

实验一：需求获取

小组成员：171860566 王盼盼

171860565 刘振国

成绩比例 1:1

实验目标

软件需求的抽取与分类。

实验基本思路

选定一个开源 IDE 项目，确定可能的信息来源，获取有效信息，使用特定技术获取潜在需求，并进行分类。

实验步骤

(1) 确定 IDE 项目

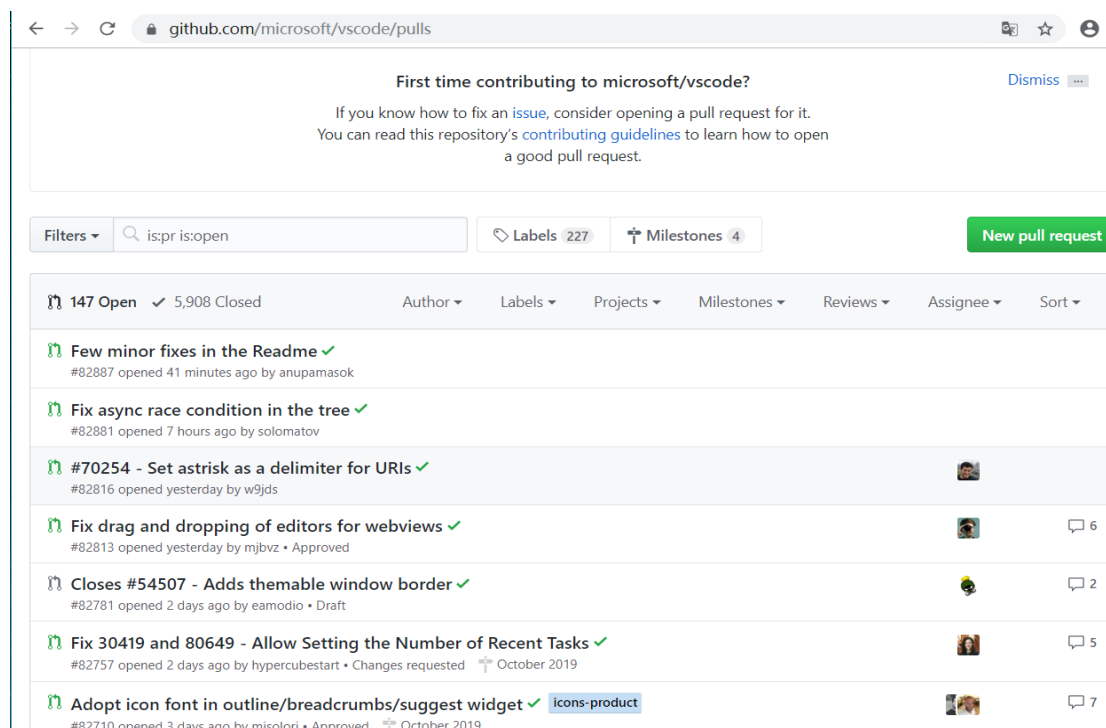
本次实验我们选取 Visual Studio Code 作为 IDE 的研究项目。

(2) 明确信息源

根据实验指导中思路一的提示，我们明确 vscode 托管在在 GitHub 上的地址作为信息来源，实验的思路是对该项目中的 Pull Request（地址：<https://github.com/microsoft/vscode/pulls>）作为一种需求，获取相关信息。

(3) 获取数据

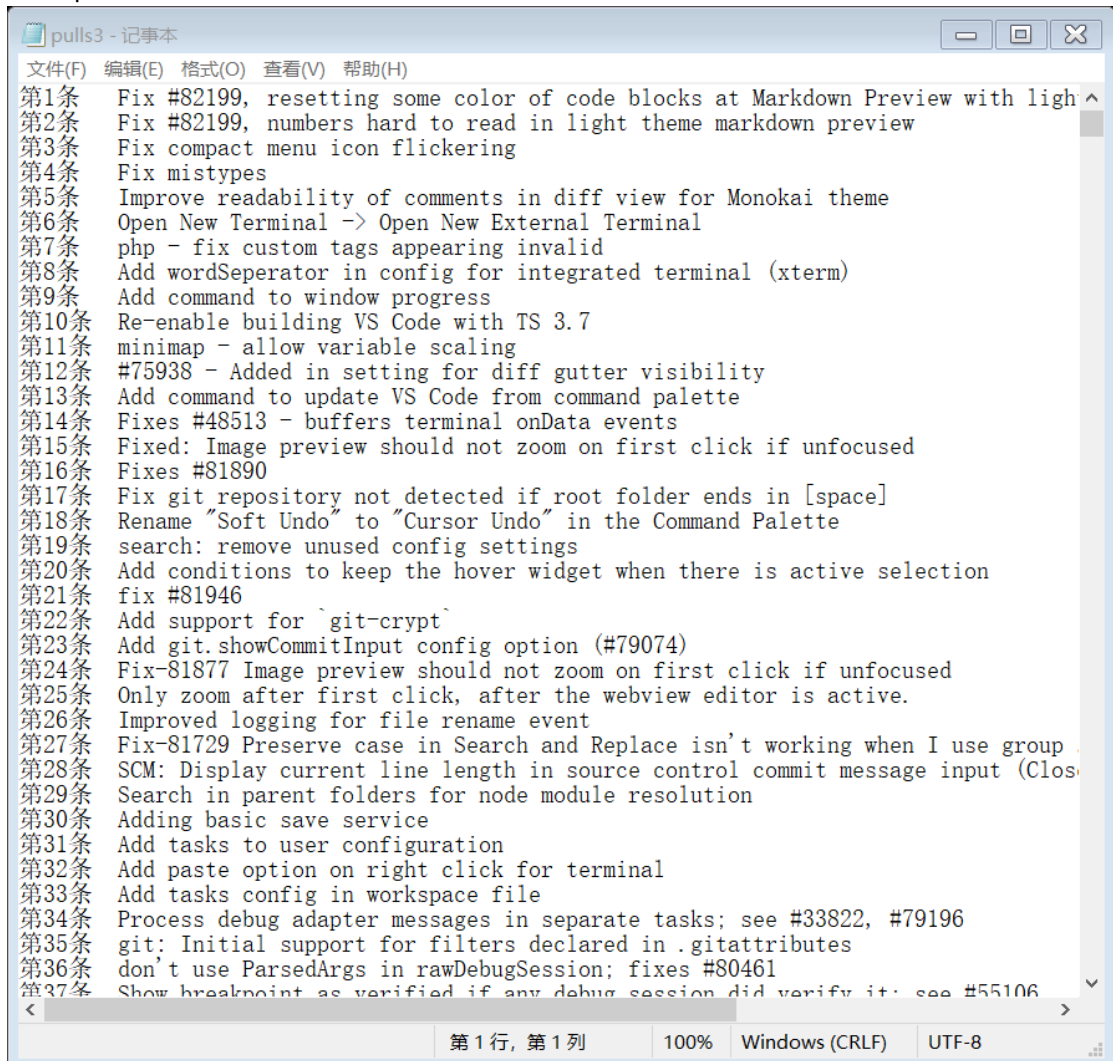
在明确信息源之后，我们在相应地址找到了 PR 的信息。



The screenshot displays the GitHub interface for the 'microsoft/vscode' repository, specifically the 'Pulls' (Pull Requests) section. At the top, there's a navigation bar with the repository name and a 'Dismiss' button. Below this, a message prompts first-time contributors to read the 'contributing guidelines'. The main area features a search bar with the filter 'is:pr is:open' and buttons for 'Labels' (227) and 'Milestones' (4). A 'New pull request' button is also visible. The pull requests are listed in a table with columns for status, title, number, author, and date. The first few entries are:

Status	Title	Number	Author	Date
Open	Few minor fixes in the Readme	#82887	anupamasok	opened 41 minutes ago
Open	Fix async race condition in the tree	#82881	solomatov	opened 7 hours ago
Open	#70254 - Set astrisk as a delimiter for URIs	#82816	w9jds	opened yesterday
Open	Fix drag and dropping of editors for webviews	#82813	mjbvz	opened yesterday
Open	Closes #54507 - Adds themable window border	#82781	eamodio	opened 2 days ago
Open	Fix 30419 and 80649 - Allow Setting the Number of Recent Tasks	#82757	hypercubestart	opened 2 days ago
Open	Adopt icon font in outline/breadcrumbs/suggest widget	#82710	misolori	opened 3 days ago

我们发现这些 PR 的标题可以概括相应的需求信息，为了获得众多的标题信息，我们使用 Python 语言写代码（文件：vs_pulls_get.py）来抓取这些信息并存入文件（pulls3.txt）。



```
pulls3 - 记事本
文件(F) 编辑(E) 格式(O) 查看(V) 帮助(H)
第1条 Fix #82199, resetting some color of code blocks at Markdown Preview with light theme
第2条 Fix #82199, numbers hard to read in light theme markdown preview
第3条 Fix compact menu icon flickering
第4条 Fix mistypes
第5条 Improve readability of comments in diff view for Monokai theme
第6条 Open New Terminal -> Open New External Terminal
第7条 php - fix custom tags appearing invalid
第8条 Add wordSeparator in config for integrated terminal (xterm)
第9条 Add command to window progress
第10条 Re-enable building VS Code with TS 3.7
第11条 minimap - allow variable scaling
第12条 #75938 - Added in setting for diff gutter visibility
第13条 Add command to update VS Code from command palette
第14条 Fixes #48513 - buffers terminal onData events
第15条 Fixed: Image preview should not zoom on first click if unfocused
第16条 Fixes #81890
第17条 Fix git repository not detected if root folder ends in [space]
第18条 Rename "Soft Undo" to "Cursor Undo" in the Command Palette
第19条 search: remove unused config settings
第20条 Add conditions to keep the hover widget when there is active selection
第21条 fix #81946
第22条 Add support for `git-crypt`
第23条 Add git.showCommitInput config option (#79074)
第24条 Fix-81877 Image preview should not zoom on first click if unfocused
第25条 Only zoom after first click, after the webview editor is active.
第26条 Improved logging for file rename event
第27条 Fix-81729 Preserve case in Search and Replace isn't working when I use group
第28条 SCM: Display current line length in source control commit message input (Close)
第29条 Search in parent folders for node module resolution
第30条 Adding basic save service
第31条 Add tasks to user configuration
第32条 Add paste option on right click for terminal
第33条 Add tasks config in workspace file
第34条 Process debug adapter messages in separate tasks; see #33822, #79196
第35条 git: Initial support for filters declared in .gitattributes
第36条 don't use ParsedArgs in rawDebugSession; fixes #80461
第37条 Show breakpoint as verified if any debug session did verify it; see #55106
```

(4) 需求分析

通过用户在使用 vs 的不同场景分析，可将需求大致分为下面几类：

- 1 编辑
- 2 编译
- 3 调试
- 4 资源管理
- 5 调试
- 6 其他

(5) 需求分类

在获得这些信息之后，为了分析不同的信息。我们首先对获得信息进行了词频分析（代码实现文件：1.cpp）获得了从高到低的出现的频率词汇。之后，在这些高频词汇中选取有意义的具有代表性的词汇，将这些词汇出现过的那些 PR 标题信息单独提取出来（代码实现文件：2.cpp），完成初步分类，再通过对词意相近或相关的

分组进行语义分析，完成二次分类：

1.编辑

- 查找指定内容
- 查找并替换指定内容
- 支持多种字体
- 显示行号、列号
- 关键字联想
- 自动对齐
- 多行注释功能
- 自定义快捷键
- 支持文件浏览器水平滚动
- 将 VSCode 设置为 git 编辑器
- 网格编辑器布局功能
- 在编辑器状态栏中显示所选行数
- 实施最近关闭的编辑器，使其在与关闭时相同的位置打开。
- 当检测到文件编码为非 UTF 时提示

2.编译

- 支持多种语言的语法
- 语法检查，提示出错位置和出错类型
- 语义检查，提示出错位置和出错类型
- 单词拼写检查功能，突出显示错误

3.调试

- 断点调试
- 带注释的反汇编
- 显示栈状态
- 跨工作区文件夹的复合调试
- 选择和启动调试配置
- 多个调试会话

4.用户界面

- 支持多种语言
- 查找所有按钮
- 按钮功能
- 翻译的快捷键
- 主题的颜色自定义
- 自定义关键字颜色
- 用户自定义新的子菜单
- 窗口（自动适应）、拆分窗口

5.资源管理

- 文件路径（在窗口标题中显示文件所在的文件夹路径）
- 从声明处转到定义的功能
- 为当前项目提供清晰的类视图
- 暂存新文件
- 文件流
- 根据文件内容建议文件名

无标题文件第一行文件名功能
预览文件
允许写入大文件（>256M）
如果项目中有未保存的文件，提交时显示警告
保留大小写但不区分大小写的文件处理
允许在工作区编辑中重命名/创建/删除文件

6.外部扩展

支持当前用户安装 Git Bash

7.版本控制系统

支持 git

仓库文件说明

data

word_frequency.txt	词频统计结果
pulls3.txt	PR 爬取结果
pullswithlink_2.txt	附链接爬取结果
其余.txt 文件	按关键词提取爬取的信息

code

Word_frequency.cpp	词频统计程序
select_keyword.cpp	按关键词分类程序
vs_pulls_get.py	vscode 的 PR 标题爬取程序
vs_pulls_get_with_site.py	附带链接的 vscode 的 PR 标题爬取程序