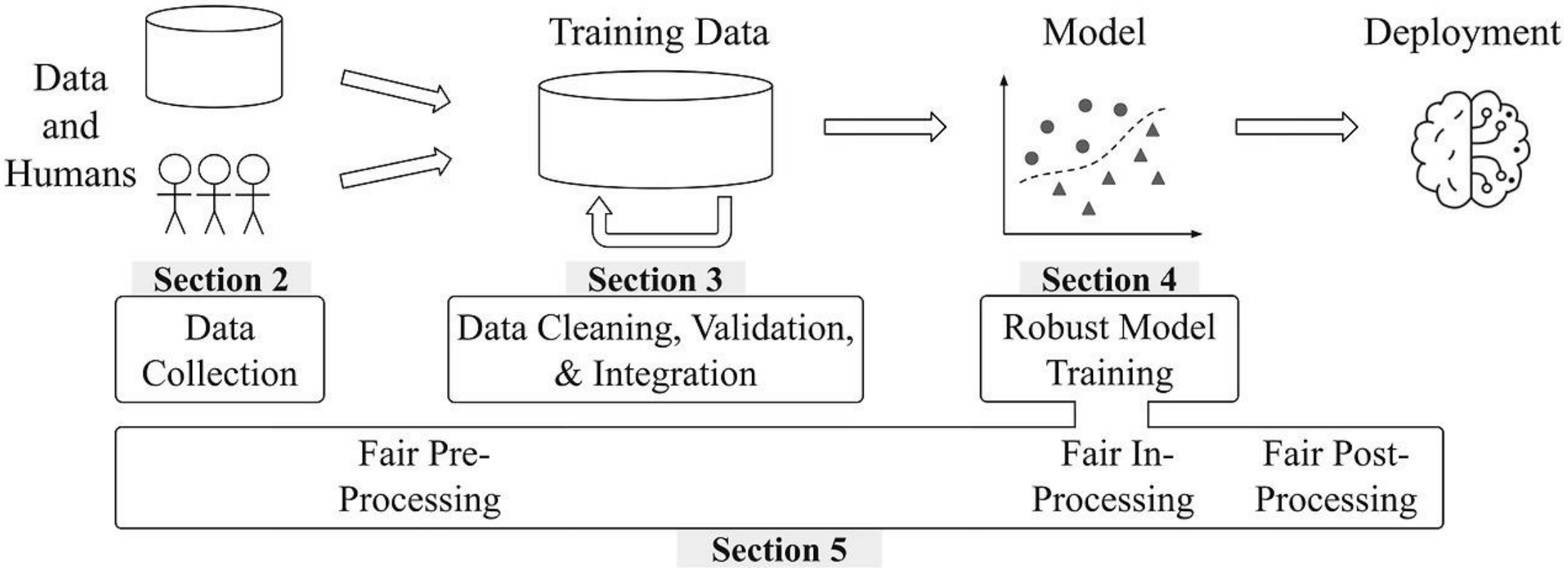
AI 실전 1주차

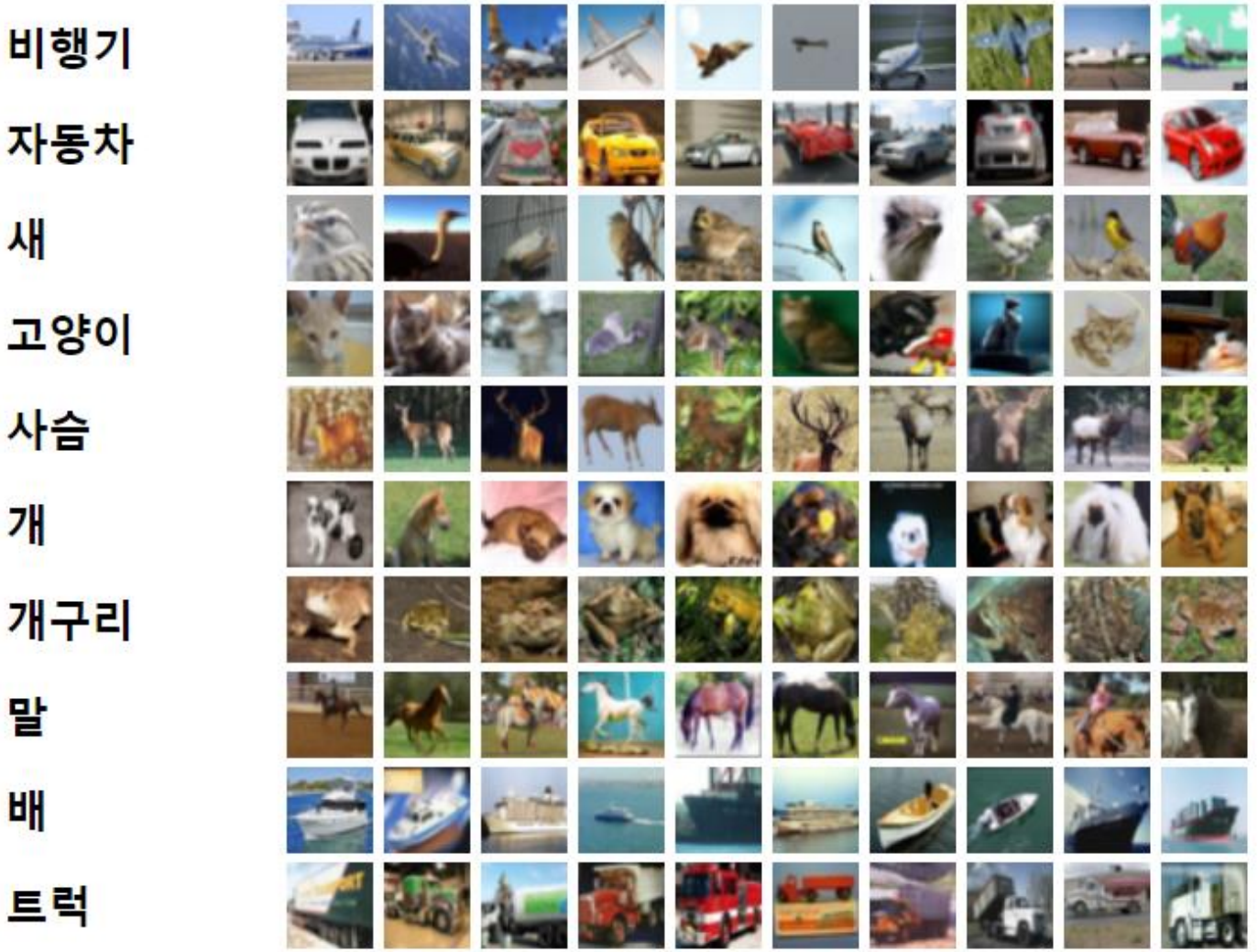
나만의 모델 설계 과정 및 진행

Feat. tensorflow

모델 제작 과정


Data
and
Humans
Training Data
Model
Deployment
Section 2
Section 3
Section 4
Data
Collection
Data Cleaning, Validation,
& Integration
Robust Model
Training
Fair Pre-
Processing
Fair In-
Processing
Fair Post-
Processing
Section 5

CIFAR-10 DataSet



필요한 라이브러리 불러오기

```
import tensorflow as tf

from tensorflow.keras import datasets, layers, models
import matplotlib.pyplot as plt
```

1. Load DataSet

```
(train_images, train_labels), (test_images, test_labels) = datasets.cifar10.load_data()
```

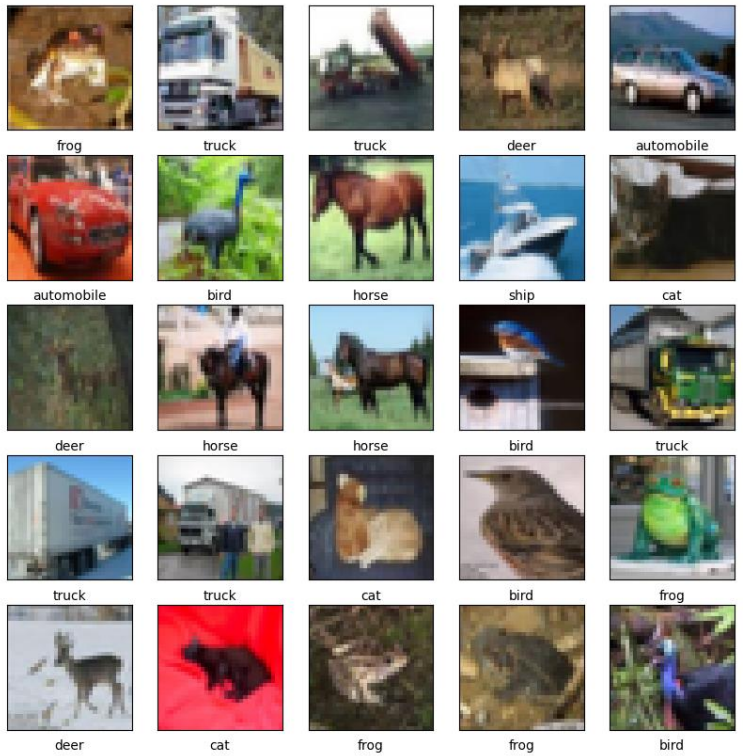
2. data preprocessing

```
train_images, test_images = train_images / 255.0, test_images / 255.0
```

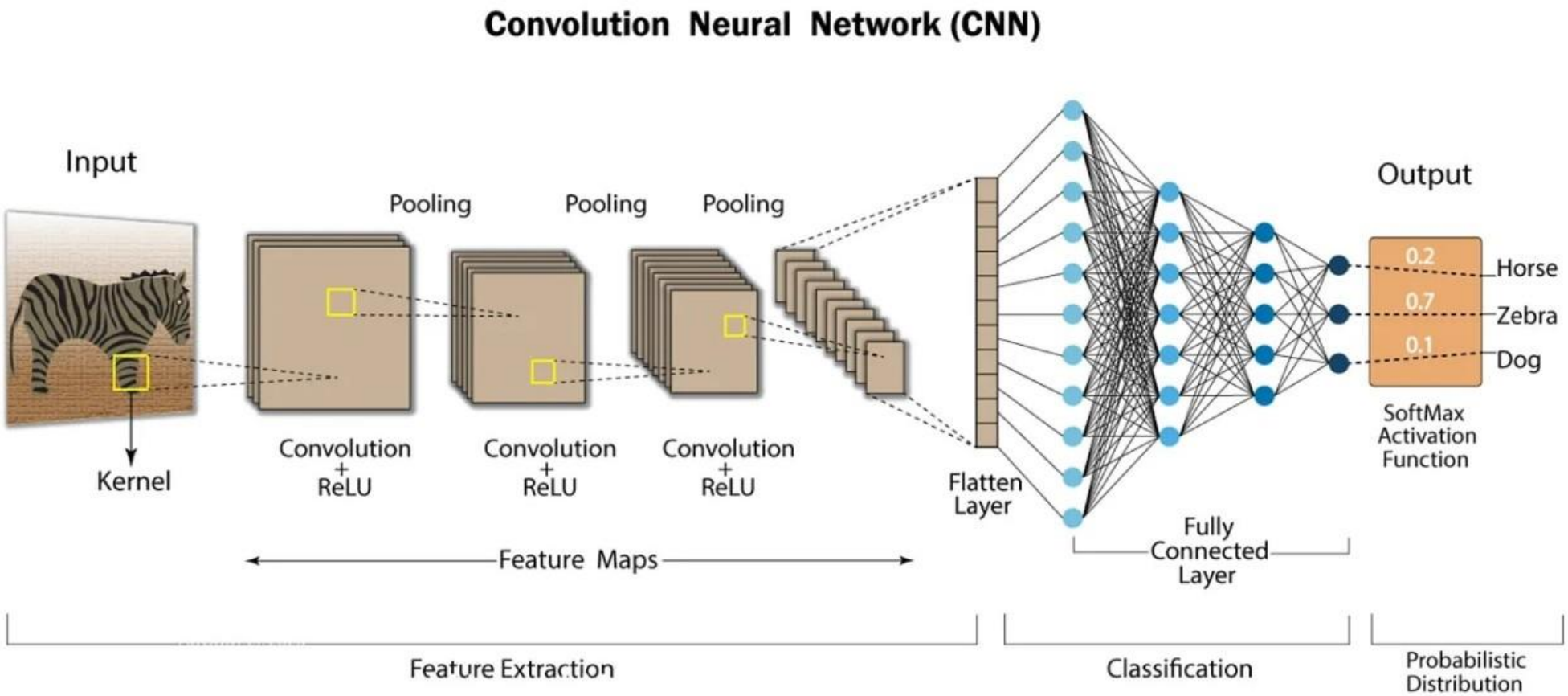
3. Check dataset

```
class_names = ['airplane', 'automobile', 'bird', 'cat', 'deer',
               'dog', 'frog', 'horse', 'ship', 'truck']

plt.figure(figsize=(10,10))
for i in range(25):
    plt.subplot(5,5,i+1)
    plt.xticks([])
    plt.yticks([])
    plt.grid(False)
    plt.imshow(train_images[i])
    # The CIFAR labels happen to be arrays,
    # which is why you need the extra index
    plt.xlabel(class_names[train_labels[i][0]])
plt.show()
```



CNN model struct



4. Model create (특성 추출)

```
model = models.Sequential()
model.add(layers.Conv2D(32, (3, 3), activation='relu', input_shape=(32, 32, 3)))
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'))
```

4.1 Model create (클래스 분류)

```
model.add(layers.Flatten())
model.add(layers.Dense(64, activation='relu'))
model.add(layers.Dense(10))
```

5. Model optimizer settings and train

```
model.compile(optimizer='adam',
              loss=tf.keras.losses.SparseCategoricalCrossentropy(from_logits=True),
              metrics=['accuracy'])

history = model.fit(train_images, train_labels, epochs=10,
                    validation_data=(test_images, test_labels))
```

과제 실습 100가지 종류 스포츠 이미지 분류하기

https://www.kaggle.com/datasets/gpiosenka/sports-classification/data

Dataset download code

```python
# Install dependencies as needed:
# pip install kagglehub[pandas-datasets]
import kagglehub
from kagglehub import KaggleDatasetAdapter

# Set the path to the file you'd like to load
file_path = ""

# Load the latest version
df = kagglehub.load_dataset(
  KaggleDatasetAdapter.PANDAS,
  "gpiosenka/sports-classification",
  file_path,
  # Provide any additional arguments like
  # sql_query or pandas_kwargs. See the
  # documenation for more information:
  # https://github.com/Kaggle/kagglehub/blob/main/README.md#kaggledatasetadapterpandas
)

print("First 5 records:", df.head())
```