

[illegible]

[illegible]

```
TRUE
## [827] TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE
TRUE
## [841] TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE
TRUE
## [855] TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE
TRUE
## [869] TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE
TRUE
## [883] TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE
TRUE
## [897] TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE
TRUE
## [911] TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE
TRUE
## [925] TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE
TRUE
## [939] TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE
TRUE
## [953] TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE
TRUE
## [967] TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE
TRUE
## [981] TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE
TRUE
## [995] TRUE TRUE TRUE TRUE TRUE TRUE

fnames <- paste0("cat.", 1001:1500, ".jpg")
file.copy(file.path(original_dataset_dir, fnames),
          file.path(validation_cats_dir))

## [1] TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE
TRUE TRUE
## [16] TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE
TRUE TRUE
## [31] TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE
TRUE TRUE
## [46] TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE
TRUE TRUE
## [61] TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE
TRUE TRUE
## [76] TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE
TRUE TRUE
## [91] TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE
TRUE TRUE
## [106] TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE
TRUE TRUE
## [121] TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE
TRUE TRUE
## [136] TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE
```

[illegible]

```
fnames <- paste0("cat.", 1501:2000, ".jpg")
file.copy(file.path(original_dataset_dir, fnames),
          file.path(test_cats_dir))
```

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

```
## [406] TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE
TRUE TRUE
## [421] TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE
TRUE TRUE
## [436] TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE
TRUE TRUE
## [451] TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE
TRUE TRUE
## [466] TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE
TRUE TRUE
## [481] TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE
TRUE TRUE
## [496] TRUE TRUE TRUE TRUE TRUE
```

Split the data into training, test and validation sets and verified the number of images in each.

```
cat("total training cat images:", length(list.files(train_cats_dir)), "\n")
## total training cat images: 1000

cat("total training dog images:", length(list.files(train_dogs_dir)), "\n")
## total training dog images: 1000

cat("total validation cat images:", length(list.files(validation_cats_dir)),
"\n")
## total validation cat images: 500

cat("total validation dog images:", length(list.files(validation_dogs_dir)),
"\n")
## total validation dog images: 500

cat("total test cat images:", length(list.files(test_cats_dir)), "\n")
## total test cat images: 500

  cat("total test dog images:", length(list.files(test_dogs_dir)), "\n")
## total test dog images: 500

library(keras)

## Warning: package 'keras' was built under R version 4.1.2

model <- keras_model_sequential() %>%
  layer_conv_2d(filters = 32, kernel_size = c(3, 3), activation = "relu",
    input_shape = c(150, 150, 3)) %>%
  layer_max_pooling_2d(pool_size = c(2, 2)) %>%
  layer_conv_2d(filters = 64, kernel_size = c(3, 3), activation = "relu") %>%
  layer_max_pooling_2d(pool_size = c(2, 2)) %>%
```

```

    layer_conv_2d(filters = 128, kernel_size = c(3, 3), activation = "relu")
%>%
    layer_max_pooling_2d(pool_size = c(2, 2)) %>%
    layer_conv_2d(filters = 128, kernel_size = c(3, 3), activation = "relu")
%>%
    layer_max_pooling_2d(pool_size = c(2, 2)) %>%
    layer_flatten() %>%
    layer_dense(units = 512, activation = "relu") %>%
    layer_dense(units = 1, activation = "sigmoid")

```

```
## Loaded Tensorflow version 2.8.0
```

```
summary(model)
```

```
## Model: "sequential"
```

```
##
```

## Layer (type)	Output Shape	Param
##		
=====		
===		
## conv2d_3 (Conv2D)	(None, 148, 148, 32)	896
##		
## max_pooling2d_3 (MaxPooling2D)	(None, 74, 74, 32)	0
##		
## conv2d_2 (Conv2D)	(None, 72, 72, 64)	18496
##		
## max_pooling2d_2 (MaxPooling2D)	(None, 36, 36, 64)	0
##		
## conv2d_1 (Conv2D)	(None, 34, 34, 128)	73856
##		
## max_pooling2d_1 (MaxPooling2D)	(None, 17, 17, 128)	0
##		
## conv2d (Conv2D)	(None, 15, 15, 128)	147584
##		
## max_pooling2d (MaxPooling2D)	(None, 7, 7, 128)	0
##		
## flatten (Flatten)	(None, 6272)	0
##		
## dense_1 (Dense)	(None, 512)	
3211776		
##		
## dense (Dense)	(None, 1)	513
##		
##		
=====		
===		
## Total params: 3,453,121		

```

## Trainable params: 3,453,121
## Non-trainable params: 0
##

model %>% compile(
  loss = "binary_crossentropy",
  optimizer = optimizer_rmsprop(learning_rate = 1e-4),
  metrics = c("acc")
)

train_datagen <- image_data_generator(rescale = 1/255)
validation_datagen <- image_data_generator(rescale = 1/255)

train_generator <- flow_images_from_directory(
  train_dir,
  train_datagen,
  target_size = c(150, 150),
  batch_size = 20,
  class_mode = "binary"
)

validation_generator <- flow_images_from_directory(
  validation_dir,
  validation_datagen,
  target_size = c(150, 150),
  batch_size = 20,
  class_mode = "binary"
)

batch <- generator_next(train_generator)
str(batch)

## List of 2
## $ : num [1:20, 1:150, 1:150, 1:3] 0.886 0.31 0.235 0.918 0.882 ...
## $ : num [1:20(1d)] 0 1 0 0 1 1 1 1 1 1 ...

history <- model %>% fit_generator(
  train_generator,
  steps_per_epoch = 100,
  epochs = 30,
  validation_data = validation_generator,
  validation_steps = 50
)

## Warning in fit_generator(., train_generator, steps_per_epoch = 100, epochs
## = 30, : `fit_generator` is deprecated. Use `fit` instead, it now accept
## generators.

```

Built and visualized the original model

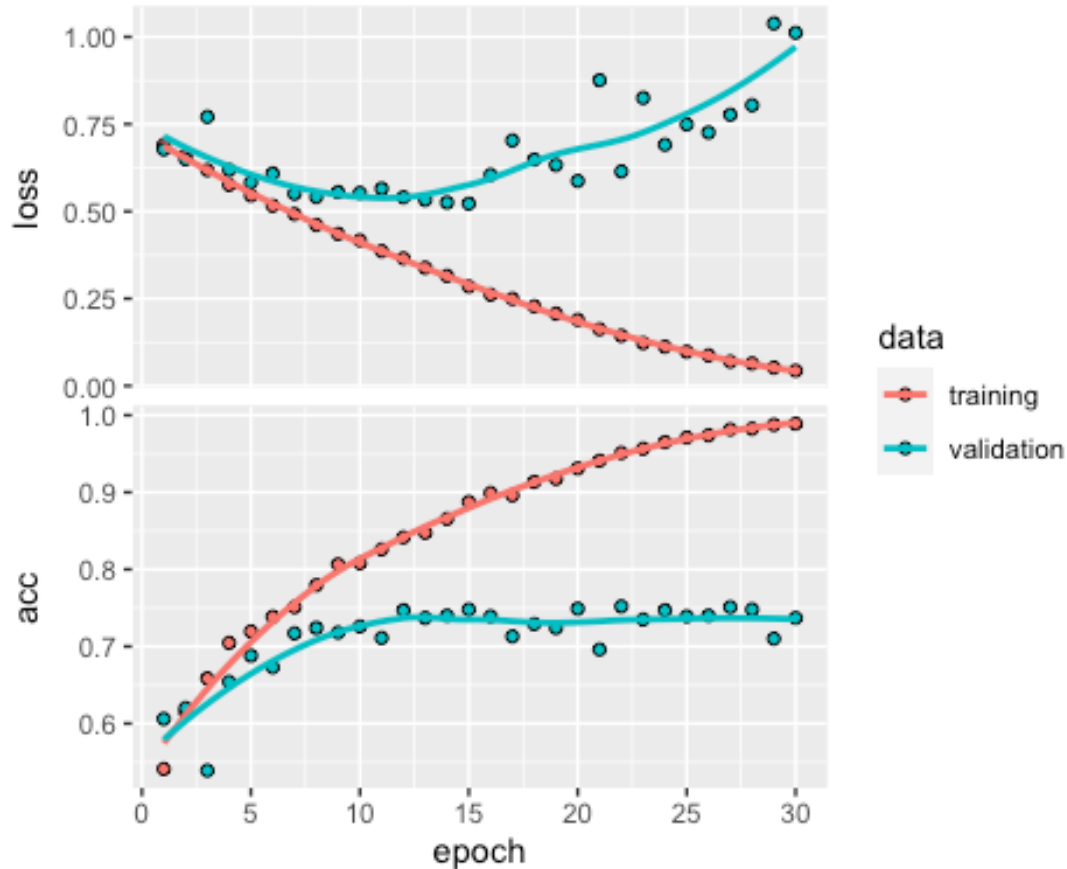
```

model %>% save_model_hdf5("cats_and_dogs_small_1.h5")

plot(history)

## `geom_smooth()` using formula 'y ~ x'

```



Created

a second model using data augmentation.

```

model2 <- keras_model_sequential() %>%
  layer_conv_2d(filters = 32, kernel_size = c(3, 3), activation = "relu",
    input_shape = c(150, 150, 3)) %>%
  layer_max_pooling_2d(pool_size = c(2, 2)) %>%
  layer_conv_2d(filters = 64, kernel_size = c(3, 3), activation = "relu") %>%
  layer_max_pooling_2d(pool_size = c(2, 2)) %>%
  layer_conv_2d(filters = 128, kernel_size = c(3, 3), activation = "relu")
%>%
  layer_max_pooling_2d(pool_size = c(2, 2)) %>%
  layer_conv_2d(filters = 128, kernel_size = c(3, 3), activation = "relu")
%>%
  layer_max_pooling_2d(pool_size = c(2, 2)) %>%
  layer_flatten() %>%
  layer_dropout(rate = 0.5) %>%
  layer_dense(units = 512, activation = "relu") %>%
  layer_dense(units = 1, activation = "sigmoid")

```



```

model2 %>% compile(
  loss = "binary_crossentropy",
  optimizer = optimizer_rmsprop(learning_rate = 1e-4),
  metrics = c("acc")
)

datagen <- image_data_generator(
  rescale = 1/255,
  rotation_range = 40,
  width_shift_range = 0.2,
  height_shift_range = 0.2,
  shear_range = 0.2,
  zoom_range = 0.2,
  horizontal_flip = TRUE
)

test_datagen <- image_data_generator(rescale = 1/255)

train_generator <- flow_images_from_directory(
  train_dir,
  datagen,
  target_size = c(150, 150),
  batch_size = 32,
  class_mode = "binary"
)

validation_generator <- flow_images_from_directory(
  validation_dir,
  test_datagen,
  target_size = c(150, 150),
  batch_size = 32,
  class_mode = "binary"
)

history <- model2 %>% fit(
  train_generator,
  epochs = 30,
  validation_data = validation_generator,
)

base_dir2 <- "~/Documents/R/cats.vs.dogs2"
dir.create(base_dir2)
train_dir2 <- file.path(base_dir2, "train")
dir.create(train_dir2)
validation_dir2 <- file.path(base_dir2, "validation")
dir.create(validation_dir2)
test_dir2 <- file.path(base_dir2, "test")
dir.create(test_dir2)

```


[illegible]

[illegible]

[illegible]


```
TRUE TRUE
## [286] TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE
TRUE TRUE
## [301] TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE
TRUE TRUE
## [316] TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE
TRUE TRUE
## [331] TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE
TRUE TRUE
## [346] TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE
TRUE TRUE
## [361] TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE
TRUE TRUE
## [376] TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE
TRUE TRUE
## [391] TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE
TRUE TRUE
## [406] TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE
TRUE TRUE
## [421] TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE
TRUE TRUE
## [436] TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE
TRUE TRUE
## [451] TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE
TRUE TRUE
## [466] TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE
TRUE TRUE
## [481] TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE
TRUE TRUE
## [496] TRUE TRUE TRUE TRUE TRUE

fnames <- paste0("cat.", 1501:2000, ".jpg")
file.copy(file.path(original_dataset_dir, fnames),
          file.path(test_cats_dir2))

## [1] TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE
TRUE TRUE
## [16] TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE
TRUE TRUE
## [31] TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE
TRUE TRUE
## [46] TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE
TRUE TRUE
## [61] TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE
TRUE TRUE
## [76] TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE
TRUE TRUE
## [91] TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE
TRUE TRUE
## [106] TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE
```

[illegible]


```
TRUE TRUE
## [496] TRUE TRUE TRUE TRUE TRUE

fnames <- paste0("dog.", 1:1300, ".jpg")
file.copy(file.path(original_dataset_dir, fnames),
          file.path(train_dogs_dir2))

## [1] TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE
TRUE
## [15] TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE
TRUE
## [29] TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE
TRUE
## [43] TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE
TRUE
## [57] TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE
TRUE
## [71] TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE
TRUE
## [85] TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE
TRUE
## [99] TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE
TRUE
## [113] TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE
TRUE
## [127] TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE
TRUE
## [141] TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE
TRUE
## [155] TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE
TRUE
## [169] TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE
TRUE
## [183] TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE
TRUE
## [197] TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE
TRUE
## [211] TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE
TRUE
## [225] TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE
TRUE
## [239] TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE
TRUE
## [253] TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE
TRUE
## [267] TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE
TRUE
## [281] TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE
TRUE
## [295] TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE
```

[illegible]

[illegible]

[illegible]

```
TRUE TRUE
## [406] TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE
TRUE TRUE
## [421] TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE
TRUE TRUE
## [436] TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE
TRUE TRUE
## [451] TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE
TRUE TRUE
## [466] TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE
TRUE TRUE
## [481] TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE
TRUE TRUE
## [496] TRUE TRUE TRUE TRUE TRUE

fnames <- paste0("dog.", 1501:2000, ".jpg")
file.copy(file.path(original_dataset_dir, fnames),
          file.path(test_dogs_dir2))

## [1] TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE
TRUE TRUE
## [16] TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE
TRUE TRUE
## [31] TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE
TRUE TRUE
## [46] TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE
TRUE TRUE
## [61] TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE
TRUE TRUE
## [76] TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE
TRUE TRUE
## [91] TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE
TRUE TRUE
## [106] TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE
TRUE TRUE
## [121] TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE
TRUE TRUE
## [136] TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE
TRUE TRUE
## [151] TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE
TRUE TRUE
## [166] TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE
TRUE TRUE
## [181] TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE
TRUE TRUE
## [196] TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE
TRUE TRUE
## [211] TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE
TRUE TRUE
## [226] TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE
```

```

TRUE TRUE
## [241] TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE
TRUE TRUE
## [256] TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE
TRUE TRUE
## [271] TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE
TRUE TRUE
## [286] TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE
TRUE TRUE
## [301] TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE
TRUE TRUE
## [316] TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE
TRUE TRUE
## [331] TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE
TRUE TRUE
## [346] TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE
TRUE TRUE
## [361] TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE
TRUE TRUE
## [376] TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE
TRUE TRUE
## [391] TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE
TRUE TRUE
## [406] TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE
TRUE TRUE
## [421] TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE
TRUE TRUE
## [436] TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE
TRUE TRUE
## [451] TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE
TRUE TRUE
## [466] TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE
TRUE TRUE
## [481] TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE
TRUE TRUE
## [496] TRUE TRUE TRUE TRUE TRUE

```

Increased the training data to 1300 images for a third model.

```

cat("total training cat images:", length(list.files(train_cats_dir2)), "\n")
## total training cat images: 1300

cat("total training dog images:", length(list.files(train_dogs_dir2)), "\n")
## total training dog images: 1300

cat("total validation cat images:", length(list.files(validation_cats_dir2)),
"\n")
## total validation cat images: 500

```

```

cat("total validation dog images:", length(list.files(validation_dogs_dir2)),
"\n")

## total validation dog images: 500

cat("total test cat images:", length(list.files(test_cats_dir2)), "\n")

## total test cat images: 500

  cat("total test dog images:", length(list.files(test_dogs_dir2)), "\n")

## total test dog images: 500

model3 <- keras_model_sequential() %>%
  layer_conv_2d(filters = 32, kernel_size = c(3, 3), activation = "relu",
               input_shape = c(150, 150, 3)) %>%
  layer_max_pooling_2d(pool_size = c(2, 2)) %>%
  layer_conv_2d(filters = 64, kernel_size = c(3, 3), activation = "relu") %>%
  layer_max_pooling_2d(pool_size = c(2, 2)) %>%
  layer_conv_2d(filters = 128, kernel_size = c(3, 3), activation = "relu")
%>%
  layer_max_pooling_2d(pool_size = c(2, 2)) %>%
  layer_conv_2d(filters = 128, kernel_size = c(3, 3), activation = "relu")
%>%
  layer_max_pooling_2d(pool_size = c(2, 2)) %>%
  layer_flatten() %>%
  layer_dense(units = 512, activation = "relu") %>%
  layer_dense(units = 1, activation = "sigmoid")

model3 %>% compile(
  loss = "binary_crossentropy",
  optimizer = optimizer_rmsprop(learning_rate = 1e-4),
  metrics = c("acc")
)

train_datagen2 <- image_data_generator(rescale = 1/255)
validation_datagen2 <- image_data_generator(rescale = 1/255)

train_generator2 <- flow_images_from_directory(
  train_dir2,
  train_datagen2,
  target_size = c(150, 150),
  batch_size = 20,
  class_mode = "binary"
)

validation_generator2 <- flow_images_from_directory(
  validation_dir2,
  validation_datagen2,
  target_size = c(150, 150),
  batch_size = 20,

```


[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible][illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

Increased the training data to 1800 images for a fourth version of the model.

```
cat("total training cat images:", length(list.files(train_cats_dir3)), "\n")
## total training cat images: 1800

cat("total training dog images:", length(list.files(train_dogs_dir3)), "\n")
## total training dog images: 1800

cat("total validation cat images:", length(list.files(validation_cats_dir3)),
"\n")
## total validation cat images: 500

cat("total validation dog images:", length(list.files(validation_dogs_dir3)),
"\n")
## total validation dog images: 500

cat("total test cat images:", length(list.files(test_cats_dir3)), "\n")
## total test cat images: 500

  cat("total test dog images:", length(list.files(test_dogs_dir3)), "\n")
## total test dog images: 500

model4 <- keras_model_sequential() %>%
  layer_conv_2d(filters = 32, kernel_size = c(3, 3), activation = "relu",
    input_shape = c(150, 150, 3)) %>%
  layer_max_pooling_2d(pool_size = c(2, 2)) %>%
  layer_conv_2d(filters = 64, kernel_size = c(3, 3), activation = "relu") %>%
  layer_max_pooling_2d(pool_size = c(2, 2)) %>%
  layer_conv_2d(filters = 128, kernel_size = c(3, 3), activation = "relu")
%>%
  layer_max_pooling_2d(pool_size = c(2, 2)) %>%
  layer_conv_2d(filters = 128, kernel_size = c(3, 3), activation = "relu")
%>%
  layer_max_pooling_2d(pool_size = c(2, 2)) %>%
  layer_flatten() %>%
  layer_dense(units = 512, activation = "relu") %>%
  layer_dense(units = 1, activation = "sigmoid")

model4 %>% compile(
  loss = "binary_crossentropy",
  optimizer = optimizer_rmsprop(learning_rate = 1e-4),
  metrics = c("acc")
)

train_datagen3 <- image_data_generator(rescale = 1/255)
validation_datagen3 <- image_data_generator(rescale = 1/255)
```

```

train_generator3 <- flow_images_from_directory(
  train_dir3,
  train_datagen3,
  target_size = c(150, 150),
  batch_size = 20,
  class_mode = "binary"
)

validation_generator3 <- flow_images_from_directory(
  validation_dir3,
  validation_datagen3,
  target_size = c(150, 150),
  batch_size = 20,
  class_mode = "binary"
)

history4 <- model4 %>% fit(
  train_generator3,
  steps_per_epoch = 100,
  epochs = 30,
  validation_data = validation_generator3,
  validation_steps = 50
)

```

This model with 1800 training images had the best performance overall.

Using a Pretrained model:

```

library(keras)

conv_base <- application_vgg16(
  weights = "imagenet",
  include_top = FALSE,
  input_shape = c(150, 150, 3)
)

summary(conv_base)

## Model: "vgg16"
##

```

## Layer (type)	Output Shape	Param
##		
##		
====		
## input_1 (InputLayer)	[(None, 150, 150, 3)]	0
## 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
##		
## block5_conv2 (Conv2D)	(None, 9, 9, 512)	2359808
##		
## block5_conv3 (Conv2D)	(None, 9, 9, 512)	2359808
##		
## block5_pool (MaxPooling2D)	(None, 4, 4, 512)	0
##		
##		
=====		
===		
## Total params: 14,714,688		
## Trainable params: 14,714,688		
## Non-trainable params: 0		
##		

```

base_dir4 <- "~/Documents/R/cats.vs.dogs4"
train_dir4 <- file.path(base_dir, "train")
validation_dir4 <- file.path(base_dir, "validation")
test_dir4 <- file.path(base_dir, "test")

datagen <- image_data_generator(rescale = 1/255)
batch_size <- 20

extract_features <- function(directory, sample_count) {

  features <- array(0, dim = c(sample_count, 4, 4, 512))
  labels <- array(0, dim = c(sample_count))

  generator <- flow_images_from_directory(
    directory = directory,
    generator = datagen,
    target_size = c(150, 150),
    batch_size = batch_size,
    class_mode = "binary"
  )

  i <- 0
  while(TRUE) {
    batch <- generator_next(generator)
    inputs_batch <- batch[[1]]
    labels_batch <- batch[[2]]
    features_batch <- conv_base %>% predict(inputs_batch)

    index_range <- ((i * batch_size)+1):((i + 1) * batch_size)
    features[index_range,,] <- features_batch
    labels[index_range] <- labels_batch

    i <- i + 1
    if (i * batch_size >= sample_count)
      break
  }

  list(
    features = features,
    labels = labels
  )
}

train <- extract_features(train_dir4, 2000)
validation <- extract_features(validation_dir4, 1000)
test <- extract_features(test_dir4, 1000)

reshape_features <- function(features) {
  array_reshape(features, dim = c(nrow(features), 4 * 4 * 512))
}

```

```

}
train$features <- reshape_features(train$features)
validation$features <- reshape_features(validation$features)
test$features <- reshape_features(test$features)

model5 <- keras_model_sequential() %>%
  layer_dense(units = 256, activation = "relu",
              input_shape = 4 * 4 * 512) %>%
  layer_dropout(rate = 0.5) %>%
  layer_dense(units = 1, activation = "sigmoid")

model5 %>% compile(
  optimizer = optimizer_rmsprop(lr = 2e-5),
  loss = "binary_crossentropy",
  metrics = c("accuracy")
)

## Warning in backcompat_fix_rename_lr_to_learning_rate(...): the `lr`
argument has
## been renamed to `learning_rate`.

history5 <- model5 %>% fit(
  train$features, train$labels,
  epochs = 30,
  batch_size = 20,
  validation_data = list(validation$features, validation$labels)
)

```

The pretrained model performed well, but is still showing signs of overfitting.

Used the pretrained model with 1300 training images:

```

base_dir6 <- "~/Documents/R/cats.vs.dogs6"

datagen <- image_data_generator(rescale = 1/255)
batch_size <- 20

extract_features <- function(directory, sample_count) {

  features <- array(0, dim = c(sample_count, 4, 4, 512))
  labels <- array(0, dim = c(sample_count))

  generator <- flow_images_from_directory(
    directory = directory,
    generator = datagen,
    target_size = c(150, 150),
    batch_size = batch_size,
    class_mode = "binary"
  )

  i <- 0

```

```

while(TRUE) {
  batch <- generator_next(generator)
  inputs_batch <- batch[[1]]
  labels_batch <- batch[[2]]
  features_batch <- conv_base %>% predict(inputs_batch)

  index_range <- ((i * batch_size)+1):((i + 1) * batch_size)
  features[index_range,,] <- features_batch
  labels[index_range] <- labels_batch

  i <- i + 1
  if (i * batch_size >= sample_count)
    break
}

list(
  features = features,
  labels = labels
)
}

train <- extract_features(train_dir2, 2600)
validation <- extract_features(validation_dir2, 1000)
test <- extract_features(test_dir2, 1000)

reshape_features <- function(features) {
  array_reshape(features, dim = c(nrow(features), 4 * 4 * 512))
}
train$features <- reshape_features(train$features)
validation$features <- reshape_features(validation$features)
test$features <- reshape_features(test$features)

model7 <- keras_model_sequential() %>%
  layer_dense(units = 256, activation = "relu",
              input_shape = 4 * 4 * 512) %>%
  layer_dropout(rate = 0.5) %>%
  layer_dense(units = 1, activation = "sigmoid")

model7 %>% compile(
  optimizer = optimizer_rmsprop(lr = 2e-5),
  loss = "binary_crossentropy",
  metrics = c("accuracy")
)

## Warning in backcompat_fix_rename_lr_to_learning_rate(...): the `lr`
argument has
## been renamed to `learning_rate`.

history7 <- model7 %>% fit(
  train$features, train$labels,

```

```

epochs = 30,
batch_size = 20,
validation_data = list(validation$features, validation$labels)
)

```

Used the pretrained model with 1800 images:

```

```r
base_dir5 <- "~/Documents/R/cats.vs.dogs5"

datagen <- image_data_generator(rescale = 1/255)
batch_size <- 20

extract_features <- function(directory, sample_count) {

 features <- array(0, dim = c(sample_count, 4, 4, 512))
 labels <- array(0, dim = c(sample_count))

 generator <- flow_images_from_directory(
 directory = directory,
 generator = datagen,
 target_size = c(150, 150),
 batch_size = batch_size,
 class_mode = "binary"
)

 i <- 0
 while(TRUE) {
 batch <- generator_next(generator)
 inputs_batch <- batch[[1]]
 labels_batch <- batch[[2]]
 features_batch <- conv_base %>% predict(inputs_batch)

 index_range <- ((i * batch_size)+1):((i + 1) * batch_size)
 features[index_range,,] <- features_batch
 labels[index_range] <- labels_batch

 i <- i + 1
 if (i * batch_size >= sample_count)
 break
 }

 list(
 features = features,
 labels = labels
)
}

```



```

train <- extract_features(train_dir3, 3600)
validation <- extract_features(validation_dir3, 1000)
test <- extract_features(test_dir3, 1000)

reshape_features <- function(features) {
 array_reshape(features, dim = c(nrow(features), 4 * 4 * 512))
}
train$features <- reshape_features(train$features)
validation$features <- reshape_features(validation$features)
test$features <- reshape_features(test$features)

model6 <- keras_model_sequential() %>%
 layer_dense(units = 256, activation = "relu",
 input_shape = 4 * 4 * 512) %>%
 layer_dropout(rate = 0.5) %>%
 layer_dense(units = 1, activation = "sigmoid")

model6 %>% compile(
 optimizer = optimizer_rmsprop(lr = 2e-5),
 loss = "binary_crossentropy",
 metrics = c("accuracy")
)

Warning in backcompat_fix_rename_lr_to_learning_rate(...): the `lr`
argument has
been renamed to `learning_rate`.

history6 <- model6 %>% fit(
 train$features, train$labels,
 epochs = 30,
 batch_size = 20,
 validation_data = list(validation$features, validation$labels)
)

```

This model resulted in the best performance overall.