 0.4129 - val_accuracy: 0.8033
Epoch 135/200
22014/22014 [==============================] - 0s 8us/step - loss: 0.4802 -
accuracy: 0.7945 - val_loss: 0.6439 - val_accuracy: 0.7635
Epoch 136/200
22014/22014 [==============================] - 0s 7us/step - loss: 0.4516 -
accuracy: 0.7953 - val_loss: 0.4203 - val_accuracy: 0.7947
Epoch 137/200
22014/22014 [==============================] - 0s 7us/step - loss: 0.4536 -
accuracy: 0.7960 - val_loss: 0.9329 - val_accuracy: 0.7605
Epoch 138/200
22014/22014 [==============================] - 0s 6us/step - loss: 0.5150 -
accuracy: 0.7935 - val_loss: 0.4412 - val_accuracy: 0.8060
Epoch 139/200
22014/22014 [==============================] - 0s 6us/step - loss: 0.4152 -
accuracy: 0.7982 - val_loss: 0.6886 - val_accuracy: 0.7859
Epoch 140/200
22014/22014 [==============================] - 0s 6us/step - loss: 0.4652 -
accuracy: 0.7960 - val_loss: 0.7678 - val_accuracy: 0.7720
Epoch 141/200
22014/22014 [==============================] - 0s 6us/step - loss: 0.4688 -
accuracy: 0.7953 - val_loss: 0.6395 - val_accuracy: 0.7841
Epoch 142/200
22014/22014 [==============================] - 0s 6us/step - loss: 0.4290 -
accuracy: 0.7941 - val_loss: 0.3783 - val_accuracy: 0.8080
Epoch 143/200
22014/22014 [==============================] - 0s 6us/step - loss: 0.4978 -
accuracy: 0.7960 - val_loss: 0.5400 - val_accuracy: 0.7903
Epoch 144/200
22014/22014 [==============================] - 0s 6us/step - loss: 0.4592 -
accuracy: 0.7948 - val_loss: 0.4490 - val_accuracy: 0.7967
Epoch 145/200
22014/22014 [==============================] - 0s 6us/step - loss: 0.4639 -
accuracy: 0.7949 - val_loss: 0.4036 - val_accuracy: 0.7929
Epoch 146/200
22014/22014 [==============================] - 0s 6us/step - loss: 0.4583 -
accuracy: 0.7954 - val_loss: 0.3759 - val_accuracy: 0.8069
Epoch 147/200
22014/22014 [==============================] - 0s 6us/step - loss: 0.4643 -
accuracy: 0.7928 - val_loss: 0.3264 - val_accuracy: 0.8379
Epoch 148/200
22014/22014 [==============================] - 0s 6us/step - loss: 0.4611 -
accuracy: 0.7992 - val_loss: 0.6712 - val_accuracy: 0.7457
Epoch 149/200
22014/22014 [==============================] - 0s 6us/step - loss: 0.4556 -
accuracy: 0.7954 - val_loss: 0.4093 - val_accuracy: 0.7977
Epoch 150/200
22014/22014 [==============================] - 0s 6us/step - loss: 0.4588 -
```

```
accuracy: 0.7977 - val_loss: 0.9183 - val_accuracy: 0.7699
Epoch 151/200
22014/22014 [==============================] - 0s 6us/step - loss: 0.4655 -
accuracy: 0.7945 - val_loss: 0.4087 - val_accuracy: 0.8069
Epoch 152/200
22014/22014 [==============================] - 0s 6us/step - loss: 0.4779 -
accuracy: 0.7963 - val_loss: 0.9944 - val_accuracy: 0.7633
Epoch 153/200
22014/22014 [==============================] - 0s 6us/step - loss: 0.4729 -
accuracy: 0.7944 - val_loss: 0.4372 - val_accuracy: 0.8149
Epoch 154/200
22014/22014 [==============================] - 0s 6us/step - loss: 0.4406 -
accuracy: 0.7978 - val_loss: 0.5721 - val_accuracy: 0.7655
Epoch 155/200
22014/22014 [==============================] - 0s 6us/step - loss: 0.4670 -
accuracy: 0.7970 - val_loss: 0.4188 - val_accuracy: 0.7907
Epoch 156/200
22014/22014 [==============================] - 0s 6us/step - loss: 0.4326 -
accuracy: 0.7989 - val_loss: 0.4767 - val_accuracy: 0.7769
Epoch 157/200
22014/22014 [==============================] - 0s 6us/step - loss: 0.4522 -
accuracy: 0.7945 - val_loss: 0.4754 - val_accuracy: 0.8024
Epoch 158/200
22014/22014 [==============================] - 0s 6us/step - loss: 0.4878 -
accuracy: 0.7994 - val_loss: 0.5389 - val_accuracy: 0.7931
Epoch 159/200
22014/22014 [==============================] - 0s 6us/step - loss: 0.4313 -
accuracy: 0.7978 - val_loss: 0.4616 - val_accuracy: 0.7795
Epoch 160/200
22014/22014 [==============================] - 0s 6us/step - loss: 0.4310 -
accuracy: 0.7980 - val_loss: 0.4394 - val_accuracy: 0.7771
Epoch 161/200
22014/22014 [==============================] - 0s 6us/step - loss: 0.4591 -
accuracy: 0.7970 - val_loss: 0.4078 - val_accuracy: 0.7927
Epoch 162/200
22014/22014 [==============================] - 0s 6us/step - loss: 0.4517 -
accuracy: 0.7979 - val_loss: 0.4746 - val_accuracy: 0.8145
Epoch 163/200
22014/22014 [==============================] - 0s 6us/step - loss: 0.4488 -
accuracy: 0.7992 - val_loss: 0.4685 - val_accuracy: 0.7721
Epoch 164/200
22014/22014 [==============================] - 0s 6us/step - loss: 0.4381 -
accuracy: 0.8002 - val_loss: 0.3809 - val_accuracy: 0.8092
Epoch 165/200
22014/22014 [==============================] - 0s 6us/step - loss: 0.5010 -
accuracy: 0.7958 - val_loss: 0.4365 - val_accuracy: 0.7863
Epoch 166/200
22014/22014 [==============================] - 0s 6us/step - loss: 0.4346 -
```

```
accuracy: 0.7974 - val_loss: 0.3579 - val_accuracy: 0.8227
Epoch 167/200
22014/22014 [==============================] - 0s 6us/step - loss: 0.4450 -
accuracy: 0.7981 - val_loss: 0.4357 - val_accuracy: 0.7949
Epoch 168/200
22014/22014 [==============================] - 0s 6us/step - loss: 0.4860 -
accuracy: 0.7974 - val_loss: 0.8447 - val_accuracy: 0.7773
Epoch 169/200
22014/22014 [==============================] - 0s 6us/step - loss: 0.4175 -
accuracy: 0.7998 - val_loss: 0.3878 - val_accuracy: 0.8137
Epoch 170/200
22014/22014 [==============================] - 0s 6us/step - loss: 0.4856 -
accuracy: 0.7933 - val_loss: 0.6967 - val_accuracy: 0.7761
Epoch 171/200
22014/22014 [==============================] - 0s 6us/step - loss: 0.4411 -
accuracy: 0.7974 - val_loss: 0.4652 - val_accuracy: 0.7872
Epoch 172/200
22014/22014 [==============================] - 0s 6us/step - loss: 0.4189 -
accuracy: 0.8002 - val_loss: 0.5426 - val_accuracy: 0.7753
Epoch 173/200
22014/22014 [==============================] - 0s 6us/step - loss: 0.4886 -
accuracy: 0.7988 - val_loss: 0.5805 - val_accuracy: 0.7841
Epoch 174/200
22014/22014 [==============================] - 0s 6us/step - loss: 0.4420 -
accuracy: 0.8002 - val_loss: 0.4373 - val_accuracy: 0.7907
Epoch 175/200
22014/22014 [==============================] - 0s 6us/step - loss: 0.4780 -
accuracy: 0.7976 - val_loss: 1.0136 - val_accuracy: 0.7612
Epoch 176/200
22014/22014 [==============================] - 0s 6us/step - loss: 0.4633 -
accuracy: 0.7968 - val_loss: 0.4633 - val_accuracy: 0.8047
Epoch 177/200
22014/22014 [==============================] - 0s 6us/step - loss: 0.4230 -
accuracy: 0.8015 - val_loss: 0.4556 - val_accuracy: 0.7687
Epoch 178/200
22014/22014 [==============================] - 0s 6us/step - loss: 0.4372 -
accuracy: 0.7974 - val_loss: 0.4813 - val_accuracy: 0.7660
Epoch 179/200
22014/22014 [==============================] - 0s 6us/step - loss: 0.4623 -
accuracy: 0.7989 - val_loss: 1.0870 - val_accuracy: 0.7467
Epoch 180/200
22014/22014 [==============================] - 0s 6us/step - loss: 0.4528 -
accuracy: 0.7966 - val_loss: 0.9072 - val_accuracy: 0.8007
Epoch 181/200
22014/22014 [==============================] - 0s 6us/step - loss: 0.4919 -
accuracy: 0.7975 - val_loss: 0.4952 - val_accuracy: 0.7963
Epoch 182/200
22014/22014 [==============================] - 0s 6us/step - loss: 0.4043 -
```

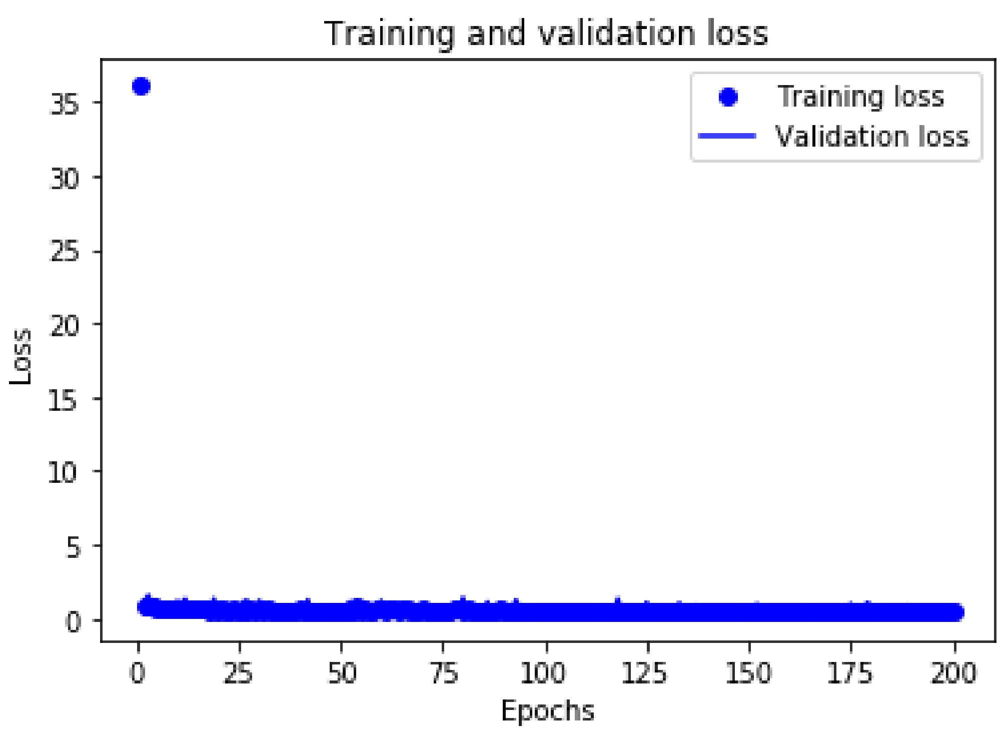
```
accuracy: 0.8037 - val_loss: 0.6241 - val_accuracy: 0.7745
Epoch 183/200
22014/22014 [==============================] - 0s 6us/step - loss: 0.4501 -
accuracy: 0.7959 - val_loss: 0.4391 - val_accuracy: 0.7988
Epoch 184/200
22014/22014 [==============================] - 0s 6us/step - loss: 0.4585 -
accuracy: 0.7989 - val_loss: 0.5457 - val_accuracy: 0.7923
Epoch 185/200
22014/22014 [==============================] - 0s 6us/step - loss: 0.4219 -
accuracy: 0.8005 - val_loss: 0.4445 - val_accuracy: 0.7797
Epoch 186/200
22014/22014 [==============================] - 0s 6us/step - loss: 0.4367 -
accuracy: 0.7998 - val_loss: 0.7778 - val_accuracy: 0.7857
Epoch 187/200
22014/22014 [==============================] - 0s 6us/step - loss: 0.4719 -
accuracy: 0.8009 - val_loss: 0.8100 - val_accuracy: 0.7568
Epoch 188/200
22014/22014 [==============================] - 0s 6us/step - loss: 0.4272 -
accuracy: 0.8008 - val_loss: 0.7881 - val_accuracy: 0.7525
Epoch 189/200
22014/22014 [==============================] - 0s 6us/step - loss: 0.4362 -
accuracy: 0.7973 - val_loss: 0.9261 - val_accuracy: 0.7908
Epoch 190/200
22014/22014 [==============================] - 0s 6us/step - loss: 0.4720 -
accuracy: 0.7994 - val_loss: 0.8012 - val_accuracy: 0.7692
Epoch 191/200
22014/22014 [==============================] - 0s 6us/step - loss: 0.4307 -
accuracy: 0.7980 - val_loss: 0.6511 - val_accuracy: 0.7641
Epoch 192/200
22014/22014 [==============================] - 0s 6us/step - loss: 0.4404 -
accuracy: 0.7995 - val_loss: 0.3875 - val_accuracy: 0.8057
Epoch 193/200
22014/22014 [==============================] - 0s 6us/step - loss: 0.4303 -
accuracy: 0.8022 - val_loss: 0.4513 - val_accuracy: 0.7816
Epoch 194/200
22014/22014 [==============================] - 0s 6us/step - loss: 0.5535 -
accuracy: 0.8020 - val_loss: 0.9022 - val_accuracy: 0.7264
Epoch 195/200
22014/22014 [==============================] - 0s 6us/step - loss: 0.4190 -
accuracy: 0.7978 - val_loss: 0.4612 - val_accuracy: 0.7912
Epoch 196/200
22014/22014 [==============================] - 0s 6us/step - loss: 0.4294 -
accuracy: 0.8030 - val_loss: 0.4430 - val_accuracy: 0.7957
Epoch 197/200
22014/22014 [==============================] - 0s 6us/step - loss: 0.4873 -
accuracy: 0.7936 - val_loss: 0.3824 - val_accuracy: 0.8125
Epoch 198/200
22014/22014 [==============================] - 0s 6us/step - loss: 0.4047 -
```

```
accuracy: 0.8053 - val_loss: 0.3887 - val_accuracy: 0.8103
Epoch 199/200
22014/22014 [==============================] - 0s 6us/step - loss: 0.4485 -
accuracy: 0.7984 - val_loss: 0.3704 - val_accuracy: 0.8160
Epoch 200/200
22014/22014 [==============================] - 0s 6us/step - loss: 0.4922 -
accuracy: 0.7963 - val_loss: 0.4279 - val_accuracy: 0.8055
```

After constructing our weighted model, we can plot the results of loss values from the training and validation set.

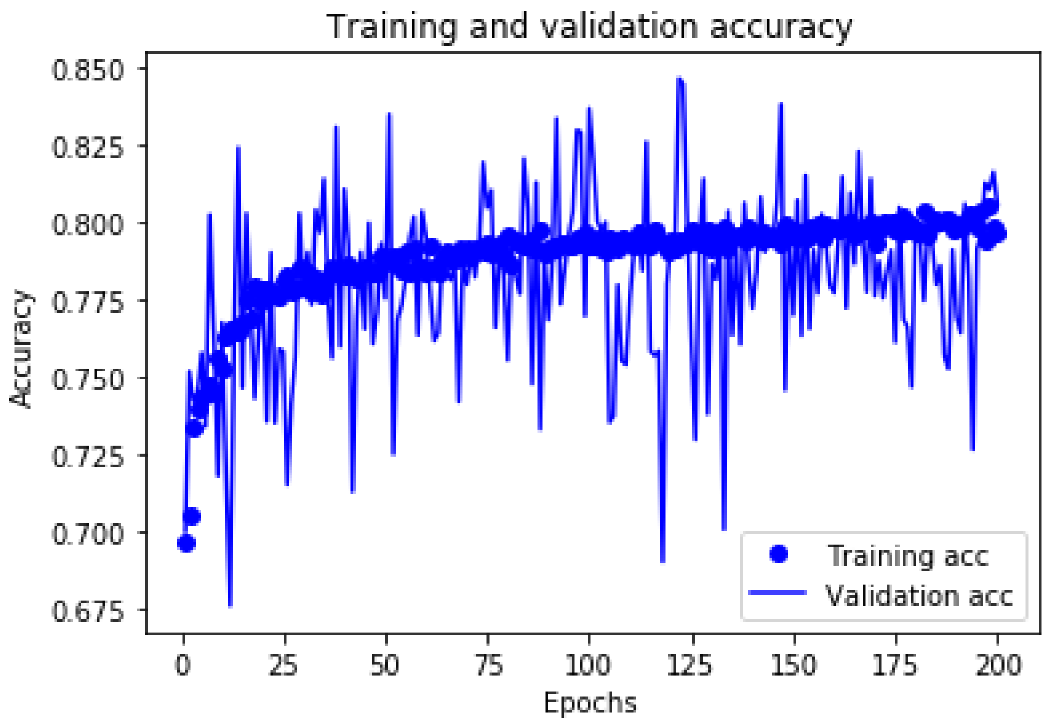
[23]:

```
# plot the results of loss values from the training set and validtion set
history_dict = history.history
loss_values = history_dict['loss']
val_loss_values = history_dict['val_loss']

epochs = range(1, len(history_dict['accuracy']) + 1)

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



Also, we can plot the results of accuracy from the training and validation set.

[24]:
```
#plt.clf()
acc = history_dict['accuracy']
val_acc = history_dict['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('Accuracy')
plt.legend()
plt.show()
```



When I used all the training data to train this model, the public score of this model is about 0.83 (the private score is also about 0.82). By using weighted model, we successfully overcome the problem of imbalanced training dataset.

Now, I can use all the training data to train this model.

[25]:
```
# Training the final model
model = models.Sequential()
model.add(layers.Dense(16, activation='relu'))
model.add(layers.Dense(16, activation='relu'))
```

```python
model.add(layers.Dense(1, activation='sigmoid'))

model.compile(optimizer='rmsprop',
              loss='binary_crossentropy',
              metrics=['accuracy'])

model.fit(X_train,y_train,
                    epochs=200,
                    batch_size=512,
                    class_weight=class_weights)
```

```
Epoch 1/200
29514/29514 [==============================] - 0s 11us/step - loss: 2.4969 -
accuracy: 0.2443
Epoch 2/200
29514/29514 [==============================] - 0s 5us/step - loss: 1.0225 -
accuracy: 0.7211
Epoch 3/200
29514/29514 [==============================] - 0s 6us/step - loss: 0.7428 -
accuracy: 0.7674
Epoch 4/200
29514/29514 [==============================] - 0s 5us/step - loss: 0.7877 -
accuracy: 0.7632
Epoch 5/200
29514/29514 [==============================] - 0s 5us/step - loss: 0.8176 -
accuracy: 0.7551
Epoch 6/200
29514/29514 [==============================] - 0s 5us/step - loss: 0.6618 -
accuracy: 0.7494
Epoch 7/200
29514/29514 [==============================] - 0s 5us/step - loss: 0.6786 -
accuracy: 0.7486
Epoch 8/200
29514/29514 [==============================] - 0s 5us/step - loss: 0.6707 -
accuracy: 0.7516
Epoch 9/200
29514/29514 [==============================] - 0s 5us/step - loss: 0.6971 -
accuracy: 0.7527
Epoch 10/200
29514/29514 [==============================] - 0s 5us/step - loss: 0.5900 -
accuracy: 0.7585
Epoch 11/200
29514/29514 [==============================] - 0s 5us/step - loss: 0.6250 -
accuracy: 0.7612
Epoch 12/200
29514/29514 [==============================] - 0s 5us/step - loss: 0.6953 -
accuracy: 0.7661
```

```
Epoch 13/200
29514/29514 [==============================] - 0s 5us/step - loss: 0.7290 -
accuracy: 0.7663
Epoch 14/200
29514/29514 [==============================] - 0s 5us/step - loss: 0.7789 -
accuracy: 0.7747
Epoch 15/200
29514/29514 [==============================] - 0s 5us/step - loss: 0.6507 -
accuracy: 0.7775
Epoch 16/200
29514/29514 [==============================] - 0s 5us/step - loss: 0.6451 -
accuracy: 0.7795
Epoch 17/200
29514/29514 [==============================] - 0s 5us/step - loss: 0.7095 -
accuracy: 0.7812
Epoch 18/200
29514/29514 [==============================] - 0s 5us/step - loss: 0.6720 -
accuracy: 0.7813
Epoch 19/200
29514/29514 [==============================] - 0s 5us/step - loss: 0.6728 -
accuracy: 0.7842
Epoch 20/200
29514/29514 [==============================] - 0s 5us/step - loss: 0.6980 -
accuracy: 0.7830
Epoch 21/200
29514/29514 [==============================] - 0s 5us/step - loss: 0.6185 -
accuracy: 0.7880
Epoch 22/200
29514/29514 [==============================] - 0s 5us/step - loss: 0.6557 -
accuracy: 0.7856
Epoch 23/200
29514/29514 [==============================] - 0s 5us/step - loss: 0.6749 -
accuracy: 0.7851
Epoch 24/200
29514/29514 [==============================] - 0s 5us/step - loss: 0.6326 -
accuracy: 0.7854
Epoch 25/200
29514/29514 [==============================] - 0s 5us/step - loss: 0.7143 -
accuracy: 0.7849
Epoch 26/200
29514/29514 [==============================] - 0s 5us/step - loss: 0.6363 -
accuracy: 0.7873
Epoch 27/200
29514/29514 [==============================] - 0s 5us/step - loss: 0.6473 -
accuracy: 0.7861
Epoch 28/200
29514/29514 [==============================] - 0s 5us/step - loss: 0.6043 -
accuracy: 0.7874
```

```
Epoch 29/200
29514/29514 [==============================] - 0s 5us/step - loss: 0.6664 -
accuracy: 0.7862
Epoch 30/200
29514/29514 [==============================] - 0s 5us/step - loss: 0.6005 -
accuracy: 0.7873
Epoch 31/200
29514/29514 [==============================] - 0s 5us/step - loss: 0.6526 -
accuracy: 0.7866
Epoch 32/200
29514/29514 [==============================] - 0s 5us/step - loss: 0.5707 -
accuracy: 0.7919
Epoch 33/200
29514/29514 [==============================] - 0s 5us/step - loss: 0.6258 -
accuracy: 0.7851
Epoch 34/200
29514/29514 [==============================] - 0s 6us/step - loss: 0.5892 -
accuracy: 0.7873
Epoch 35/200
29514/29514 [==============================] - 0s 8us/step - loss: 0.6526 -
accuracy: 0.7871
Epoch 36/200
29514/29514 [==============================] - 0s 8us/step - loss: 0.5644 -
accuracy: 0.7896
Epoch 37/200
29514/29514 [==============================] - 0s 7us/step - loss: 0.6396 -
accuracy: 0.7853
Epoch 38/200
29514/29514 [==============================] - 0s 7us/step - loss: 0.5428 -
accuracy: 0.7896
Epoch 39/200
29514/29514 [==============================] - 0s 7us/step - loss: 0.5967 -
accuracy: 0.7893
Epoch 40/200
29514/29514 [==============================] - 0s 6us/step - loss: 0.6315 -
accuracy: 0.7866
Epoch 41/200
29514/29514 [==============================] - 0s 12us/step - loss: 0.7360 -
accuracy: 0.7818
Epoch 42/200
29514/29514 [==============================] - 0s 9us/step - loss: 0.5369 -
accuracy: 0.7910
Epoch 43/200
29514/29514 [==============================] - 0s 7us/step - loss: 0.5897 -
accuracy: 0.7900
Epoch 44/200
29514/29514 [==============================] - 0s 6us/step - loss: 0.7010 -
accuracy: 0.7827
```

```
Epoch 45/200
29514/29514 [==============================] - 0s 6us/step - loss: 0.5194 -
accuracy: 0.7885
Epoch 46/200
29514/29514 [==============================] - 0s 6us/step - loss: 0.5953 -
accuracy: 0.7861
Epoch 47/200
29514/29514 [==============================] - 0s 7us/step - loss: 0.5561 -
accuracy: 0.7885
Epoch 48/200
29514/29514 [==============================] - 0s 7us/step - loss: 0.6185 -
accuracy: 0.7856
Epoch 49/200
29514/29514 [==============================] - 0s 7us/step - loss: 0.5096 -
accuracy: 0.7904
Epoch 50/200
29514/29514 [==============================] - 0s 6us/step - loss: 0.6016 -
accuracy: 0.7862
Epoch 51/200
29514/29514 [==============================] - 0s 6us/step - loss: 0.5732 -
accuracy: 0.7894
Epoch 52/200
29514/29514 [==============================] - 0s 7us/step - loss: 0.5825 -
accuracy: 0.7858
Epoch 53/200
29514/29514 [==============================] - 0s 6us/step - loss: 0.5734 -
accuracy: 0.7887
Epoch 54/200
29514/29514 [==============================] - 0s 7us/step - loss: 0.5663 -
accuracy: 0.7881
Epoch 55/200
29514/29514 [==============================] - 0s 6us/step - loss: 0.5670 -
accuracy: 0.7881
Epoch 56/200
29514/29514 [==============================] - 0s 6us/step - loss: 0.5778 -
accuracy: 0.7879
Epoch 57/200
29514/29514 [==============================] - 0s 7us/step - loss: 0.5794 -
accuracy: 0.7886
Epoch 58/200
29514/29514 [==============================] - 0s 6us/step - loss: 0.5440 -
accuracy: 0.7915
Epoch 59/200
29514/29514 [==============================] - 0s 10us/step - loss: 0.5683 -
accuracy: 0.7915
Epoch 60/200
29514/29514 [==============================] - 0s 14us/step - loss: 0.5496 -
accuracy: 0.7912
```

```
Epoch 61/200
29514/29514 [==============================] - 0s 6us/step - loss: 0.5322 -
accuracy: 0.7930
Epoch 62/200
29514/29514 [==============================] - 0s 7us/step - loss: 0.5490 -
accuracy: 0.7905
Epoch 63/200
29514/29514 [==============================] - 0s 7us/step - loss: 0.5406 -
accuracy: 0.7895
Epoch 64/200
29514/29514 [==============================] - 0s 6us/step - loss: 0.5442 -
accuracy: 0.7910
Epoch 65/200
29514/29514 [==============================] - 0s 7us/step - loss: 0.5710 -
accuracy: 0.7884
Epoch 66/200
29514/29514 [==============================] - 0s 6us/step - loss: 0.5616 -
accuracy: 0.7891
Epoch 67/200
29514/29514 [==============================] - 0s 7us/step - loss: 0.5211 -
accuracy: 0.7923
Epoch 68/200
29514/29514 [==============================] - 0s 6us/step - loss: 0.5815 -
accuracy: 0.7907
Epoch 69/200
29514/29514 [==============================] - 0s 7us/step - loss: 0.5249 -
accuracy: 0.7903
Epoch 70/200
29514/29514 [==============================] - 0s 7us/step - loss: 0.6096 -
accuracy: 0.7884
Epoch 71/200
29514/29514 [==============================] - 0s 7us/step - loss: 0.4853 -
accuracy: 0.7938
Epoch 72/200
29514/29514 [==============================] - 0s 7us/step - loss: 0.5340 -
accuracy: 0.7884
Epoch 73/200
29514/29514 [==============================] - 0s 6us/step - loss: 0.5940 -
accuracy: 0.7881
Epoch 74/200
29514/29514 [==============================] - 0s 9us/step - loss: 0.5186 -
accuracy: 0.7905
Epoch 75/200
29514/29514 [==============================] - 0s 7us/step - loss: 0.5796 -
accuracy: 0.7927
Epoch 76/200
29514/29514 [==============================] - 0s 6us/step - loss: 0.5374 -
accuracy: 0.7924
```

```
Epoch 77/200
29514/29514 [==============================] - 0s 6us/step - loss: 0.5206 -
accuracy: 0.7910
Epoch 78/200
29514/29514 [==============================] - 0s 6us/step - loss: 0.5410 -
accuracy: 0.7882
Epoch 79/200
29514/29514 [==============================] - 0s 6us/step - loss: 0.5392 -
accuracy: 0.7912
Epoch 80/200
29514/29514 [==============================] - 0s 7us/step - loss: 0.5730 -
accuracy: 0.7880
Epoch 81/200
29514/29514 [==============================] - 0s 7us/step - loss: 0.5303 -
accuracy: 0.7911
Epoch 82/200
29514/29514 [==============================] - 0s 7us/step - loss: 0.5390 -
accuracy: 0.7901
Epoch 83/200
29514/29514 [==============================] - 0s 7us/step - loss: 0.5141 -
accuracy: 0.7896
Epoch 84/200
29514/29514 [==============================] - 0s 6us/step - loss: 0.5009 -
accuracy: 0.7926
Epoch 85/200
29514/29514 [==============================] - 0s 7us/step - loss: 0.5200 -
accuracy: 0.7914
Epoch 86/200
29514/29514 [==============================] - 0s 7us/step - loss: 0.5227 -
accuracy: 0.7903
Epoch 87/200
29514/29514 [==============================] - 0s 6us/step - loss: 0.4901 -
accuracy: 0.7939
Epoch 88/200
29514/29514 [==============================] - 0s 7us/step - loss: 0.5285 -
accuracy: 0.7918
Epoch 89/200
29514/29514 [==============================] - 0s 7us/step - loss: 0.5062 -
accuracy: 0.7923
Epoch 90/200
29514/29514 [==============================] - 0s 7us/step - loss: 0.5067 -
accuracy: 0.7938
Epoch 91/200
29514/29514 [==============================] - 0s 6us/step - loss: 0.5771 -
accuracy: 0.7895
Epoch 92/200
29514/29514 [==============================] - 0s 6us/step - loss: 0.4958 -
accuracy: 0.7931
```

```
Epoch 93/200
29514/29514 [==============================] - 0s 6us/step - loss: 0.5400 -
accuracy: 0.7925
Epoch 94/200
29514/29514 [==============================] - 0s 6us/step - loss: 0.4981 -
accuracy: 0.7911
Epoch 95/200
29514/29514 [==============================] - 0s 6us/step - loss: 0.5032 -
accuracy: 0.7944
Epoch 96/200
29514/29514 [==============================] - 0s 7us/step - loss: 0.5437 -
accuracy: 0.7894
Epoch 97/200
29514/29514 [==============================] - 0s 6us/step - loss: 0.5148 -
accuracy: 0.7933
Epoch 98/200
29514/29514 [==============================] - 0s 6us/step - loss: 0.4797 -
accuracy: 0.7926
Epoch 99/200
29514/29514 [==============================] - 0s 5us/step - loss: 0.5382 -
accuracy: 0.7909
Epoch 100/200
29514/29514 [==============================] - 0s 5us/step - loss: 0.5360 -
accuracy: 0.7924
Epoch 101/200
29514/29514 [==============================] - 0s 5us/step - loss: 0.4825 -
accuracy: 0.7941
Epoch 102/200
29514/29514 [==============================] - 0s 6us/step - loss: 0.4718 -
accuracy: 0.7957
Epoch 103/200
29514/29514 [==============================] - 0s 5us/step - loss: 0.5090 -
accuracy: 0.7911
Epoch 104/200
29514/29514 [==============================] - 0s 5us/step - loss: 0.4999 -
accuracy: 0.7949
Epoch 105/200
29514/29514 [==============================] - 0s 5us/step - loss: 0.4991 -
accuracy: 0.7915
Epoch 106/200
29514/29514 [==============================] - 0s 6us/step - loss: 0.4900 -
accuracy: 0.7940
Epoch 107/200
29514/29514 [==============================] - 0s 7us/step - loss: 0.5446 -
accuracy: 0.7921
Epoch 108/200
29514/29514 [==============================] - 0s 7us/step - loss: 0.4974 -
accuracy: 0.7924
```

```
Epoch 109/200
29514/29514 [==============================] - 0s 6us/step - loss: 0.4465 -
accuracy: 0.7962
Epoch 110/200
29514/29514 [==============================] - 0s 6us/step - loss: 0.4783 -
accuracy: 0.7941
Epoch 111/200
29514/29514 [==============================] - 0s 6us/step - loss: 0.5792 -
accuracy: 0.7926
Epoch 112/200
29514/29514 [==============================] - 0s 5us/step - loss: 0.4671 -
accuracy: 0.7949
Epoch 113/200
29514/29514 [==============================] - 0s 6us/step - loss: 0.5194 -
accuracy: 0.7917
Epoch 114/200
29514/29514 [==============================] - 0s 6us/step - loss: 0.4543 -
accuracy: 0.7984
Epoch 115/200
29514/29514 [==============================] - 0s 6us/step - loss: 0.4787 -
accuracy: 0.7944
Epoch 116/200
29514/29514 [==============================] - 0s 6us/step - loss: 0.4725 -
accuracy: 0.7957
Epoch 117/200
29514/29514 [==============================] - 0s 6us/step - loss: 0.4857 -
accuracy: 0.7951
Epoch 118/200
29514/29514 [==============================] - 0s 6us/step - loss: 0.4873 -
accuracy: 0.7924
Epoch 119/200
29514/29514 [==============================] - 0s 7us/step - loss: 0.4635 -
accuracy: 0.7948
Epoch 120/200
29514/29514 [==============================] - 0s 6us/step - loss: 0.4583 -
accuracy: 0.7929
Epoch 121/200
29514/29514 [==============================] - 0s 6us/step - loss: 0.4652 -
accuracy: 0.7940
Epoch 122/200
29514/29514 [==============================] - 0s 6us/step - loss: 0.4953 -
accuracy: 0.7930
Epoch 123/200
29514/29514 [==============================] - 0s 8us/step - loss: 0.4936 -
accuracy: 0.7921
Epoch 124/200
29514/29514 [==============================] - 0s 6us/step - loss: 0.4563 -
accuracy: 0.7959
```

```
Epoch 125/200
29514/29514 [==============================] - 0s 7us/step - loss: 0.4798 -
accuracy: 0.7921
Epoch 126/200
29514/29514 [==============================] - 0s 6us/step - loss: 0.4544 -
accuracy: 0.7943
Epoch 127/200
29514/29514 [==============================] - 0s 7us/step - loss: 0.4715 -
accuracy: 0.7933
Epoch 128/200
29514/29514 [==============================] - 0s 7us/step - loss: 0.4801 -
accuracy: 0.7938
Epoch 129/200
29514/29514 [==============================] - 0s 9us/step - loss: 0.4594 -
accuracy: 0.7978
Epoch 130/200
29514/29514 [==============================] - 0s 9us/step - loss: 0.4605 -
accuracy: 0.7940
Epoch 131/200
29514/29514 [==============================] - 0s 8us/step - loss: 0.4734 -
accuracy: 0.7923
Epoch 132/200
29514/29514 [==============================] - 0s 6us/step - loss: 0.4741 -
accuracy: 0.7940
Epoch 133/200
29514/29514 [==============================] - 0s 7us/step - loss: 0.4755 -
accuracy: 0.7938
Epoch 134/200
29514/29514 [==============================] - 0s 7us/step - loss: 0.4546 -
accuracy: 0.7967
Epoch 135/200
29514/29514 [==============================] - 0s 7us/step - loss: 0.4530 -
accuracy: 0.7943
Epoch 136/200
29514/29514 [==============================] - 0s 7us/step - loss: 0.4593 -
accuracy: 0.7952
Epoch 137/200
29514/29514 [==============================] - 0s 8us/step - loss: 0.4437 -
accuracy: 0.7955
Epoch 138/200
29514/29514 [==============================] - 0s 6us/step - loss: 0.4850 -
accuracy: 0.7929
Epoch 139/200
29514/29514 [==============================] - 0s 10us/step - loss: 0.4367 -
accuracy: 0.7954
Epoch 140/200
29514/29514 [==============================] - 0s 7us/step - loss: 0.4624 -
accuracy: 0.7944
```

```
Epoch 141/200
29514/29514 [==============================] - 0s 7us/step - loss: 0.4498 -
accuracy: 0.7966
Epoch 142/200
29514/29514 [==============================] - 0s 8us/step - loss: 0.4432 -
accuracy: 0.7952
Epoch 143/200
29514/29514 [==============================] - 0s 8us/step - loss: 0.4683 -
accuracy: 0.7951
Epoch 144/200
29514/29514 [==============================] - 0s 8us/step - loss: 0.4401 -
accuracy: 0.7954
Epoch 145/200
29514/29514 [==============================] - 0s 9us/step - loss: 0.4562 -
accuracy: 0.7935
Epoch 146/200
29514/29514 [==============================] - 0s 8us/step - loss: 0.4283 -
accuracy: 0.7963
Epoch 147/200
29514/29514 [==============================] - 0s 9us/step - loss: 0.4710 -
accuracy: 0.7940
Epoch 148/200
29514/29514 [==============================] - 0s 7us/step - loss: 0.4399 -
accuracy: 0.7963
Epoch 149/200
29514/29514 [==============================] - 0s 6us/step - loss: 0.4598 -
accuracy: 0.7949
Epoch 150/200
29514/29514 [==============================] - 0s 6us/step - loss: 0.4573 -
accuracy: 0.7962
Epoch 151/200
29514/29514 [==============================] - 0s 8us/step - loss: 0.4442 -
accuracy: 0.7946
Epoch 152/200
29514/29514 [==============================] - 0s 8us/step - loss: 0.4334 -
accuracy: 0.7960
Epoch 153/200
29514/29514 [==============================] - 0s 10us/step - loss: 0.4359 -
accuracy: 0.7972
Epoch 154/200
29514/29514 [==============================] - 0s 9us/step - loss: 0.4517 -
accuracy: 0.7950
Epoch 155/200
29514/29514 [==============================] - 0s 11us/step - loss: 0.4346 -
accuracy: 0.7969
Epoch 156/200
29514/29514 [==============================] - 0s 8us/step - loss: 0.4455 -
accuracy: 0.7947
```

```
Epoch 157/200
29514/29514 [==============================] - 0s 8us/step - loss: 0.4227 -
accuracy: 0.7986
Epoch 158/200
29514/29514 [==============================] - 0s 7us/step - loss: 0.4300 -
accuracy: 0.7967
Epoch 159/200
29514/29514 [==============================] - 0s 13us/step - loss: 0.4742 -
accuracy: 0.7982
Epoch 160/200
29514/29514 [==============================] - 0s 7us/step - loss: 0.4289 -
accuracy: 0.7963
Epoch 161/200
29514/29514 [==============================] - 0s 7us/step - loss: 0.4237 -
accuracy: 0.7968
Epoch 162/200
29514/29514 [==============================] - 0s 6us/step - loss: 0.4368 -
accuracy: 0.7936
Epoch 163/200
29514/29514 [==============================] - 0s 6us/step - loss: 0.4239 -
accuracy: 0.7972
Epoch 164/200
29514/29514 [==============================] - 0s 6us/step - loss: 0.4096 -
accuracy: 0.7989
Epoch 165/200
29514/29514 [==============================] - 0s 6us/step - loss: 0.4634 -
accuracy: 0.7953
Epoch 166/200
29514/29514 [==============================] - 0s 6us/step - loss: 0.4330 -
accuracy: 0.7968
Epoch 167/200
29514/29514 [==============================] - 0s 6us/step - loss: 0.4395 -
accuracy: 0.7951
Epoch 168/200
29514/29514 [==============================] - 0s 6us/step - loss: 0.4091 -
accuracy: 0.8000
Epoch 169/200
29514/29514 [==============================] - 0s 6us/step - loss: 0.4721 -
accuracy: 0.7956
Epoch 170/200
29514/29514 [==============================] - 0s 6us/step - loss: 0.4171 -
accuracy: 0.7994
Epoch 171/200
29514/29514 [==============================] - 0s 6us/step - loss: 0.4296 -
accuracy: 0.7955
Epoch 172/200
29514/29514 [==============================] - 0s 6us/step - loss: 0.4300 -
accuracy: 0.7968
```

```
Epoch 173/200
29514/29514 [==============================] - 0s 6us/step - loss: 0.4149 -
accuracy: 0.7981
Epoch 174/200
29514/29514 [==============================] - 0s 6us/step - loss: 0.4120 -
accuracy: 0.8000
Epoch 175/200
29514/29514 [==============================] - 0s 5us/step - loss: 0.4362 -
accuracy: 0.7972
Epoch 176/200
29514/29514 [==============================] - 0s 6us/step - loss: 0.4339 -
accuracy: 0.7946
Epoch 177/200
29514/29514 [==============================] - 0s 8us/step - loss: 0.4267 -
accuracy: 0.7965
Epoch 178/200
29514/29514 [==============================] - 0s 6us/step - loss: 0.4039 -
accuracy: 0.7982
Epoch 179/200
29514/29514 [==============================] - 0s 6us/step - loss: 0.4362 -
accuracy: 0.7958
Epoch 180/200
29514/29514 [==============================] - 0s 6us/step - loss: 0.4225 -
accuracy: 0.7983
Epoch 181/200
29514/29514 [==============================] - 0s 6us/step - loss: 0.4103 -
accuracy: 0.7967
Epoch 182/200
29514/29514 [==============================] - 0s 6us/step - loss: 0.4141 -
accuracy: 0.7984
Epoch 183/200
29514/29514 [==============================] - 0s 6us/step - loss: 0.4683 -
accuracy: 0.7922
Epoch 184/200
29514/29514 [==============================] - 0s 6us/step - loss: 0.3851 -
accuracy: 0.8017
Epoch 185/200
29514/29514 [==============================] - 0s 7us/step - loss: 0.4485 -
accuracy: 0.7970
Epoch 186/200
29514/29514 [==============================] - 0s 7us/step - loss: 0.3967 -
accuracy: 0.7983
Epoch 187/200
29514/29514 [==============================] - 0s 6us/step - loss: 0.4105 -
accuracy: 0.7978
Epoch 188/200
29514/29514 [==============================] - 0s 6us/step - loss: 0.4160 -
accuracy: 0.7984
```

```
Epoch 189/200
29514/29514 [==============================] - 0s 6us/step - loss: 0.4399 -
accuracy: 0.7961
Epoch 190/200
29514/29514 [==============================] - 0s 6us/step - loss: 0.4008 -
accuracy: 0.7982
Epoch 191/200
29514/29514 [==============================] - 0s 5us/step - loss: 0.4342 -
accuracy: 0.7964
Epoch 192/200
29514/29514 [==============================] - 0s 6us/step - loss: 0.4019 -
accuracy: 0.7970
Epoch 193/200
29514/29514 [==============================] - 0s 6us/step - loss: 0.4654 -
accuracy: 0.7954
Epoch 194/200
29514/29514 [==============================] - 0s 7us/step - loss: 0.4043 -
accuracy: 0.7964
Epoch 195/200
29514/29514 [==============================] - 0s 6us/step - loss: 0.3987 -
accuracy: 0.7987
Epoch 196/200
29514/29514 [==============================] - 0s 6us/step - loss: 0.4506 -
accuracy: 0.7935
Epoch 197/200
29514/29514 [==============================] - 0s 6us/step - loss: 0.4081 -
accuracy: 0.7981
Epoch 198/200
29514/29514 [==============================] - 0s 6us/step - loss: 0.4105 -
accuracy: 0.7986
Epoch 199/200
29514/29514 [==============================] - 0s 6us/step - loss: 0.4280 -
accuracy: 0.7954
Epoch 200/200
29514/29514 [==============================] - 0s 6us/step - loss: 0.4013 -
accuracy: 0.7990
```

```
[25]: <keras.callbacks.callbacks.History at 0x11cbbb090>
```

After training our final model, we can then use this model to predict our final answer (use our test dataset).

```
[26]: #Using a trained network to generate predictions on new data
y_pred_probability=model.predict(X_test)
y_pred=(y_pred_probability>0.5).astype(int)
answer=pd.DataFrame(y_pred)
```

After some dataframe operations, we can then export our answer.

Now, let's see the prediction result of our final model.

The prediction result:

As we talked before, by using our final weighted model, in Kaggle competition, the public score of this model is about 0.83 (the private score is also about 0.82).

In my opinion, this model's great performance in Kaggle competition means this model can effectively predict whether a person makes over 50K a year. That is to say, if we want to predict if a person makes over 50K a year, our final model is trustable and reasonable.

Learning progress and reflection:

To be honest, when I first built a NN model for this assignment, the result was very bad. Even if I used the weighted model technique, I still got bad results (accuracy scores are very unstable).

In order to overcome this situation, I started to do explorative data analysis and I found that 'fnlwgt' is almost unrelated with 'Target'.

Therefore, I decided to drop 'fnlwgt'. Fortunately, my models started to improve and got trustable predictions.

In short, never forget to do EDA before modeling.

Feedback for the teaching team's reference:

Thanks for your teaching and helping. I have learned a lot from this journey.