实验一 需求获取

一、小组成员与得分分配

彭玉　　181860073　　33%

黄思明　181860030　　34%

赫龙飞　181860034　　33%

二、实验目的

对开源项目VSCode获取有效的信息，爬虫抽取其潜在需求，进行需求的分析与分类。

三、实验方法

数据获取方法：使用一个python爬虫程序抓取Visual Studio Code 的 Issue上一百条用户的问答

需求分析方法：将每个问答作为一个需求，直接对问答进行人工分类从而获得可能的需求

四、实验过程

4.1 确定项目

选定开源项目Visual Studio Code。
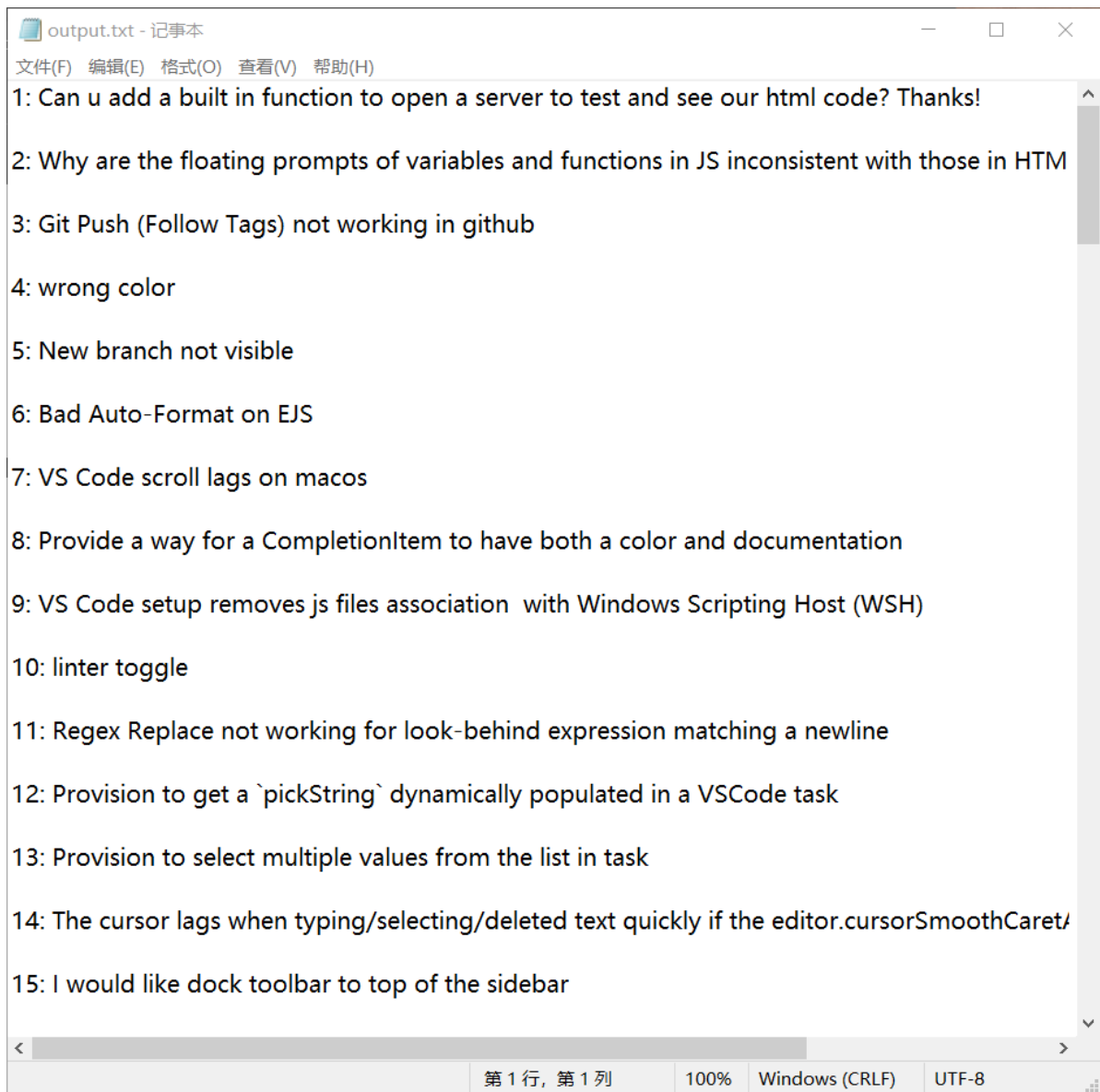
4.2 数据获取

爬虫抓取了Visual Studio Code 的 Issue上一百条用户的问答。

爬虫代码：

```python
def getHTMLText(self, page):
    try:
        r = requests.get(self.url+page, headers=self.header, timeout=30)
        print(r.status_code)
        r.raise_for_status()
        print(r.apparent_encoding)
        r.encoding = r.apparent_encoding
        return r.text
    except:
        traceback.print_exc()
        sys.exit("the url refused requests")

def dealHTMLText(self, text):
    soup = BeautifulSoup(text, 'lxml')
    for tr in soup.find_all('div', class_='flex-auto min-width-0 p-2 pr-3 pr-md-2'):
        content = str(tr.find('a', class_='link-gray-dark v-align-middle no-underline h4 js-navigation-open').string)
        print(content)
        self.resList.append(content)
```

数据文件：

1: Can u add a built in function to open a server to test and see our html code? Thanks!

2: Why are the floating prompts of variables and functions in JS inconsistent with those in HTM

3: Git Push (Follow Tags) not working in github

4: wrong color

5: New branch not visible

6: Bad Auto-Format on EJS

7: VS Code scroll lags on macos

8: Provide a way for a CompletionItem to have both a color and documentation

9: VS Code setup removes js files association with Windows Scripting Host (WSH)

10: linter toggle

11: Regex Replace not working for look-behind expression matching a newline

12: Provision to get a `pickString` dynamically populated in a VSCode task

13: Provision to select multiple values from the list in task

14: The cursor lags when typing/selecting/deleted text quickly if the editor.cursorSmoothCaretA

15: I would like dock toolbar to top of the sidebar

4.3 需求分析与分类

a) 外观需求

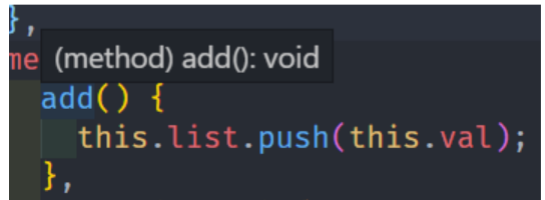2:Why are the floating prompts of variables and functions in JS inconsistent with those in HTML?

变量和函数的悬浮提示在JS和HTML中不一致。

in JS:

in HTML:

    在不同环境中悬浮提示不相同，会使界面整体美观性下降，也有可能在使用时带给用户歧义。此条需求是用户对于软件的美观性和在不同环境的一致性的需求。

此类需求还有以下issue：

4: wrong color

8: Provide a way for a CompletionItem to have both a color and documentation

10: linter toggle

15: I would like dock toolbar to top of the sidebar

18: Unwanted HTML tag highlight since recent updates

24: Readonly editor hover appears weird for first line

25: Semantic Highlighting can't be disabled

27: The extension manager should display to which marketplace it connects

29: Toggle visibility of Editor Groups

34: Getting artefacts on mouse move over window

35: Double scrollbars in diff views should feel like 1 thing

37: Icons have half opacity in Running Extensions

42: Windows explorer icons should be updated with the icons for the current VS Code file icon theme

47: Suggest detail does not shrink to fit content when resized

52: Import from other project to test.component.spec.ts is underlined with red (Angular)

55: Word wrap for variables in Watch view

63: Improve user experience when settings sync requires client update

65: Intellisense incorrectly highlighting auto variable in C++

66: Highlight the row selected from the "Find All Reference" list

67: if i save the file allignement keeps changing

73: Bash syntax highlighting breaks after in-line pipe to while-read loop

74: Custom menu: consider to align accounts menu like gear and home menu

75: Scrolling in a notebook diff editor is sluggish

89: Pinned Tabs: offer a mode where unpin icon is hidden

95: Sync debug and editor hover fade out times

b）使用功能需求

3:Git Push (Follow Tags) not working in github

Git Push命令在github中不生效。

- VSCode Version: 1.50.1
- OS Version: Ubuntu 18 Bionic Beaver

Steps to Reproduce:

1. Commit
2. Git Create Tag
3. Git Push (Follow Tags)

My code successful push to github, but no tags and releases file created in github.

Does this issue occur when all extensions are disabled?: Yes/No
YES (except GIT extension)

此条需求为用户对于功能的需求，虽然可能是由于用户自身出现了不当操作导致了该问题，但是作为知名开源项目，需要尽可能完善软件的种种功能以保证用户的使用。

26: Allow filter in quick navigation (Ctrl-Q)

Now: Ctrl-Q shows quick navigation and no fast jump to some item
Request: show quick navigation and allow typing letters to jump to an item

对于此条需求，用户根据自己的使用习惯提出了快捷键的需求，希望在原有功能的基础上添加完善自己的更多需求。

此类需求还有以下issue：

5: New branch not visible

7: VS Code scroll lags on macos

9: VS Code setup removes js files association with Windows Scripting Host (WSH)

11: Regex Replace not working for look-behind expression matching a newline

12: Provision to get a `pickString` dynamically populated in a VSCode task

13: Provision to select multiple values from the list in task

14: The cursor lags when typing/selecting/deleted text quickly if the editor.cursorSmoothCaretAnimation option is set to true

20: Should "Report Issue" work for built-in extensions in Running Extensions view?

21: OPEN_TO_SIDE_COMMAND_ID should not use fileservice to detect folders

22: Editors: revisit "preview" editors default behaviour

31: Multiple snippet transforms

33: settings sync gets turned off because of too many requests

38: Compound configuration with two C applications; sometimes doesn't launch both application

41: Visual indication of upload progress when drag&dropping into remote host

43: language server support for big files when largeFileOptimization is disabled

44: Running `code .` on WSL will try to download an outdated version of VSC

45: unable to login with corporate microsoft account

49: Should peek editor close on ESC before suggest?

50: Consider to remove migrateFromOldCachedViewContainersValue

51: Inconsistent Behavior Inside the Custom Open File Dialog of Remote Extensions when Navigating with a Screen Reader

60: Action icons that are submenus don't handle empty well

61: Bot should not lock author-verification-requested issues until after the iteration

62: Wsl won't allow me to open terminal for python

64: On the integrated terminal, `ctrl + click` doesn't correctly open the file links with file URI schemes (file:///).

68: Slow file handling on remote

69: TypeScript string unions autocompletion cut off because of redundant info

70: Please add support engine version to api

72: Inform users about incompatibility while opting in to sync Insiders with Stable

76: Clicking link in text block navigates to link AND opens it for edit

77: Remote SSH extension locks up low memory servers

78: Application is 'fuzzy' - nothing is crisp

80: Preserve text input between 'Git: checkout to' and Quick Open providers

84: VSCode is too slow on all actions related to SCM

86: System Sound on Windows dialog boxes lags for the first dialog window

87: Smoke test: changes 'workbench.action.toggleSidebarPosition' command key binding and verifies it

88: Can not load installed extension and recommend extension

92: Copy/Cut broken on safari on iPadOS

94: unbound breakpoint with preview debugger and runtimeArguments launch config

97: find and replace bad

100: Extensions - Improve notification dialog when installing an extension

c) 软件正确性需求

6: Bad Auto-Format on EJS

EJS上的自动格式错误

I was writing the following lines on a .ejs file:

```
<script>
    const mapToken = '<%-process.env.MAPBOX_TOKEN%>';
    const campground = <%-JSON.stringify(campground) %>;
</script>
```

The '<%-...' on the third line would become '<% -...' whenever I was saving, i.e. white-space is added between the '%' and the '-', resulting in bad ejs syntax. This would not happen in other lines such as this for example:
```
<%- include('../partials/navbar')%>
```

nor does it happen in the second line in the code at the beginning.

In addition, it also marks said third line with red-squigly line below the '<' and '>' (I think it is below it) no matter if with or without the white space.

I should clarify that the code that I cited at the beginning works fine.

Good luck and have a nice day!

VS Code version: Code 1.50.1 (d2e414d , 2020-10-13T15:06:15.712Z)

此条需求为用户对于软件某些功能的正确性的需求。在用户使用软件过程中，不能出现bug以及一些意料之外的错误是基本的需求。

此类需求还有以下issue：

16: ${fileDirname} does not resolve correctly

17: natvis does not support enum class template argument types

19: [1031/181118.450:ERROR:registration_protocol_win.cc(103)] CreateFile: 系统找不到指定的文件。(0x2)

23: Not Getting Auto Information

28: grid-template-area not working

30: VCS sync error bad handling

32: Safari: terminal focus context key not updated after hide

36: EDITING PROBLEM. :(

39: Can't connect to extensions marketplace

40: 100% CPU Usage - Code Helper (renderer)

46: [regression] Files served by readonly filesystem show double '(read-only)' suffix in titlebar again

48: "Restart to Update" does not restart on macOS 11.0.1 (Big Sur)

53: Can't save workspace locally when using remote extension

54: configurationDefaults contribution changes JSON auto-complete behavior

56: Emmet doesn't remove snippets if user moves snippet file to another directory

57: Left overs in terminal after deleting text

58: Dropdowns don't work with touch screen

59: Webview repeats some keyboard shortcut execution

71: Locaolization bug:zh-cn: Preference: files.trimTrailingWhitespace

79: execSync('code --version').toString() returns node.js version instead of vs code version

81: File save time stamp is not correct

82: TS Server fatal error: Cannot read property 'charCount' of undefined

83: Debug hover expands over cursor

85: Syntax highlighting in HTML is wrong when @page rule contains more than one nested rule

90: windows crashed

91: xterm.css is out of date

96: Notebook drag titles are in front of scrollbar

98: Format on save for TypeScript files sometimes indents JSDoc comments incorrectly

99: Invalid predictions should be invalidated after some timeout

五、实验结果与效果分析

实验结果：

　　此次实验我们只抽取了一百条问答，并把每条问答直接当做一个需求进行分析。最终，把这一百条问答人工划分为了三类需求：外观需求、使用功能需求、软件正确性需求。具体分类内容见上。

效果分析：

　　此次实验满足了基本要求，最后的效果呈现还有许多不足，具体如下：

　　1、由于只抽取了一百条问答，数据并没有涵盖到需求的所有类别，不够全面。

　　2、对需求进行人工分类，难免存在人为失误，但又能力有限，还不会使用机器学习中的分类方法或者自然语言处理技术。

　　3、分类的标准较为主观，是根据每一个问题及其官方提供的解答作为分类依据，很大部分靠的是自我理解，去没有一套统一的分类标准。

六、 总结

本次实验完成了基本任务，但还有很多需要改进的方面。虽然使用了爬虫获取数据源，减轻了一定的工作量，但是在分类过程中对一百条问答是人工分类，还是存在一定工作量和人为误差，希望在下次实验中更有改进。