

Звіт з лабораторної роботи №2

Кишинський Олександр

КНС 21

Завдання

Побудувати рекомендаційну систему, яка б містила 5 об'єктів та 5-6 параметрів рекомендацій. Розставити власноруч рейтинг рекомендацій.

1) Створюємо тестовий Dataset з певною інформацією:

Object	Param1	Param2	Param3	Param4	Param5
1	1	1	0	1	0
2	0	1	0	0	1
3	1	0	1	0	0
4	1	0	1	1	1
5	1	1	1	0	1

2) Створюємо пустий датафрейм з рекомендаціями та записуємо в нього дані з нашого Dataset'у. Вказуємо що колонка яка відповідає за об'єкт (віддільного користувача) буде виступати в якості індекса. Створюємо ексель файл.

```
1 import pandas as pd
2
3 recommendations = pd.DataFrame()
4
5 df = pd.read_excel('lab2.xlsx')
6 df = df.astype(str)
7
8 df = df.set_index("Object").T
9 df.to_excel('DataSet.xlsx')
10
11 df = pd.read_excel(io: 'DataSet.xlsx', index_col=0)
12 recList = list()
```

- 3) Робимо вибірку найбільш корельованої (взаємопов'язаної один з одною) змінної за допомогою методу `idxmax()`. Вибираємо відповідне значення кореляції (в нашому випадку приклад дуже простий – тому там є нуль), значення додається до списку `recList`, і пізніше ми видаляємо рядок з обробленою змінною. Цикл буде виконуватися до тих пір поки не буде опрацьована кожна змінна.

```
while k != 1:
    name = tempMatr.idxmax().item()
    value = tempMatr[0][tempMatr.idxmax().item()]
    recList.append([row, name, value])
    tempMatr = tempMatr.drop(labels=[tempMatr.idxmax().item()], axis=0)
    k += 1
```

- 4) Форматуємо дані і отримуємо вихідний файл з рекомендаціями:

```
recommendations = pd.concat(objs=[recommendations, pd.DataFrame(recList)], ignore_index=True)
recommendations.columns = ['Джерело', 'Рекомендація', 'Індекс кореляції']
recommendations = recommendations.sort_values(by='Індекс кореляції', ascending=False)
recommendations.to_excel(excel_writer='result.xlsx', index=False)
```

	A	B	C
1	Джерело	Рекомендація	Індекс кореляції
2	2	5	0.40824829
3	3	5	0.40824829
4	5	2	0.40824829
5	4	3	0.40824829
6	1	2	-0.16666667