▼ 빅데이터 기반 AI 응용 솔루션 개발자 전문과정

교과목명: 딥러닝알고리즘 구현

• 평가일: 2022. 11. 22

성명: 김영선점수: 100

25000 25000

from google.colab import drive
drive.mount('/content/drive')

Mounted at /content/drive

Q1. 사람이 문장을 읽는 것처럼 이전에 나온 것을 기억하면서 단어별로 또는 한눈에 들어오는 만큼씩 처리하여 문장에 있은 의미를 자연스럽게 표현하려는 목적으로 과거 정보를 사용하고 새롭게 얻은 정보를 계속 업데이트 하는 방식이 순환 신경망(RNN) 이다. SimpleRNN을 활용하여 IMDB 영화 리뷰 데이터에 대하여 아래 사항을 수행하세요.

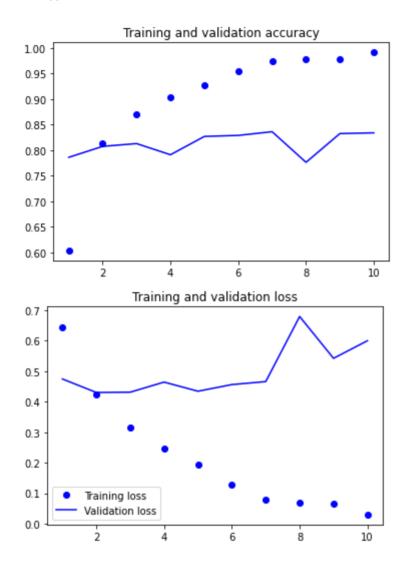
- 데이터 전처리: max_features 10000, maxlen = 500, batch_size 32
- 케라스를 사용하여 입력 시퀀스에 대한 마지막 출력만 반환하는 방식으로 모델링.(embedding 층 입력 (max_features, 32))
- 학습 및 검증 옵션 : epochs 10, batch_size 128, 검증 데이터 20% ※ 학습시간 20분
- 훈련과 검증의 손실과 정확도를 그래프로 표현
- 검증 정확도를 확인하고 동 사례에 SimpleRNN 모델의 적합 여부 및 개선 방안에 대하여 기술하세요.

```
(25000, 500)
(25000, 500)
```

```
from keras.layers import Dense
from keras.models import Sequential
from keras.layers import Embedding, SimpleRNN
model = Sequential()
model.add(Embedding(max features, 32))
model.add(SimpleRNN(32))
model.add(Dense(1,activation='sigmoid'))
model.compile(optimizer='rmsprop',loss='binary crossentropy',metrics=['acc'])
history = model.fit(input train, y train,
           epochs=10,
           batch size=128,
           validation split=0.2)
  Epoch 1/10
  Epoch 2/10
  Epoch 3/10
  Epoch 4/10
  Epoch 5/10
  Epoch 6/10
  Epoch 7/10
  Epoch 8/10
  Epoch 9/10
  Epoch 10/10
  import matplotlib.pyplot as plt
acc = history.history['acc']
val acc = history.history['val acc']
loss = history.history['loss']
val loss = history.history['val loss']
epochs = range(1, len(acc)+1)
plt.plot(epochs,acc,'bo',label='Training acc')
plt.plot(epochs,val acc,'b', label='Validation acc')
plt.title('Training and validation accuracy')
plt.figure()
plt.plot(epochs, loss, 'bo', label='Training loss')
```

```
plt.plot(epochs, val_loss, 'b', label='Validation loss')
plt.title('Training and validation loss')
plt.legend()
```



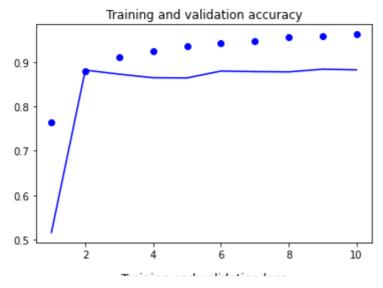


Q2. Q1 문제를 LSTM 모델을 적용하여 수행하세요

- 모델링, 학습 및 검증
- 결과 시각화

Epoch 1/10

```
Epoch 2/10
  Epoch 3/10
  Epoch 4/10
  Epoch 5/10
  Epoch 6/10
  Epoch 7/10
  Epoch 8/10
  Epoch 9/10
  Epoch 10/10
  import matplotlib.pyplot as plt
acc = history.history['acc']
val acc = history.history['val acc']
loss = history.history['loss']
val loss = history.history['val loss']
epochs = range(1, len(acc)+1)
plt.plot(epochs,acc,'bo',label='Training acc')
plt.plot(epochs,val acc,'b', label='Validation acc')
plt.title('Training and validation accuracy')
plt.figure()
plt.plot(epochs, loss, 'bo', label='Training loss')
plt.plot(epochs, val loss, 'b', label='Validation loss')
plt.title('Training and validation loss')
plt.legend()
plt.show()
```



O3. MNIST 숫자 이미지 데이터에 대하여 CNN 모델을 사용하여 아래사항을 수행하세요

- Conv2D와 MaxPooling2D 층을 사용하여 컨브넷을 생성(채널의 수 32개 또는 64개)
- 출력 텐서를 완전 연결 네트워크에 주입
- 10개의 클래스 분류하기 위한 분류기 추가
- 컨브넷 학습 및 평가

```
# 간단한 컨브넷 만들기
from keras import layers
from keras import models

model = models.Sequential()
model.add(layers.Conv2D(32, (3,3), activation='relu',input_shape=(28,28,1)))
model.add(layers.MaxPooling2D((2,2)))
model.add(layers.Conv2D(64, (3,3), activation='relu'))
model.add(layers.MaxPooling2D((2,2)))
model.add(layers.Conv2D(64, (3,3), activation='relu'))

# 전체 네트워크 확인
model.summary()
```

Model: "sequential 3"

| Layer (type) | Output Shape | Param # |
|--|--------------------|---------|
| conv2d (Conv2D) | (None, 26, 26, 32) | 320 |
| <pre>max_pooling2d (MaxPooling2D)</pre> | (None, 13, 13, 32) | 0 |
| conv2d_1 (Conv2D) | (None, 11, 11, 64) | 18496 |
| <pre>max_pooling2d_1 (MaxPooling 2D)</pre> | (None, 5, 5, 64) | 0 |
| conv2d_2 (Conv2D) | (None, 3, 3, 64) | 36928 |

Trainable params: 55,744 Non-trainable params: 0

```
# 컨브넷 위에 분류기 추가하기
model.add(layers.Flatten())
model.add(layers.Dense(64, activation='relu'))
model.add(layers.Dense(10, activation='softmax'))
# MNIST 이미지에 컨브넷 훈련하기
from keras.datasets import mnist
from keras.utils import to categorical
(train images, train labels), (test images, test labels) = mnist.load data()
train images = train images.reshape((60000, 28, 28, 1))
train images = train images.astype('float32') / 255 # 0과 1사이의 값을 가지는 float타입으로
test images = test images.reshape((10000, 28, 28, 1))
test images = test images.astype('float32') / 255
train labels = to categorical(train labels)
test labels = to categorical(test labels)
model.compile(optimizer='rmsprop', # 입력된 데이터와 손실함수를 기반으로 네트워크를 업데이트하는 머
          loss='categorical crossentropy', # 신경망의 성능을 측정하는 방법으로 네트워크가
          metrics=['accuracy']) # 훈련과 테스트 과정을 모니터링할 지표
model.fit(train images, train labels, epochs=5, batch size=64)
   Downloading data from <a href="https://storage.googleapis.com/tensorflow/tf-keras-datas">https://storage.googleapis.com/tensorflow/tf-keras-datas</a>
   Epoch 1/5
   Epoch 2/5
   Epoch 3/5
   Epoch 4/5
   Epoch 5/5
   938/938 [============== ] - 4s 5ms/step - loss: 0.0197 - accura
   <keras.callbacks.History at 0x7f9c1bd39e90>
# 테스트 데이터에서 모델 평가
test loss, test acc = model.evaluate(test images, test labels)
test_acc
   0.9902999997138977
```

Q4. cats_and_dogs_small으로 축소한 데이터 셋으로 사전 훈련된 네트워크를 사용하여 강아지 고양이 분류 과제를 아래와 같이 수행하세요.

- ImageNet 데이터셋에 훈련된 VGG16 네트워크의 합성곱 기반 층을 사용하여 유용한 특성 추출하고 이 특성으로 분류기 훈련
- ImageDataGenerator 사용 (※ 소요시간 20분)
- VGG 매개변수
 - weights는 모델을 초기화할 가중치 체크포인트를 지정 : 'imagenet'
 - ∘ include_top은 네트워크의 최상위 완전 연결 분류기를 포함할지 안할지를 지정 : False
 - ∘ input_shape은 네트워크에 주입할 이미지 텐서의 크기 :(150.150,3)
- 데이터 증식을 사용하지 않는 방법으로 수행
- 완전 연결 분류기를 정의하고 규제를 위해 드롭아웃 사용: 0.5

from tensorflow.keras.applications import VGG16

conv_base = VGG16(weights='imagenet', # imagenet으로 학습한 가중치 include_top=False, input_shape=(150, 150, 3))

conv_base.summary()

Model: "vgg16"

| Layer (type) | Output Shape | Param # |
|----------------------------|----------------------|---------|
| input_2 (InputLayer) | | |
| block1_conv1 (Conv2D) | (None, 150, 150, 64) | 1792 |
| block1_conv2 (Conv2D) | (None, 150, 150, 64) | 36928 |
| block1_pool (MaxPooling2D) | (None, 75, 75, 64) | 0 |
| block2_conv1 (Conv2D) | (None, 75, 75, 128) | 73856 |
| block2_conv2 (Conv2D) | (None, 75, 75, 128) | 147584 |
| block2_pool (MaxPooling2D) | (None, 37, 37, 128) | 0 |
| block3_conv1 (Conv2D) | (None, 37, 37, 256) | 295168 |
| block3_conv2 (Conv2D) | (None, 37, 37, 256) | 590080 |
| block3_conv3 (Conv2D) | (None, 37, 37, 256) | 590080 |
| block3_pool (MaxPooling2D) | (None, 18, 18, 256) | 0 |
| block4_conv1 (Conv2D) | (None, 18, 18, 512) | 1180160 |
| block4_conv2 (Conv2D) | (None, 18, 18, 512) | 2359808 |
| block4_conv3 (Conv2D) | (None, 18, 18, 512) | 2359808 |
| block4_pool (MaxPooling2D) | (None, 9, 9, 512) | 0 |
| block5_conv1 (Conv2D) | (None, 9, 9, 512) | 2359808 |

2359808

(None, 9, 9, 512)

block5 conv2 (Conv2D)

```
block5 conv3 (Conv2D)
                          (None, 9, 9, 512)
                                                       2359808
     block5 pool (MaxPooling2D) (None, 4, 4, 512)
    _____
    Total params: 14,714,688
    Trainable params: 14,714,688
    Non-trainable params: 0
%cd /content/drive/MyDrive/강의/m9 딥러닝 기본/dataset
    /content/drive/MyDrive/강의/m9 딥러닝 기본/dataset
!ls
    aclImdb.zip
                        cats and dogs small 1.h5 glove.6B.100d.txt
    cats and dogs small cats and dogs small 2.h5 glove.6B.zip
# 데이터 내려받기
import os, shutil
import numpy as np
from tensorflow.keras.preprocessing.image import ImageDataGenerator
base dir = 'cats and dogs small'
train dir = os.path.join(base dir, 'train')
validation dir = os.path.join(base dir, 'validation')
test dir = os.path.join(base dir,'test')
datagen = ImageDataGenerator(rescale=1./255)
batch size = 20
def extract features(directory, sample count):
   features = np.zeros(shape=(sample count, 4, 4, 512))
   labels = np.zeros(shape=(sample count))
   generator = datagen.flow from directory(
       directory,
       target size=(150, 150),
       batch size=batch size,
       class mode='binary')
   i = 0
   for inputs batch, labels batch in generator:
       features batch = conv base.predict(inputs batch) # conv base에 데이터를 주입하고
       features[i * batch size : (i + 1) * batch size] = features batch
       labels[i * batch size : (i + 1) * batch size] = labels batch
       if i * batch size >= sample count:
           # 제너레이터는 루프 안에서 무한하게 데이터를 만들어내므로 모든 이미지를 한 번씩 처리하고 나면
           break
   return features, labels
```

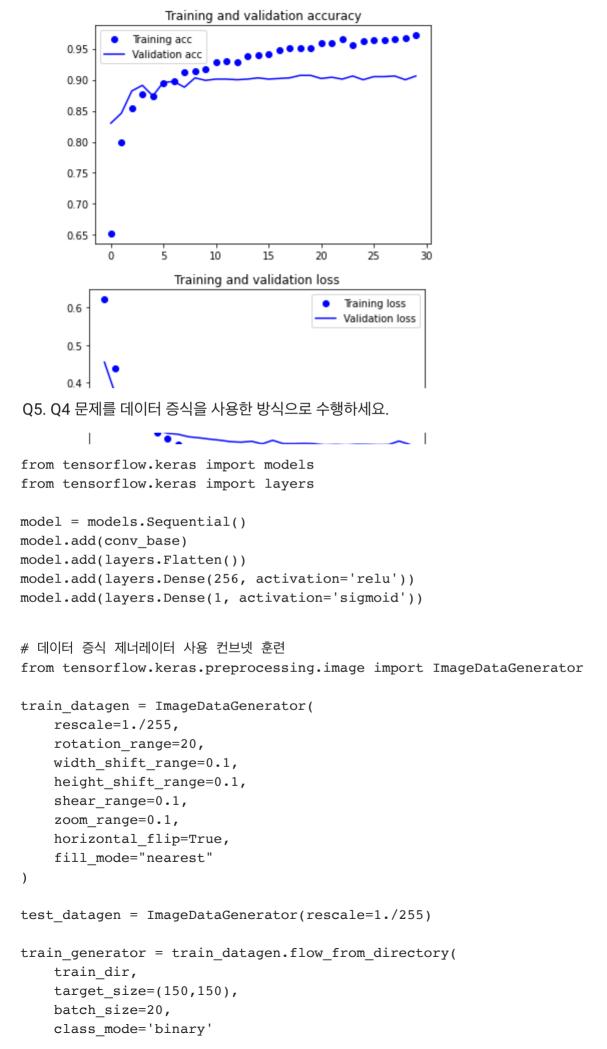
train_features, train_labels = extract_features(train_dir, 2000)
validation_features, validation_labels = extract_features(validation_dir, 1000)
test_features, test_labels = extract_features(test_dir, 1000)

```
1/1 |-----| - Va 21ma/acep
1/1 [=======] - 0s 24ms/step
1/1 [======= ] - 0s 21ms/step
1/1 [======] - 0s 20ms/step
1/1 [======= ] - 0s 20ms/step
1/1 [======= ] - 0s 24ms/step
1/1 [=======] - 0s 19ms/step
1/1 [======] - 0s 22ms/step
1/1 [======= 1 - 0s 23ms/step
1/1 [=======] - 0s 20ms/step
1/1 [======] - 0s 19ms/step
1/1 [======= ] - 0s 21ms/step
1/1 [=======] - 0s 20ms/step
1/1 [======] - 0s 21ms/step
1/1 [======= ] - 0s 19ms/step
1/1 [======] - 0s 24ms/step
1/1 [======= ] - 0s 19ms/step
1/1 [======= ] - 0s 25ms/step
1/1 [======] - 0s 20ms/step
1/1 [=======] - 0s 18ms/step
1/1 [======= ] - 0s 23ms/step
1/1 [=======] - 0s 18ms/step
1/1 [=======] - 0s 18ms/step
1/1 [======] - 0s 18ms/step
1/1 [=======] - 0s 20ms/step
1/1 [======] - 0s 19ms/step
1/1 [======] - 0s 21ms/step
1/1 [======] - 0s 27ms/step
1/1 [======] - 0s 20ms/step
1/1 [======] - 0s 20ms/step
1/1 [====== ] - 0s 21ms/step
1/1 [=======] - 0s 21ms/step
1/1 [======] - 0s 21ms/step
1/1 [=======] - 0s 19ms/step
1/1 [======] - 0s 20ms/step
1/1 [======= ] - 0s 21ms/step
1/1 [======] - 0s 21ms/step
1/1 [======] - 0s 18ms/step
1/1 [======] - 0s 20ms/step
1/1 [======] - 0s 19ms/step
1/1 [=======] - 0s 19ms/step
1/1 [======] - 0s 19ms/step
1/1 [======] - 0s 22ms/step
1/1 [======] - 0s 19ms/step
1/1 [======] - 0s 20ms/step
1/1 [======= ] - 0s 30ms/step
1/1 [=======] - 0s 20ms/step
1/1 [======] - 0s 19ms/step
1/1 [======= ] - 0s 18ms/step
1/1 [======] - 0s 19ms/step
1/1 [======] - 0s 19ms/step
1/1 [======] - 0s 19ms/step
1/1 [======] - 0s 21ms/step
1/1 [======] - 0s 20ms/step
1/1 [======] - 0s 20ms/step
1/1 [======] - 0s 20ms/step
```

Found 1000 images belonging to 2 classes.

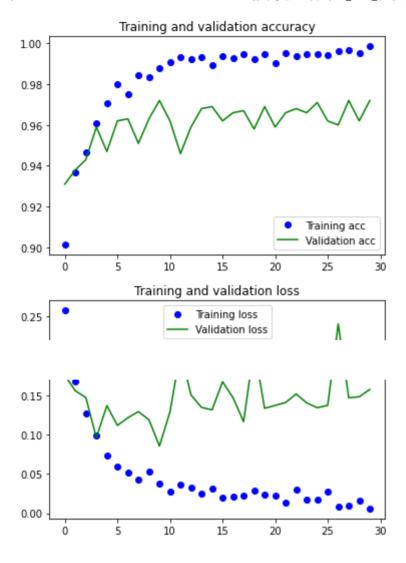
```
1/1 [======] - 0s 19ms/step
  1/1 [======] - 0s 23ms/step
train features = np.reshape(train features, (2000, 4 * 4 * 512))
validation features = np.reshape(validation_features, (1000, 4 * 4 * 512))
test features = np.reshape(test features, (1000, 4 * 4 * 512))
from tensorflow.keras import models
from tensorflow.keras import layers
from tensorflow.keras import optimizers
model = models.Sequential()
model.add(layers.Dense(256, activation='relu', input dim=4 * 4 * 512))
model.add(layers.Dropout(0.5))
model.add(layers.Dense(1, activation='sigmoid'))
model.compile(optimizer=optimizers.RMSprop(lr=2e-5),
      loss='binary crossentropy',
      metrics=['acc'])
history = model.fit(train features, train labels,
         epochs=30,
         batch size=20,
         validation data=(validation features, validation labels))
  Epoch 1/30
  Epoch 2/30
  Epoch 3/30
  Epoch 4/30
  Epoch 5/30
  Epoch 6/30
  Epoch 7/30
  Epoch 8/30
  Epoch 9/30
  100/100 [============] - 0s 4ms/step - loss: 0.2260 - acc: C
  Epoch 10/30
  Epoch 11/30
  Epoch 12/30
  Epoch 13/30
  Epoch 14/30
  Epoch 15/30
  Epoch 16/30
```

```
Epoch 17/30
  Epoch 18/30
  Epoch 19/30
  Epoch 20/30
  Epoch 21/30
  Epoch 22/30
  Epoch 23/30
  Epoch 24/30
  Epoch 25/30
  100/100 [============== ] - 0s 5ms/step - loss: 0.1129 - acc: C
  Epoch 26/30
  Epoch 27/30
  100/100 [============== ] - 0s 4ms/step - loss: 0.1080 - acc: C
  Epoch 28/30
  100/100 [=============== ] - 0s 4ms/step - loss: 0.1013 - acc: C
  Epoch 29/30
  import matplotlib.pyplot as plt
acc = history.history['acc']
val acc = history.history['val acc']
loss = history.history['loss']
val loss = history.history['val loss']
epochs = range(len(acc))
plt.plot(epochs, acc, 'bo', label='Training acc')
plt.plot(epochs, val acc, 'b', label='Validation acc')
plt.title('Training and validation accuracy')
plt.legend()
plt.figure()
plt.plot(epochs, loss, 'bo', label='Training loss')
plt.plot(epochs, val loss, 'b', label='Validation loss')
plt.title('Training and validation loss')
plt.legend()
plt.show()
```



```
22. 11. 22. 오후 4:15
                                    딥러닝알고리즘구현_1118_김영선.ipynb - Colaboratory
   )
   validation generator = test datagen.flow from directory(
       validation dir,
       target size=(150,150),
       batch size=20,
       class mode='binary'
   )
   model.compile(loss='binary crossentropy',
                 optimizer=optimizers.RMSprop(lr=2e-5),
                 metrics=['acc'])
   history = model.fit generator(
         train generator,
         steps per epoch=100,
         epochs=30,
         validation data=validation generator,
         validation steps=50,
         verbose=2)
        Found 2000 images belonging to 2 classes.
        Found 1000 images belonging to 2 classes.
        /usr/local/lib/python3.7/dist-packages/keras/optimizers/optimizer v2/rmsprop.r
         super(RMSprop, self). init (name, **kwargs)
        /usr/local/lib/python3.7/dist-packages/ipykernel launcher.py:41: UserWarning:
        Epoch 1/30
        100/100 - 33s - loss: 0.2578 - acc: 0.9015 - val loss: 0.1754 - val acc: 0.931
        Epoch 2/30
        100/100 - 28s - loss: 0.1676 - acc: 0.9370 - val loss: 0.1557 - val acc: 0.938
        Epoch 3/30
        100/100 - 28s - loss: 0.1274 - acc: 0.9465 - val loss: 0.1471 - val acc: 0.943
        Epoch 4/30
        100/100 - 27s - loss: 0.0985 - acc: 0.9610 - val loss: 0.0961 - val acc: 0.959
        Epoch 5/30
        100/100 - 28s - loss: 0.0729 - acc: 0.9705 - val loss: 0.1368 - val acc: 0.947
       Epoch 6/30
        100/100 - 28s - loss: 0.0594 - acc: 0.9800 - val loss: 0.1116 - val acc: 0.962
        Epoch 7/30
        100/100 - 27s - loss: 0.0521 - acc: 0.9750 - val loss: 0.1215 - val acc: 0.963
        Epoch 8/30
        100/100 - 27s - loss: 0.0431 - acc: 0.9845 - val loss: 0.1292 - val acc: 0.951
       Epoch 9/30
        100/100 - 28s - loss: 0.0533 - acc: 0.9835 - val loss: 0.1189 - val acc: 0.963
        Epoch 10/30
        100/100 - 27s - loss: 0.0372 - acc: 0.9880 - val loss: 0.0854 - val acc: 0.972
        Epoch 11/30
        100/100 - 27s - loss: 0.0275 - acc: 0.9910 - val loss: 0.1288 - val acc: 0.962
       Epoch 12/30
        100/100 - 28s - loss: 0.0360 - acc: 0.9930 - val loss: 0.2058 - val acc: 0.946
       Epoch 13/30
        100/100 - 27s - loss: 0.0326 - acc: 0.9920 - val loss: 0.1503 - val acc: 0.959
       Epoch 14/30
        100/100 - 28s - loss: 0.0250 - acc: 0.9930 - val loss: 0.1345 - val acc: 0.968
        Epoch 15/30
        100/100 - 28s - loss: 0.0308 - acc: 0.9895 - val loss: 0.1314 - val acc: 0.969
        Epoch 16/30
        100/100 - 28s - loss: 0.0193 - acc: 0.9935 - val_loss: 0.1673 - val_acc: 0.962
```

```
Epoch 17/30
    100/100 - 27s - loss: 0.0215 - acc: 0.9925 - val loss: 0.1468 - val acc: 0.966
    Epoch 18/30
    100/100 - 28s - loss: 0.0224 - acc: 0.9945 - val_loss: 0.1164 - val acc: 0.967
    Epoch 19/30
    100/100 - 27s - loss: 0.0285 - acc: 0.9920 - val_loss: 0.2115 - val acc: 0.958
    Epoch 20/30
    100/100 - 28s - loss: 0.0234 - acc: 0.9945 - val loss: 0.1334 - val acc: 0.969
    Epoch 21/30
    100/100 - 27s - loss: 0.0227 - acc: 0.9905 - val loss: 0.1370 - val acc: 0.959
    Epoch 22/30
    100/100 - 27s - loss: 0.0136 - acc: 0.9950 - val loss: 0.1408 - val acc: 0.966
    Epoch 23/30
    100/100 - 29s - loss: 0.0294 - acc: 0.9935 - val loss: 0.1517 - val acc: 0.968
    Epoch 24/30
    100/100 - 27s - loss: 0.0177 - acc: 0.9945 - val loss: 0.1403 - val acc: 0.966
    Epoch 25/30
    100/100 - 28s - loss: 0.0167 - acc: 0.9945 - val loss: 0.1342 - val acc: 0.971
    Epoch 26/30
    100/100 - 27s - loss: 0.0277 - acc: 0.9940 - val loss: 0.1371 - val acc: 0.962
    Epoch 27/30
model.save('/content/drive/MyDrive/강의/m9 딥러닝 기본/dataset/cats and dogs small 2.h5
acc = history.history['acc']
val acc = history.history['val acc']
loss = history.history['loss']
val_loss = history.history['val loss']
epochs = range(len(acc))
plt.plot(epochs, acc, 'bo', label='Training acc')
plt.plot(epochs, val acc, 'g', label='Validation acc')
plt.title('Training and validation accuracy')
plt.legend()
plt.figure()
plt.plot(epochs, loss, 'bo', label='Training loss')
plt.plot(epochs, val loss, 'g', label='Validation loss')
plt.title('Training and validation loss')
plt.legend()
plt.show()
```



Colab 유료 제품 - 여기에서 계약 취소

✓ 1초 오후 4:13에 완료됨