data-overview

December 9, 2021

1 Data overview

On présentera ici les principales caractèristiques du jeu de données CIFAR

[7]:
```
import tensorflow as tf
import tensorflow.keras
import numpy as np
from matplotlib import pyplot as plt

import random

# Attribution d'une graine.
random.seed(564654)
```

[9]:
```
# Chargement du jeu de données
(x_train, y_train), (x_test, y_test) = tf.keras.datasets.cifar10.load_data()
assert x_train.shape == (50000, 32, 32, 3)
assert x_test.shape == (10000, 32, 32, 3)
assert y_train.shape == (50000, 1)
assert y_test.shape == (10000, 1)
```

```
A local file was found, but it seems to be incomplete or outdated because the
auto file hash does not match the original value of
6d958be074577803d12ecdefd02955f39262c83c16fe9348329d7fe0b5c001ce so we will re-
download the data.
Downloading data from https://www.cs.toronto.edu/~kriz/cifar-10-python.tar.gz
170500096/170498071 [==============================] - 68s 0us/step
170508288/170498071 [==============================] - 68s 0us/step
```

[10]:
```
label_names = {0: "airplane",
    1: "automobile",
    2: "bird",
    3: "cat",
    4: "deer",
    5: "dog",
    6: "frog",
    7: "horse",
    8: "ship",
```

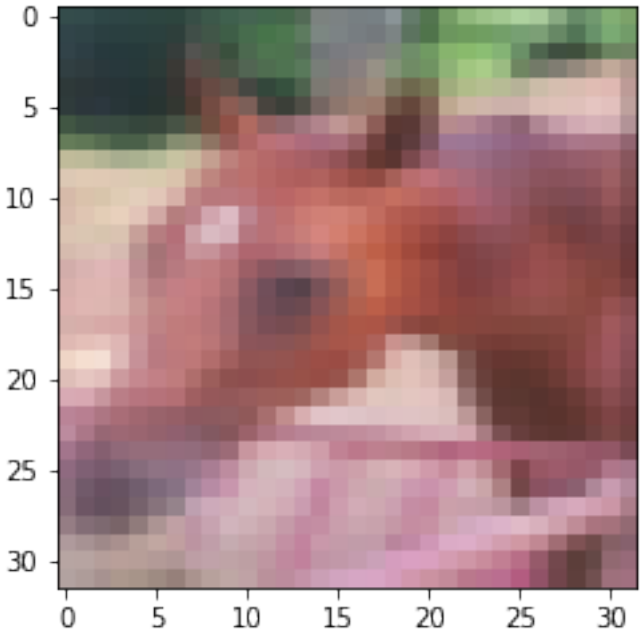
```
    9: "truck"}
```

Définire une fonction qui prend une image du dataset et l'affiche

[11]:
```
def display_picture(picture):
    plt.imshow(picture)
    plt.show()
```

Affichage d'une image du dataset quelconque

[17]:
```
i = random.randint(0,len(x_train))
display_picture(x_train[i])
print(label_names[y_train[i][0]])
```




```
horse
```

1.1 Liste des categories

[18]:
```
for i, label in label_names.items():
    print(i, label)
```

```
0 airplane
1 automobile
2 bird
3 cat
4 deer
```

```
5 dog
6 frog
7 horse
8 ship
9 truck
```