

基于知识图谱的 python 学习平台

0 背景

网络的发展和普及也促成了新的教学形式的出现，人们逐渐将视线关注于网络在线学习。由于教育资源分配不均，不同地区的学习者获得的教育水平也不同。作为一种不受时空限制、资源丰富的教学形式，互联网学习被视为一种获取知识的新途径（康准、王德军，2018）。在这一方面，使用可视化技术展现知识图谱，够描述知识资源及其结构关系，使用户更清晰地了解领域知识的脉络和走向。然而目前网络学习平台鲜有利用知识图谱进行数据组织。Python 是一种广泛使用的解释型、高级编程、通用型编程语言，由吉多·范罗苏姆创造，第一版发布于 1991 年。Python 的设计哲学强调代码的可读性和简洁的语法。相比于 C++ 或 Java，Python 让开发者能够用更少的代码表达想法。不管是小型还是大型程序，该语言都试图让程序的结构清晰明了。

由于 Python 语言的简洁性、易读性以及可扩展性，Python 在国内外受到 IT 行业人员的热烈追捧。截止至 2019 年 9 月，Python 已荣登 IT 行业编程语言排行榜第三名。国内外的大学也都纷纷开设起 Python 课程。

将知识图谱与 Python 学习平台相结合，能更好地帮助用户了解学习 Python 课程。Python 学习平台，旨在为想学习 Python 的用户群体，提供更好的帮助和更加便利的支持。

1 可行性报告

1.1 行业市场分析

随着 python 语言使用率的提高以及信息化建造的推动，越来越多的线上 python 学习平台 app 出现。

在传统的 python 线上学习平台中，往往知识点比较散，知识点间的关联不够紧密，用户往往很难通过搜索精确找到自己所需的知识点；并且它们往往功能比较混杂，通常是几个板块混在一起。大多数平台更注重的是问题的问答和求解功能，而不是基础知识的普及功能。虽然也有部分平台有基础知识教程，但是也缺少相关的搜索和推荐功能，以及论坛的问答功能。

由上可知，虽然现在市面上已经有不少的 python 学习平台，但是能真正能做到既有基础知识点个性化推荐和智能搜索，又有论坛问答功能的平台还是不多，因此，只要我们能够把个性化推荐和智能搜索的优势发挥出来，基于知识图谱的 python 学习平台将会有比较好的前景。

1.2 竞争对手分析

目前市面上比较典型的 python 学习平台应用有：Python 编程狮、python 教程、Python 学习通。

①Python 编程狮：

“Python 编程狮”是 W3Cschool 编程狮旗下专门为零基础 Python 编程爱好者打造的一款入门工具 App，致力于帮助初学者入门 Python 编程语言。主要功能有 1.微课即学即练：游戏化闯关学习，为初学者打造。2.丰富教程：Python 内容、知识点丰富。3.Python 题库：在练习中巩固，查缺补漏。4.在线编译：实现手机端写代码。

②python 教程：

“python 教程”应用是一款带视频功能的 python 入门学习苹果应用，另有图文并茂教程。教程的内容从最基础到高级阶段。主要功能有：1.带视频教程学习 2.图文并茂教程学习 3.视频收藏功能。

③Python 学习通：

“Python”学习通包含以下主要内容：1.Python 编程基础，面向对象编程学习资料。2.Python 数据处理，Python Numpy 和 Python Panda 学习资料。3.测试题库，提供对 Python 编程的知识点和技术点进行全面测试的题目。

通过对以上三个典型的 python 学习平台应用进行分析，可以看出目前的 Python 学习平台大多都有题库练习、在线编译和视频学习功能，这是我们应用没有的功能；但是尚未发现有平台有关于知识点个性化推荐、智能搜索以及论坛问答功能，因此，我们应该发挥出我们知识图谱的优势。

1.3 自身条件分析

我们的组员都是来自计算机学院的学生，都对 python 的知识有一定的了解，能更好地作为学生从用户得角度去做需求分析。在技术层面上我们已经初步构建好了知识图谱，有组员熟悉服务器的搭建过程，也有组员有安卓开发的基础，小组的分工也比较明确，采用迭代式开发方法，前端和后端并行开发。在时间上，组内组员本学期课程都不多，时间较充足，如果有多的时间可能会新加功能板块。不过现网上能找到关于知识图谱构建的学习平台应用并不多，相关资料和参考文献比较少，这可能是需要我们去克服的一个难点。

2 产品定位及目标

本产品定位为学习类型，适合想要学习 Python 知识的用户群体。无论用户是零基础入门 Python，还是想深入了解 Python，都可以在该平台满足用户的需求。同时，该平台也是 Python 爱好者的聚集地，在平台上能互相讨论。

综合各种原因，本产品的目标用户应具备以下特点和个性：

1、从未接触过编程或者从未学习 Python 语言，但想要从零开始入门 Python；有过一定 Python 使用经验，但是想要深入学习尚未掌握的 Python 知识。

2、想要搜索学习 Python 相关方面知识，比如爬虫、WEB 开发、数据分析、

人工智能等等。

- 3、 熟悉手机的各种新型应用，经常使用学习类软件，能够持续每天学习。
- 4、 社交网络的忠实用户，喜欢使用网络交流。

3 产品内容策划

3.1 核心功能模块

(1) 个性化推荐

该功能通过用户曾经学习浏览过的知识点，从知识图谱中寻找该知识点，并智能推荐该用户可能感兴趣的知识点或下一阶段将要学习的知识点推荐给用户，在推荐页面供用户选择浏览。

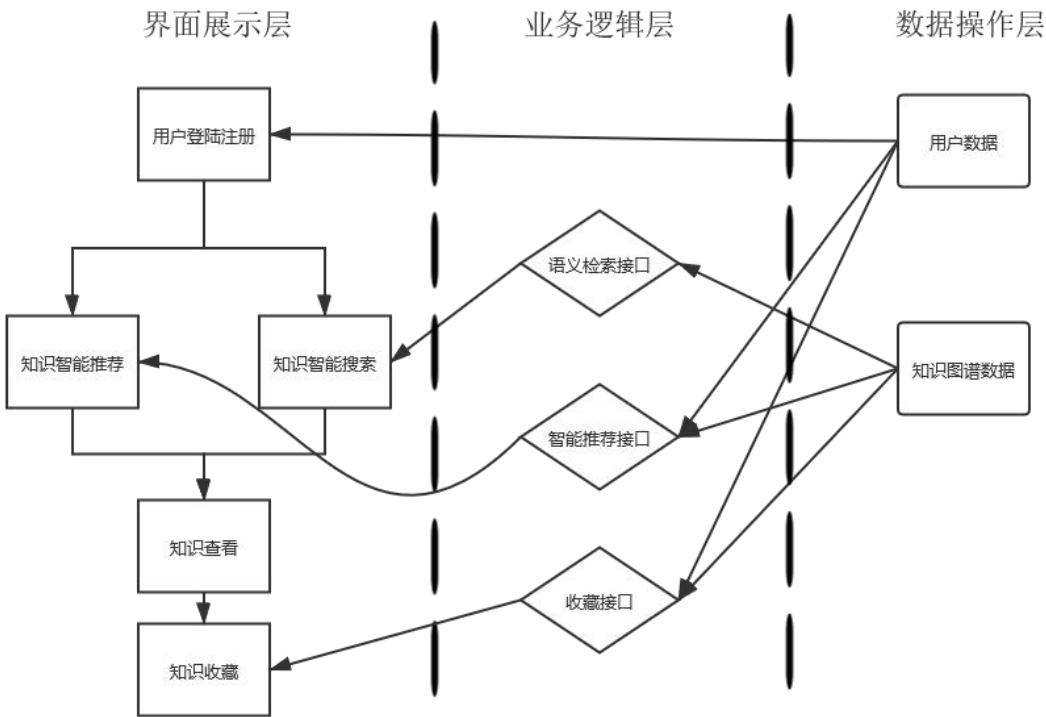
(2) 知识搜索

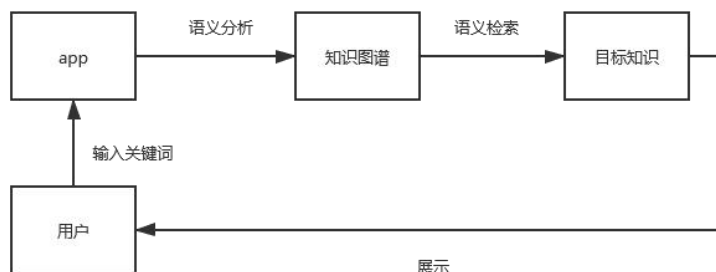
通过用户输入关键词，然后通过语义分析，在知识库中寻找相关知识，并返回给用户，供用户选择。

(3) 交流论坛

用户群体可以再此进行沟通交流。用户可以在此进行提问，其他用户可以在其疑问后发言帮助解答。用户也可以在此分享学习心得体会。

3.2 处理流程和数据流程图





4 技术解决方案

（1）前端：

开发语言：Java

开发工具：Android Studio

Gradle 版本：5.4.1

（2）后端：

服务器框架：Spring Boot

①数据库管理系统：

1.MySQL：用于存储用户信息、论坛内容

2.Neo4j：图数据库，用于存储知识图谱

②知识图谱构建方法：

1.使用 Python 爬虫爬取有关 Python 的知识内容

2.使用基于预训练 Bert 神经网络模型的 Bert-BiLSTM-CRF 从第一步中爬取到的内容进行实体抽取

3.使用 K-Means 聚类方法分析知识点间的关系，并对知识点进行分类。

4.将数据处理成 csv 格式文件，导入 Neo4j 图数据库。

③智能搜索：

1.对用户输入的查询语句进行分词处理，提取关键词

2.对关键词进行同义转换，通过用户输入的关键词匹配图数据库中的具体节点

3.若不能匹配到对应的节点，则利用词汇相似性算法匹配与关键词相似度最高的概念作为搜索结果。

5 推广方案

为了使更多志同道合的用户加入该平台，现备有如下推广方案：

1、在高校内推广：在学生之间推广，为学生提供一个沟通学习的平台，利用小

- 圈子逐步扩散，逐渐让更多学生了解到该平台。
- 2、平台互推：联合其他的学习平台进行互推。
 - 3、平台内推广：通过一定的奖励机制，鼓励用户将平台分享到互联网平台上，吸引新的用户。

6 运营策划书

本产品为普通的学习用户提供免费服务，商业运营模式如下：

1、广告服务

收取定广告费用，为各类学习平台和学习机构做广告推广，使更多用户涌入到企业的商业活动。帮助企业发现忠诚用户和发展潜在的用户，扩张产品服务的营销渠道，提高产品影响力。

2、用户会员

用户可以选择开通会员获取更好的服务，如学习计划的制订，学习内容的多元化推荐、更多的个人论坛等。