

음식 분류 Documents

★ 사용법 [↗](#)

- ! 해당 코드를 실행하면 크롤링 된 음식점 데이터에서 음식 부분만 추출하여 해당 음식을 어느 카테고리를 분류하여 집어 넣을지 선택하는 코드

해당 코드를 실행하면 아래의 사진처럼 동작하는데 해당하는 음식에 맞는 카테고리 번호를 입력하면 해당 카테고리 텍스트 파일에 자동으로 저장됨.



★ 코드 설명 [↗](#)

```
1 import re
2 import os
3 # import pandas as pd
4 import json
5 import sys
6 import csv
7
8 # 크롤링된 음식점 데이터
9 file_path = r'C:\Users\com\Desktop\전국데이터\음식\KakaoMap(Scrape)_Restaurant_Seoul_20220304_51030.csv'
10
11 # def restart():
12 #     os.execl(sys.executable, sys.executable, *sys.argv)
13
14 # 총 13개의 카테고리
15 category_list = ['한식', '일식', '중식', '양식', '기타 외국식', '기타 음식', '기타', '제과점', '주점', '치킨', '카페&베
16                 '제외']
17 cate_food_list = {}
18 lines = []
19
20 # 크롤링된 데이터에서 음식(메뉴)부분만을 추출하기 위한 부분
21 for cate_idx, cate in enumerate(category_list):
22     if not cate in cate_food_list:
23         cate_food_list[cate] = []
24     category_file = r'C:\Users\com\Desktop\전국데이터\음식\음식카테고리\' + cate + '.txt'
25
26     food_f = open(category_file, 'r', encoding='utf-8')
27     f_rdr = csv.reader(food_f)
28     for f_idx, food in enumerate(f_rdr):
29         food = str(food)
30         food = food.replace('"', '').replace("'", '').strip()
31         if food in cate_food_list[cate]:
32             continue
33         else:
34             cate_food_list[cate].append(food)
```

```

35
36 # -----
37
38 f = open(file_path, 'r', encoding='utf-8')
39 rdr = csv.reader(f)
40 lines, remain_list = [], []
41
42 # 해당 음식점에 메뉴(음식)이 없을 경우
43 for main_idx, line in enumerate(rdr):
44     cate_to_food = []
45     if main_idx == 0:
46         line[0] = 'Unknownname'
47         lines.append(line)
48         continue
49     if line[1] == 'GoogleMap' or line[10] == '정보없음':
50         lines.append(line)
51         continue
52     try:
53         if line[10] != '{}':
54             search_food = eval(line[10]).keys()
55
56             for food_idx, food in enumerate(search_food):
57                 food = re.sub('[^ㄱ-ㅣ가-힣]+', '', food)
58                 food = food.strip()
59                 for cate in category_list:
60                     if food in cate_food_list[cate]:
61                         cate_to_food.append({food: cate})
62                         break
63             else:
64                 if food in remain_list:
65                     continue
66                 if food == '':
67                     continue
68             else:
69                 # 해당 음식과 카테고리 리스트를 표출하기 위함
70                 remain_list.append(food)
71                 print(food)
72                 print('한식 : 1\t 중식 : 2\t 일식 : 3\t 양식 : 4\t 기타 외국식(동남아) : 5\t 해물,생선 : 6\t '
73                       '패스트푸드 : 8\t 치킨 : 9\t 기타 음식(분식) : 10\t 카페 : 11\t 주점 : 12\t 기타(사찰) :
74
75                 food_dict = {'1': '한식', '2': '중식', '3': '일식', '4': '양식', '5': '기타 외국식', '6': '
76                             '8': '패스트푸드', '9': '치킨', '10': '기타 음식', '11': '카페, 베이커리', '12'
77                 s = input().strip()
78                 if s == '':
79                     s = food_dict['0']
80                     remain_list.append(food)
81                 else:
82                     s = food_dict[s]
83
84                 tmp = open(r'C:\Users\com\Desktop\전국데이터\음식\음식카테고리\\' + s + '.txt', 'a', encodi
85                 tmp.write('\n' + food)
86                 tmp.close()
87
88             # # restart()
89
90         else:
91             lines.append(line)
92             continue
93     except:

```

```
93         lines.append(line)
94         continue
95
96         line[10] = cate_to_food + remain_list
97
98         lines.append(line)
99
100 f = open('gimcheon_restaurant_processing.csv', 'w', encoding='utf-8', newline='')
101 wr = csv.writer(f)
102 wr.writerows(lines)
103
104 f.close()
```