

Python programming and practice

Initial Medical Diagnosis Program

Progress Report : 2

Date : 2023-11-26

Name : 조예지

ID : 222162

1. Introduction

1) Background

When someone feels unwell, they often think about going to the hospital, but sometimes it can be confusing to decide which specific department to visit. If you end up hearing advice like going from dermatology to orthopedics, it means you have to allocate time to visit another department. To address this issue, I have come up with the concept of an initial medical diagnosis program. If a person enters their symptoms, the program will provide a rough prediction of the associated illness and suggest which type of hospital or department they should visit. I believe that such a program would be more practical for users, helping them make informed decisions when seeking medical care.

2) Project goal

The goal is to analyze a CSV file containing past patient symptoms and the corresponding diseases linked to those symptoms. The objective is to predict a patient's symptoms based on the analysis and recommend a suitable hospital for the identified symptoms.

3) Differences from existing programs

Existing programs primarily focus on predicting the disease based on a patient's symptoms. Consequently, patients need to revisit the information to determine which hospital to visit after receiving the prediction results. This program, however, differs by having a feature that informs users 'which hospital to visit' right from the beginning, aiming to streamline the process and reduce the need for additional searches compared to traditional programs.

2. Functional Requirement

1) Function 1 (user_input_validator.py) (Validation of input values)

- It validates the input values.
- If a correct value is entered, it returns that value; otherwise, it prints a message instructing to enter the correct format and prompts for input again.

(1) Detailed function 1 (validate_yes_no_input())

- Function that only accepts Yes or No.
- User input is stored in lowercase, so it is case-insensitive.
- If the user enters Yes or No, the respective value is returned; otherwise, it prints a message instructing to enter the correct format and prompts for input again.

(2) Detailed function 2 (validate_integer_input())

- Function that only accepts integers.
- It uses the `isdigit()` method to check if the entered value is a number, and if so, it converts the value to an integer and returns it.
- If the input value is not a number, it prints a message instructing to enter the correct format and prompts for input again.

(3) Detailed function 3 (validate_gender_input())

- Function that only accepts gender (Male or Female).
- User input is stored in lowercase, so it is case-insensitive.
- If the user enters Male or Female, the respective value is returned; otherwise, it prints a message instructing to enter the correct format and prompts for input again.

2) Function 2 (medical_diagnosis.py)

(Disease diagnosis & Hospital recommendation)

- It loads a CSV file containing patient data, identifies the patient with the most similar symptoms based on the information provided by the user, checks the disease of that patient, and outputs it. Then, it recommends an appropriate hospital.

class MedicalDiagnosis:

(1) Detailed function 1 (__init__())

- It takes the CSV file path, loads the data, and initializes the user input and the list of matched patients.

(2) Detailed function 2 (load_data())

- It reads data from the given CSV file and returns it in the form of a dictionary. The header and data are stored in a dictionary with keys 'header' and 'data', respectively.

(3) Detailed function 3 (calculate_similarity())

- This method calculates the similarity between the given patient and user input.
- For each item, the method increments the score by 1 if the user input and patient data values match.
- The final accumulated score is returned.

(4) Detailed function 4 (find_top_disease())

- Iterates through each row in 'self.data'.
- Calculates the similarity between user input and each row using the 'calculate_similarity' method.
- Appends a tuple containing the disease name and similarity score to the 'similarities' list.
- Sorts the list of tuples based on similarity in descending order.
- Retrieves the top N diseases based on similarity and returns the result.

(5) Detailed function 5 (recommend_hospital())

- Recommends a hospital based on the extracted disease.

3. Progress

1) Feature implementation

(1) user_input_validator.py

validate_yes_no_input():

Input: Receive a parameter called message. This is a message to output to the user.

Output: The user receives a "Yes" or "No" response.

Action: Through an infinite loop, the user will continue to receive input until they enter "yes" or "no" correctly. User input is verified by converting to lowercase letters, and if it is not correct, it outputs an error message and receives input again.

validate_integer_input():

Input: Receive a parameter called message. This is a message to output to the user.

Output: Enter an integer from the user.

Action: Through an infinite loop, the user continues to receive input until an integer is entered. Use the isdigit() method to verify that the input consists of integers, and if it is not an integer, it outputs an error message and receives the input again.

validate_gender_input():

Input: Receive a parameter called message. This is a message to output to the user.

Output: The user receives a response with "Male" or "Female".

Action: Through an infinite loop, the input is continuously received until the user enters "mal" or "female" correctly. User input is verified by converting to lowercase letters, and if it is not correct, it outputs an error message and receives input again.

- What you have learned applied (e.g. repetitions, conditional statements, functions, classes, modules, etc.)

Repeat statement, conditional statement, function, module

- Code Screenshot

```

PY202309-P > Sources > user_input_validator.py
1  # Yes 또는 No로 응답을 입력받는 함수
2  def validate_yes_no_input(message):
3      while True:
4          user_input = input(f"{message}").lower()
5          if user_input in ['yes', 'no']:
6              return user_input
7          else:
8              print("올바른 형식으로 입력하세요. 'Yes' 또는 'No'로 입력하세요.")
9
10 # 정수 입력 받는 함수
11 def validate_integer_input(message):
12     while True:
13         user_input = input(f"{message}")
14         if user_input.isdigit():
15             return int(user_input)
16         else:
17             print("숫자 형식으로 입력하세요.")
18
19 # 성별 유효성 확인 함수
20 def validate_gender_input(message):
21     while True:
22         user_input = input(f"{message}").lower()
23         if user_input in ['male', 'female']:
24             return user_input
25         else:
26             print("올바른 형식으로 입력하세요. 'Male' 또는 'Female'로 입력하세요.")
27

```

(2) medical_diagnosis.py

__init__():

Input: 'user_input': A dictionary containing user-provided symptoms and their values.

'csv_filename': The filename of the CSV file containing medical data.

Output: Initializes the instance variables 'user_input' and 'csv_filename'.

Action: Initializes the object with the provided user input and CSV filename.

load_data():

Input: None explicitly (uses 'self.csv_filename').

Output: Populates the instance variable 'self.data' with the contents of the CSV file.

Action: Opens the specified CSV file and reads its contents using 'csv.DictReader'.
Converts the reader object into a list of dictionaries and stores it in 'self.data'.

calculate_similarity():

Input: 'row': A dictionary representing a row of medical data.

Output: Returns an integer representing the similarity between 'self.user_input' and the provided 'row'.

Action: Compares each key-value pair in 'self.user_input' with the corresponding values in the given 'row'. Increments the similarity count whenever a match is found.

find_top_diseases():

Input: 'top_n': An optional parameter specifying the number of top diseases to retrieve.

Output: Returns a list of tuples, each containing the name of a disease and its similarity score.

Action: Iterates through each row in 'self.data'.

Calculates the similarity between user input and each row using the 'calculate_similarity' method.

Appends a tuple containing the disease name and similarity score to the 'similarities' list.

Sorts the list of tuples based on similarity in descending order.

Retrieves the top N diseases based on similarity and returns the result.

- What you have learned applied (e.g. repetitions, conditional statements, functions, classes, modules, etc.)

repetitions, conditional statements, functions, classes, modules

- Code Screenshot

```

PY202309-P > Sources > medical_diagnosis.py > ...
1  import csv
2
3  class MedicalDiagnosis:
4      def __init__(self, user_input, csv_filename):
5          self.user_input = user_input
6          self.csv_filename = csv_filename
7
8      def load_data(self):
9          with open(self.csv_filename, 'r') as file:
10             reader = csv.DictReader(file)
11             self.data = list(reader)
12
13     def calculate_similarity(self, row):
14         # 일치하는 증상의 개수를 세어 유사도로 사용
15         similarity = sum(1 for key, value in self.user_input.items() if value == row[key])
16         return similarity
17
18     def find_top_diseases(self, top_n=3):
19         similarities = []
20
21         for row in self.data:
22             similarity = self.calculate_similarity(row)
23             disease_name = row['Disease']
24             similarities.append((disease_name, similarity))
25
26         # 유사도를 기준으로 내림차순 정렬
27         similarities.sort(key=lambda x: x[1], reverse=True)
28
29         # Top N 질병 추출
30         top_diseases = similarities[:top_n]
31
32         return top_diseases
33

```

2) Test Results

(1) main.py

- Verify that only valid values are received among the user's inputs
- Test Results Screenshot


```

올바른 형식으로 입력하세요. 'Yes' 또는 'No'로 입력하세요.
발열이 있나요? (Yes/No) :yes
기침이 있나요? (Yes/No) :no
피로감이 있나요? (Yes/No) :20
올바른 형식으로 입력하세요. 'Yes' 또는 'No'로 입력하세요.
피로감이 있나요? (Yes/No) :yes
호흡 곤란이 있나요? (Yes/No) :d
올바른 형식으로 입력하세요. 'Yes' 또는 'No'로 입력하세요.
호흡 곤란이 있나요? (Yes/No) :no
나이를 입력하세요. :20
성별을 입력하세요. (Male/Female) :dfk
올바른 형식으로 입력하세요. 'Male' 또는 'Female'로 입력하세요.
성별을 입력하세요. (Male/Female) :male
수축기 혈압을 입력하세요. :200
총 콜레스테롤 수치를 입력하세요. :k
숫자 형식으로 입력하세요.
총 콜레스테롤 수치를 입력하세요. :200

```

- Compare your input with the csv file to see if you can return a disease with similar symptoms
- Test Results Screenshot

```

C:\Users\82102\Desktop\PY202309-P>C:/Python311/python.exe c:/Users/82102/Desktop/PY202309-P/Sources
/main.py
발열이 있나요? (Yes/No): yes
기침이 있나요? (Yes/No): no
피로감이 있나요? (Yes/No): yes
호흡 곤란이 있나요? (Yes/No): yes
나이를 입력하세요: 19
성별을 입력하세요 (Male/Female): female
수축기 혈압을 입력하세요: 50
총 콜레스테롤 수치를 입력하세요: 180
Top-3 Diseases:
Influenza: Similarity - 7
Pneumonia: Similarity - 6
Influenza: Similarity - 5

```

4. Changes in Comparison to the Plan

1. calculate_similarity_score() -----> calculate_similarity()

Change each key-value pair in self.user_input to compare with its corresponding key in a given row

2. find_similar_disease() -----> find_top_diseases()

Change to find and return top N similar diseases

5. Schedule

- Progress Indication

Work		11/3	11/26	12/10	12/24
Create a proposal		----->			
user_input_validator.py	validate_yes_no_input() validate_integer_input() validate_gender_input()		----->	Modify to save first letter uppercase	
medical_diagnosis.py	init() load_data() calculate_similarity() find_top_diseases()			----->	
	recommend_hospital()				----->