数据的预处理

数据往往需要预先的处理，才能更好的用于神经网络的训练。

One-Hot Encoding

正如上述所说，如果是定性的数据，往往

In [ ]:
```
d
```

In [ ]:
```

```

In [ ]:
```

```

In [9]:
```
import pickle
import os

datasets = os.listdir('./others/datasets/')
datasets = [f for f in datasets if (f.startswith('Dataset') and f.endswith('.p'))]
datasets.sort()

datapath = os.path.join(f'./others/datasets/{datasets[20]}')
with open(datapath, 'rb') as f:
    data = pickle.load(f)
X_train    = data['X_train']
y_train    = data['y_train']
X_valid    = data['X_valid']
y_valid    = data['y_valid']
X_test     = data['X_test']
y_test     = data['y_test']
data_name  = data['name']

N_class    = data['n_class']
N_feature  = data['n_feature']
N_train    = X_train.shape[0]
N_valid    = X_valid.shape[0]
N_test     = X_test.shape[0]

print(f'Dataset "{data_name}" has {N_feature} input features and {N_class} classes.\nThere are {N_train} training example
```

```
Dataset "Pendigits" has 16 input features and 10 classes.
There are 6595 training examples, 2198 valid examples, and 2199 test examples in the dataset.
```

In [ ]:
```

```