Want to learn machine learning? Use my **machine learning flashcards**.

One-Hot Encode Features With Multiple Labels

20 Dec 2017

Preliminaries

```python
# Load libraries
from sklearn.preprocessing import MultiLabelBinarizer
import numpy as np
```

Create Data

```python
# Create NumPy array
y = [('Texas', 'Florida'),
     ('California', 'Alabama'),
     ('Texas', 'Florida'),
     ('Delware', 'Florida'),
     ('Texas', 'Alabama')]
```

One-hot Encode Data

```python
# Create MultiLabelBinarizer object
one_hot = MultiLabelBinarizer()

# One-hot encode data
one_hot.fit_transform(y)
```

```python
array([[0, 0, 0, 1, 1],
       [1, 1, 0, 0, 0],
       [0, 0, 0, 1, 1],
       [0, 0, 1, 1, 0],
       [1, 0, 0, 0, 1]])
```

CHRIS ALBON