

Lab#2, NLP@CGU Spring 2025

This is due on 2025/03/17 15:30, submit to Teams as a PDF (lab2.pdf) (File>Print>Save as PDF).

IMPORTANT: After copying this notebook to your Google Drive, please paste a link to it below. To get a publicly-accessible link, hit the *Share* button at the top right, then click "Get shareable link" and copy over the result. If you fail to do this, you will receive no credit for this lab!

LINK: *paste your link here*

<https://colab.research.google.com/drive/16pp9XaNoLURlew5rBy1XtfWeGRrkJrAW?usp=sharing>

1 開始使用 AI 編寫或生成程式碼。

Student ID: B1129028

Name: 陳奕劭

按兩下 (或按 Enter 鍵) 即可編輯

Question 1 (100 points)

Implementing Trie in Python.

Trie is a very useful data structure. It is commonly used to represent a dictionary for looking up words in a vocabulary.

For example, consider the task of implementing a search bar with auto-completion or query suggestion. When the user enters a query, the search bar will automatically suggests common queries starting with the characters input by the user.



按兩下 (或按 Enter 鍵) 即可編輯

```
1 import requests
2 from bs4 import BeautifulSoup
3
4 def get_page():
5     try:
6         resp = requests.get("https://www.cgu.edu.tw")
7     except:
8         resp = None
9
10    if resp and resp.status_code == 200:
11        print(resp.status_code)
12        print(resp.text)
13        soup = BeautifulSoup(resp.text, 'html.parser')
14        print(soup)
15        try:
16            h1 = soup.find('h1')
17        except:
18            h1 = None
19        if h1:
20            print(soup.find('h1'))
21            print(soup.find('h1').text)
22        try:
23            h2 = soup.find('h2')
```

```

24         except:
25             h2 = None
26         if h2:
27             print(soup.find('h2').text)
28         else:
29             print('h2 tag not found!')
30
31
32 get_page()
33
34 # YOUR CODE HERE!
35 # IMPLEMENTIG TRIE IN PYTHON
36
37 class TrieNode:
38
39     def __init__(self, char):
40         self.char = char
41         self.children = []
42         self.is_end = False
43         self.count = 0
44
45 class Trie(object):
46
47     def __init__(self):
48         self.root = TrieNode("")
49
50     def insert(self, word):
51         node = self.root
52         for char in word:
53             found = False
54             for child in node.children:
55                 if child.char == char:
56                     node = child
57                     found = True
58                     break
59             if not found:
60                 new_node = TrieNode(char)
61                 node.children.append(new_node)
62                 node = new_node
63         node.is_end = True
64         node.count += 1 # Increment word count at this node
65
66     def dfs(self, node, prefix):
67         result = []
68         if node.is_end:
69             result.append((prefix, node.count))
70         for child in node.children:
71             result.extend(self.dfs(child, prefix + child.char))
72         return result
73
74     def query(self, x):
75         node = self.root
76         for char in x:
77             found = False
78             for child in node.children:
79                 if child.char == char:
80                     node = child
81                     found = True
82                     break
83             if not found:
84                 return [] # No words match the prefix
85         return self.dfs(node, x) # Perform DFS from the last matched node
86
87 # # DO NOT MODIFY THE VARIABLES
88 obj = Trie()
89 obj.insert("長庚資工")
90 obj.insert("長大")
91 obj.insert("長庚")
92 obj.insert("長庚")
93 obj.insert("長庚大學")
94 obj.insert("長庚科技大學")
95
96 # # DO NOT MODIFY THE BELOW LINE!
97 # # THE RESULTS : [(words, count), (words, count)]
98 print(obj.query("長"))

```

```
99 # [('長庚', 2), ('長庚資工', 1), ('長庚大學', 1), ('長庚科技大學', 1), ('長大', 1)]
100
101 print(obj.query("長庚"))
102 # [('長庚', 2), ('長庚資工', 1), ('長庚大學', 1), ('長庚科技大學', 1)]

    }
    });

    //取得多頁籤的列表資料
    $.ajax({
        url: '/ch/Home/ContentBlockMultiTabDetail',
        type: "POST",
        data: { nodeId: id, blockId: blockId, blockContentType: blockContentType },
        success: function (data) {
            $("#myTabContent_All").removeClass("active");
            var tabContentId = "#myTabContent_" + blockId;
            $(tabContentId).addClass("active");
            $(tabContentId).html(data);
            //載入圖片
            $('img.lazyload').lazy();
            //CheckImg(); //預設圖
        }
    });
}

$('.RedirectPage').click(function () {
    var url = $(this).attr('data-action');

    if (url !== undefined)
        window.location.href = url;
});

Date.prototype.addDays = function (days) {
    this.setDate(this.getDate() + days);
    return this;
}
</script>
<script>
    setTimeout(function () {
        AOS.init({
            once: true,
            delay: 200,
            duration: 200
        })
    }, 500)
</script>
</link></body>
</html>

<h1 class="logo me-auto">
<a href="/">

</a>
</h1>
```

如果你是

```
[('長庚', 2), ('長庚資工', 1), ('長庚大學', 1), ('長庚科技大學', 1), ('長大', 1)]
[('長庚', 2), ('長庚資工', 1), ('長庚大學', 1), ('長庚科技大學', 1)]
```