

Рубежный контроль

Лю Ченхао

Группа: ИУ5И-22М

Вариант:17

Задача1:17. Задача2:37.

```
import numpy as np
import pandas as pd
from sklearn.datasets import load_iris
from sklearn.preprocessing import PowerTransformer
from sklearn.feature_selection import SelectPercentile, mutual_info_classif

# Загрузка набора данных
data = load_iris()
X = pd.DataFrame(data.data, columns=data.feature_names)
y = data.target

# Применение преобразования Yeo-Johnson к одному числовому признаку (первый признак)
pt = PowerTransformer(method='yeo-johnson')
X_transformed = X.copy()
X_transformed.iloc[:, 0] = pt.fit_transform(X.iloc[:, [0]])

# Вывод данных до и после преобразования
print("Исходные данные:\n", X.head())
print("Данные после преобразования Yeo-Johnson:\n", X_transformed.head())
```

Задача1:17

```
Исходные данные:
   sepal length (cm)  sepal width (cm)  petal length (cm)  petal width (cm)
0                5.1                3.5                1.4                0.2
1                4.9                3.0                1.4                0.2
2                4.7                3.2                1.3                0.2
3                4.6                3.1                1.5                0.2
4                5.0                3.6                1.4                0.2

Данные после преобразования Yeo-Johnson:
   sepal length (cm)  sepal width (cm)  petal length (cm)  petal width (cm)
0        -0.895690                3.5                1.4                0.2
1        -1.185173                3.0                1.4                0.2
2        -1.487921                3.2                1.3                0.2
3        -1.644609                3.1                1.5                0.2
4        -1.038838                3.6                1.4                0.2
```

Задача2:37.

```

# Вывод данных до и после преобразования
print("Исходные данные:\n", X.head())
print("Данные после преобразования Yeo-Johnson:\n", X_transformed.head())

# Использование класса SelectPercentile и метода на основе взаимной информации для отбора признаков
selector = SelectPercentile(mutual_info_classif, percentile=5)
X_selected = selector.fit_transform(X_transformed, y)

# Получение индексов выбранных признаков
selected_features = selector.get_support(indices=True)

# Вывод выбранных признаков
print("Индексы выбранных признаков:", selected_features)
print("Выбранные признаки:\n", X_transformed.iloc[:, selected_features].head())

```

Индексы выбранных признаков: [3]

Выбранные признаки:

	petal width (cm)
0	0.2
1	0.2
2	0.2
3	0.2
4	0.2

