

Кодирование признаков

Алгоритм кодирования в sklearn

1. Импорт выбранного кодировщика

```
[ ] from sklearn.preprocessing import ВЫБРАННЫЙ_КОДИРОВЩИК
```

2. Инициализация кодировщика

```
[ ] encoder = ВЫБРАННЫЙ_КОДИРОВЩИК()
```

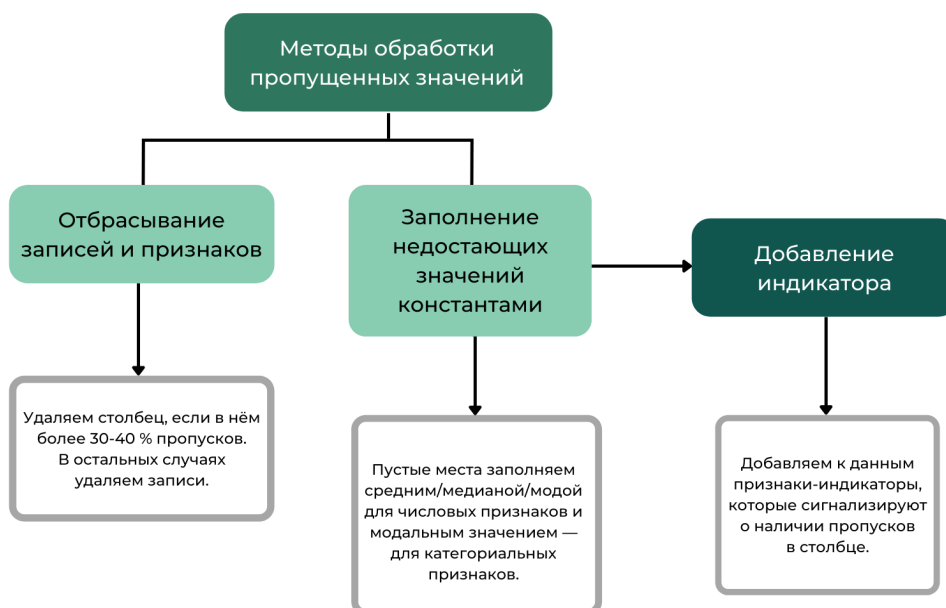
3. Обучение (подгонка) кодировщика на выбранном признаке

```
[ ] ВЫБРАННЫЙ_КОДИРОВЩИК.fit(data['ВЫБРАННЫЙ_ПРИЗНАК'])
```

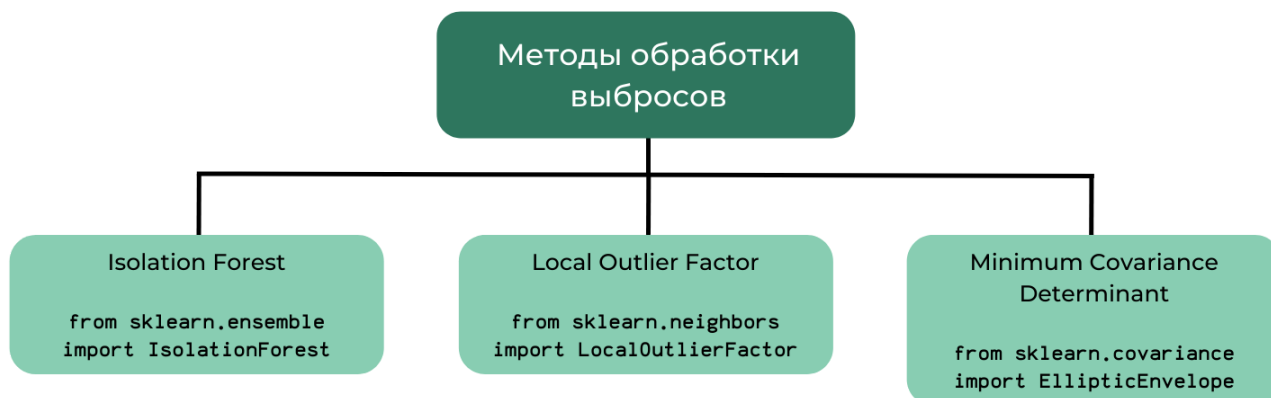
4. Преобразование выбранного признака

```
[ ] ВЫБРАННЫЙ_КОДИРОВЩИК.transform(data['ВЫБРАННЫЙ_ПРИЗНАК'])
```

Кодирование признаков



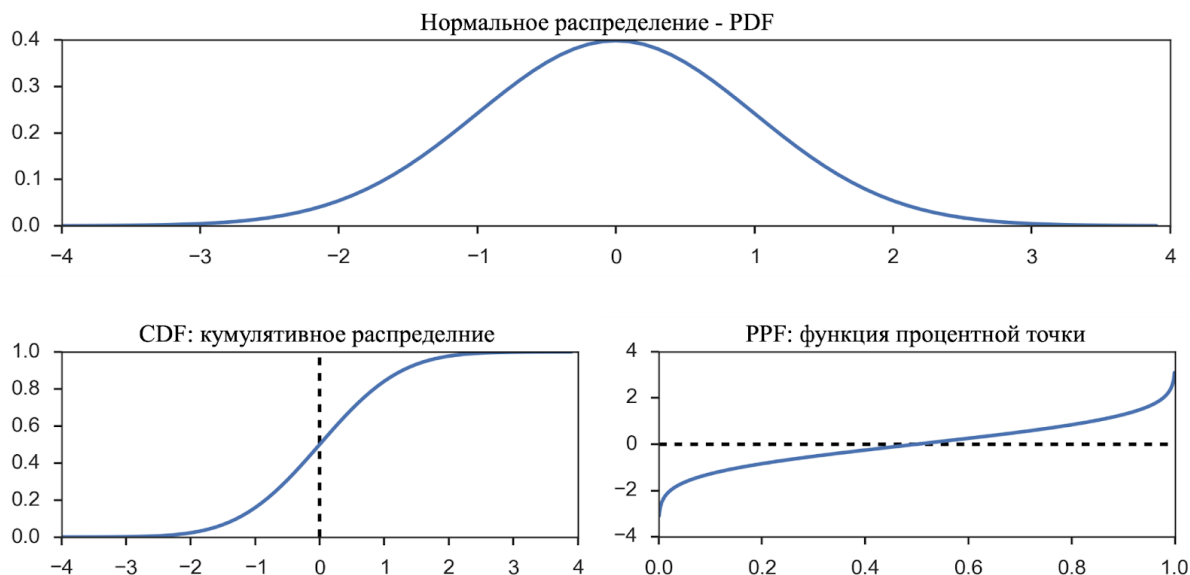
Обработка выбросов



Трансформация распределений признаков

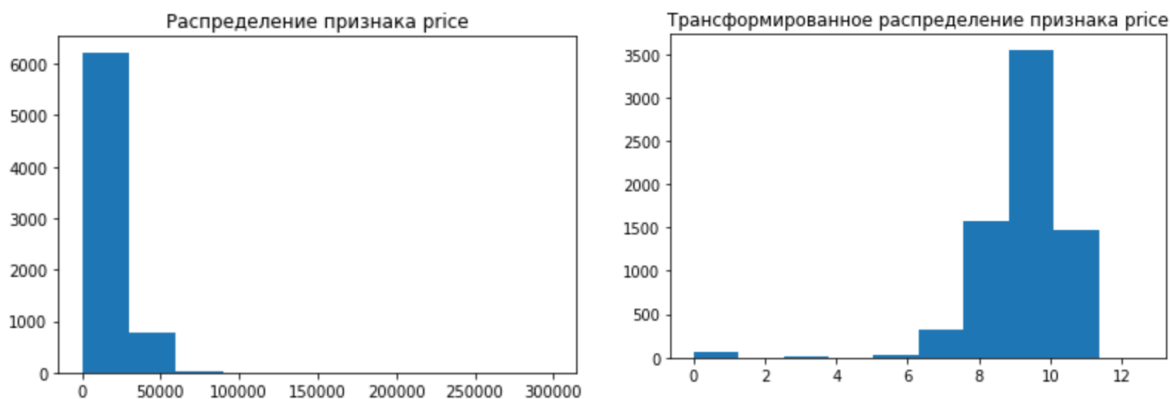
Quantile Transformer

Преобразование распределения переменной в нормальное или равномерное.



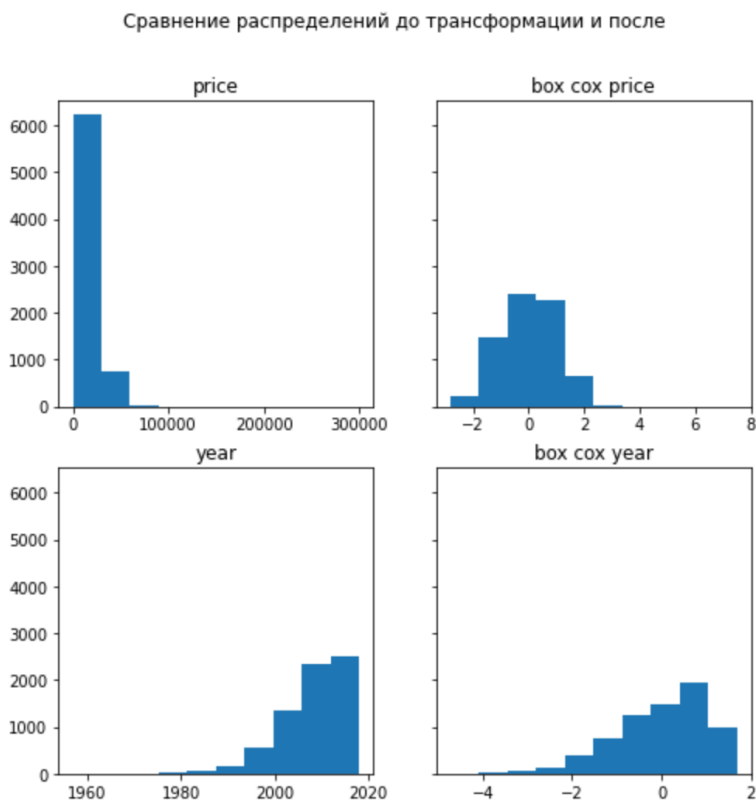
Log Transformer

Преобразование асимметричного распределения в нормальное/менее асимметричное.



Box-Cox Transformer

Преобразование распределения в нормальное.



Отбор признаков



Методы отбора признаков на основе фильтров

