one_hot_encoding

July 12, 2022

```python
[38]: import pandas as pd

      df = pd.read_csv(' 1500.csv')
```

```python
[39]: df.head()
```

```
[39]:    Unnamed: 0                         /                    \
      0         0   20 3870    998cc  16,465Km               5
      1         1   53 5162    998cc  42,082Km               5
      2         2   67 0998  3,342cc  62,265Km          4    5
      3         3  106 5607  1,995cc  38,233Km   SUV        5
      4         4   63 4826  1,998cc  56,155Km   SUV        7


      0  2018   9,800,000             (JA)
      1  2015   6,500,000
      2  2016  26,700,000      DHG330   AWD
      3  2019  25,800,000    TL 2.02WD
      4  2017  24,200,000       2.0 2WD
```

```python
[40]: data = df[[' ',' ']]
      data.head()
```

```
[40]:
      0        9,800,000
      1        6,500,000
      2       26,700,000
      3  SUV  25,800,000
      4  SUV  24,200,000
```

```python
[41]: temp=data[' '].str.replace(' ',',')
      temp=temp.str.replace(',','')
      temp=temp.astype(int)
```

```python
[42]: temp2 =pd.get_dummies(data[' '],
                            prefix=None,#
                            prefix_sep='_',
```

```
                             dummy_na=False,
                             columns=None,
                             sparse=False,
                             drop_first=False,
                             dtype=None)
```

[43]: `temp2.head()`

[43]:
```
    RV  SUV
0   0    0   1    0    0    0    0    0    0    0
1   0    0   1    0    0    0    0    0    0    0
2   0    0   0    1    0    0    0    0    0    0
3   0    1   0    0    0    0    0    0    0    0
4   0    1   0    0    0    0    0    0    0    0
```

[44]: `temp2['price'] = temp`

[45]: `temp2.head()`

[45]:
```
    RV  SUV                                         price
0   0    0   1    0    0    0    0    0    0    0   9800000
1   0    0   1    0    0    0    0    0    0    0   6500000
2   0    0   0    1    0    0    0    0    0    0  26700000
3   0    1   0    0    0    0    0    0    0    0  25800000
4   0    1   0    0    0    0    0    0    0    0  24200000
```