# 小说推荐工具需求说明书

修改记录

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| 修改编号 | 版本 | 修改内容 | 修改人 | 修改日期 |
| 1 | 1.0.0 | 新建需求说明 | 黄俊远 | 2020.8.3 |
|  |  |  |  |  |

目录

[小说推荐工具需求说明书 1](#_Toc47395768)

[1． 简介 2](#_Toc47395769)

[1.1 目的 2](#_Toc47395770)

[1.2 范围 2](#_Toc47395771)

[1.3 参考资料 2](#_Toc47395772)

[1.4 概述 2](#_Toc47395773)

[2． 整体说明 2](#_Toc47395774)

[2.1 产品总体效果 2](#_Toc47395777)

[2.2 假设与依赖关系 2](#_Toc47395778)

[2.3 系统运行环境 3](#_Toc47395779)

[3． 具体需求 3](#_Toc47395780)

[3.1 功能性 3](#_Toc47395782)

[3.2 可用性 4](#_Toc47395783)

[3.3 可靠性 4](#_Toc47395784)

[3.4 安全性 4](#_Toc47395785)

[3.5 性能 4](#_Toc47395786)

[3.6 可支持性 4](#_Toc47395787)

[3.7 设计约束 4](#_Toc47395788)

[3.8 接口 4](#_Toc47395789)

[4． 界面原型 5](#_Toc47395790)

[5． 需求优先级 5](#_Toc47395791)

[6． 软件功能列表 5](#_Toc47395792)

## 简介

### 目的

小说推荐工具，目的是根据已有的小说内容，推荐类似的小说，减少了去各大网站翻找小说的麻烦。

### 范围

本工具主要适用于小说爱好者，以及缺少时间和精力去筛选小说等文学作品的人。

### 参考资料

Python爬虫案例 <https://blog.csdn.net/weixin_44552215/article/details/103668739>

文章相似度比较算法 <http://www.ruanyifeng.com/blog/2013/03/cosine_similarity.html>

### 概述

本产品主要解决了两个问题：

1. 在线小说阅读受网站约束 ，可能伴随收费、广告、加载慢等问题。本产品将网络小说下载到本地，保存为纯粹的文本文件，能有效解决以上问题。
2. 在线小说一般按照类别、标签的方式进行分类推荐。对于小说内容的分类较少。本产品使用内容中的关键字进行分析比较，具有深度学习的效果，能在不断的筛选过程中不断提升相似度比较的准确性。真正找到用户感兴趣的内容，让用户停不下来。

## 整体说明



### 产品总体效果

第一期的产品尽量简单。使用命令行的形式，输入小说文件的路径，返回相似度前几的小说路径及对应的相似度。

### 假设与依赖关系

本产品需要本地服务器作为存储小说和计算相似度的载体，初期使用个人PC实现。

本产品需要使用数据库进行数据存储，使用mysql数据库既免费又好用。

本产品使用的文章相似度算法需要对算法进行训练，初期训练的时间和数量不足，推荐效果可能会不尽人意。

### 系统运行环境

#### 硬件环境

服务器一台

CPU intel

内存 8G

#### 软件环境

Windows操作系统

Mysql数据库

Python开发环境

Pycharm IDE

第三方网络库、算法库等

## 具体需求



### 功能性

#### 爬虫功能

提供小说网站的home地址，能根据home地址自动下载，分类小说文件

#### 文章特征提取功能

根据本地保存的txt小说文件，获取出文章的关键字特征，存入数据库

#### 相似度比较功能

根据文章特征，比较文章相似度。包含1V1比较；1VN比较。

### 可用性

能正确下载大部分网站的小说内容，并正确排版。

相似度比较算法能大致分析出文章的相似度

### 可靠性

小说爬虫和相似度比较算法分开执行。小说以本地文件、数据库等方式保存，系统断电，崩溃等问题不会损失大量数据。

### 安全性

暂无

### 性能

暂无

### 可支持性

程序分模块编写，统一接口，方便后续维护修改

### 设计约束

给予win下的Python开发，数据库使用mysql

### 接口

#### 爬虫功能

输入为小说网站主址

输出为本地小说文件和数据库中的小说描述信息

#### 文章特征提取功能

输入为本地小说文件

输出为文章关键字的向量信息

#### 1V1相似度比较功能

输入为两部本地小说文件

输出为小说相似度值，在0和1之间

#### 1VN相似度比较功能

输入为一部本地小说文件，数字N

输出为已下载的小说中相似度最高的N部小说的名字。

## 界面原型

暂无

## 需求优先级

|  |  |  |
| --- | --- | --- |
| 需求编号 | 需求描述 | 优先级 |
| FUN-1 | 爬虫 | 高 |
| FUN-2 | 推荐算法 | 中 |
| FUN-3 | 人机交互 | 中 |
| SUPP-1 | 性能优化 | 低 |

## 软件功能列表

|  |  |  |
| --- | --- | --- |
| 功能编号 | 功能描述 | 产品版本 |
| FUN-1 | 小说爬虫 | 1.0 |
| FUN-2 | 相似度比较算法啊 | 1.0 |
| FUN-3 | 控制台输入输出 | 1.0 |
| FUN-4 | 比较算法性能，数据库访问性能 | 1.1 |