# Practice Task: One-Hot Encoding

In one-hot encoding, each item in a sequence (e.g. `list`) of discrete items is **replaced by a vector where exactly one value is `1` and all other values are `0`**.

If an item occurs **multiple times**, it should be encoded to **the same one-hot vector** every time. **Different items** should be encoded to **different one-hot vectors**.

In other words from Wikipedia:

```
In natural language processing, a one-hot vector is a 1 × N matrix (vector) used to distinguish each word i
n a vocabulary from every other word in the vocabulary. The vector consists of 0s in all cells with the exc
eption of a single 1 in a cell used uniquely to identify the word. One-hot encoding ensures that machine le
arning does not assume that higher numbers are more important. For example, the value '8' is bigger than th
e value '1', but that does not make '8' more important than '1'. The same is true for words: the value 'lau
ghter' is not more important than 'laugh'.
```

One-hot encoding also has **applications beyond natural language processing**, e.g. when encoding categorical data in machine learning or statistics or when designing digital circuits.

```python
fruits = ["banana", "apple", "orange", "orange", "banana"]
one_hot_encoded_fruits = one_hot_encode(fruits)

print("one_hot_encoded_fruits", one_hot_encoded_fruits.shape, ":\n\n", one_hot_encoded_fruits, "\n")

fibs = [0, 1, 1, 2, 3, 5]
one_hot_encoded_fibs = one_hot_encode(fibs)

print("one_hot_encoded_fibs", one_hot_encoded_fibs.shape, ":\n\n", one_hot_encoded_fibs, "\n")
```

```
one_hot_encoded_fruits (5, 3) :

 [[0 0 1]
 [0 1 0]
 [1 0 0]
 [1 0 0]
 [0 0 1]]

one_hot_encoded_fibs (6, 5) :

 [[1 0 0 0 0]
 [0 1 0 0 0]
 [0 1 0 0 0]
 [0 0 1 0 0]
 [0 0 0 1 0]
 [0 0 0 0 1]]
```

```python
import numpy as np

def one_hot_encode(sequence):
    """ Takes an iterable sequence with arbitrary items as input and returns a one-hot encoded np.ndarray. """

    pass
```