

# 基于B-S架构与KNN算法的自主学习交流平台设计

吴房胜<sup>1</sup>, 何世清<sup>2</sup>

(1.安徽工商职业学院信息工程学院,合肥 231131;2.安徽农业大学经济技术学院,合肥 230011)

**摘要:**开发一套基于B-S架构的自主学习交流平台,为用户创建一个友好的学习交流环境。平台采用Java语言开发,通过MySQL5.7对数据进行存储,依托Tomcat7.0服务器发布,最终通过浏览器展现给用户。开发过程中引入了KNN算法,可对乱填信息注册平台的用户有效拦截过滤,自主预测平台中用户的年龄结构、热门话题等信息,收集热门话题创建学习兴趣小组,并通过用户对学习资源的评价等行为的分析,对表现较好的用户给予更优质的服务与优惠,从而更好地提高平台的质量。

**关键词:**KNN算法,学习平台,信息过滤,行为预测

**中图分类号:**TP311.52 **文献标志码:**A **文章编号:**1673-1891(2020)03-0043-04

## Design of Autonomous Learning and Communication Platform Based on B-S Architecture and KNN Algorithm

WU Fangsheng<sup>1</sup>, HE Shiqing<sup>2</sup>

(1. School of Information Engineering, Anhui Business Vocational College, Hefei, Anhui 231131, China;

2. School of Economics and Technology, Anhui Agricultural University, Hefei, Anhui 230011, China)

**Abstract:** In this paper we developed an autonomous learning and communication platform based on B-S architecture to create a friendly learning and communication environment for users. The platform is developed with java language. The data is stored through MySQL5.7, released on Tomcat7.0 server, and finally displayed to the user through a browser. The KNN algorithm is introduced in the development process, which can effectively block and filter users who randomly fill in the information registration platform, autonomously predict users' age brackets, their hot topics and other information, and collect hot topics to create learning groups of different interests. Through analysis of user behaviors such as their evaluation of study resources, better services and discounts are given to users with better performance, thereby quality of the platform will be improved.

**Keywords:** KNN algorithm; learning platform; information filtering; behavior prediction

## 0 引言

在信息时代,自主学习交流平台为广大学习者提供了一个交流平台,让广大用户在学习中遇到的各种困扰,通过交流平台可快速交流解决,具有较强的实践价值和现实意义。自主学习交流平台整体上易于操作,系统的可见性非常好,均为图形化界面,许多页面都是单击选择即可输入的模式,对用户的要求不高,超级管理员会一些基本的电脑操作便可。该平台从功能上主要包括用户提问与回答、用户搜索问题、管理员管理问题、管理用户等主要功能模块<sup>[1]</sup>。

## 1 系统概述

该自主学习交流平台采用的是B-S三层架构开发模式,对平台的和后台数据库进行了分离,这样能很好的保障数据的安全性和平台的稳定运行。前端和后台数据的交互如图1所示。

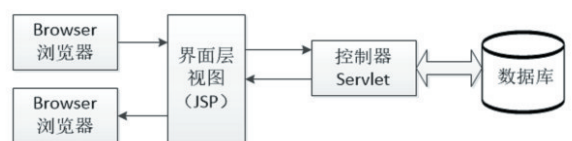


图1 前后台数据交互图

平台主题分为前台功能和后台功能,每个功能

收稿日期:2020-02-17

基金项目:安徽省高校自然科学研究重点项目(KJ2019A1166);安徽省省级质量工程项目(2018sjjd114);安徽省省级教学研究项目(2019jyxm0746);安徽工商职业学院校级重点教研项目(2018xjyyZD02)。

作者简介:吴房胜(1983—),男,安徽安庆人,讲师,硕士研究生,研究方向:物联网与智能控制技术。

的子模块都进行集中式管理,使每个模块除了完成自己的子功能以外,不同模块之间还要存在一定的联系,大大提高了平台的稳定性。该平台总体功能如图2所示<sup>[2]</sup>。

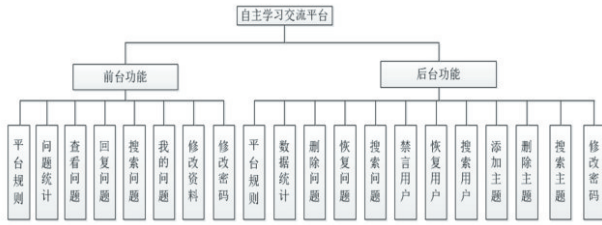


图2 自主学习交流平台总体功能图

自主学习平台有超级管理员和普通用户两种身份。游客在平台首页点击注册按钮之后,按要求正确填写相关信息,并选择遵守平台规则之后即可成功注册自己的账号,登录自己账户来到平台首页,平台主要由平台规则、讨论区、我要提问、搜索问题、我的问题、个人资料和修改密码等功能模块组成。

管理员在浏览器输入 127.0.0.1:8080/admin 来到登录界面,正确输入账号和密码,进入平台管理首页,超级管理员可看到平台规则,可对当天的数据进行统计,还可对用户提出的问题进行管理,对不符合平台规则的问题进行删除操作,并对不遵守平台规则的用户进行禁言处理,查看平台自主预测的用户的年龄结构、所在地、热门话题等信息,收集热门话题创建学习兴趣小组,并利用平台通过用户对学习资源的评价及点赞等行为的分析,超级管理员可对表现较好的用户给予更优质的服务与优惠,可以添加新的主题和删去用户们不需要的主题,还可查看所有注册会员的所有信息,在用户忘记登录密码时,联系管理员即可找回登录密码<sup>[3]</sup>。

普通用户注册账号并登陆成功之后,便进入自主学习交流平台首页,在此可以看到各个功能模块的展示(图3)。超级管理员输入用户名和密码登录成功后,即可进入平台管理员首页(图4)。

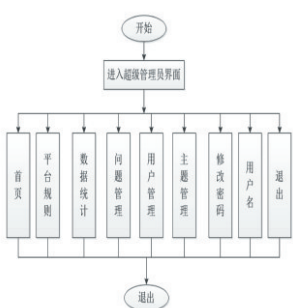
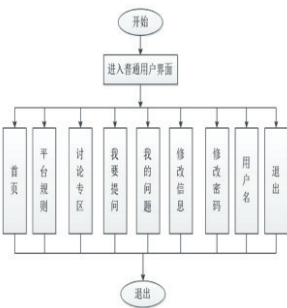


图3 普通用户模块结构

图4 超级管理员模块结构

## 2 KNN 算法概述

KNN算法中文名称叫最近邻算法,是将给定的检验元组与其相似的训练元组实施比较进行学习,该算法是一种应用于分类与回归的无母数统计算法,主要应用在信息检索与过滤、网页分类、数据回归及产品推荐等。该算法是假设一个样本,在特征空间中的  $k$  个最相似(最邻近)的样本中的绝大部分都归类于其中一个类别,那么这个样本也归类于这个类别,该算法所选取的邻居均为已正确分类的对象。简单来讲就是有一堆已分类好的数据信息  $\omega_1$ 、 $\omega_2$ 、和  $\omega_3$ (图5),当加入一个新数据信息  $x_u$  时,  $x_u$  和训练数据中的每个点算垂直距离,再挑选与  $x_u$  距离最近的  $K$  个点,看这  $K$  个点属于哪种类型,如果  $x_u$  中多数点属于  $\omega_1$ ,则根据少数服从多数原则,将  $x_u$  这个新数据归类到  $\omega_1$  中<sup>[4]</sup>。

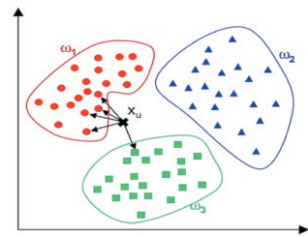


图5 KNN算法解析图

通常测试的样本和训练样本的相似度采用欧式距离进行测量,2个文本向量  $a$  与  $b$  之间的欧氏距离表达式如式(1)。

$$D(a, b) = \sqrt{\sum_i (a_i - b_i)^2} \quad (1)$$

如果训练样本个数能够趋于无穷大,并且  $K$  值固定,则这  $K$  个近邻全部将收敛于  $x_u$ 。类似于最近邻规则,  $K$  个近邻的标记全为随机数据,概率  $P(\omega_i | x_u)$  ( $i=1, 2, \dots, K$ ) 全为互相独立的<sup>[5]</sup>。如果  $P(\omega_m | x_u)$  是那个较大的后验概率,按照贝叶斯分类原理,则选择  $\omega_m$  这个类别。最近邻规则根据  $P(\omega_m | x_u)$  这个概率选择类别,而按照  $K$  近邻规则,必须当  $K$  个最近邻中的大部分标记在  $\omega_m$  类别中,方可判定为  $\omega_m$  类别,这种断定的概率为式(2)。

$$\sum_{i=(K+1)/2}^K \binom{K}{i} P(\omega_i | x_u)^i [1 - P(\omega_m | x_u)]^{K-i} \quad (2)$$

一般  $K$  值越大,选择  $\omega_m$  类别的概率就越大。

## 3 系统软件设计

在用户进入自主交流平台的登录界面之后,输入对应的账号和密码,系统对输入的信息进行判

断,并作出相应的应答(图6)<sup>[6]</sup>。

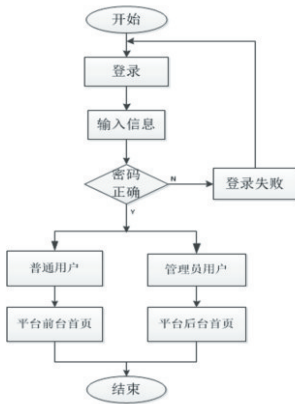


图6 用户登录

在Java编程中,首先计算测试样本与每个训练样本之间的距离值 $D$ ,算出 $K$ 个最临近样本中的最大距离值 $D_{max}$ ,将 $D$ 与 $D_{max}$ 值进行比较,如果 $D$ 比 $D_{max}$ 小,即作为 $K$ 最近邻样本,如此往复,统计出每个类标号出现的次数,将次数最多的作为测试样本的类标号(图7)<sup>[7]</sup>。

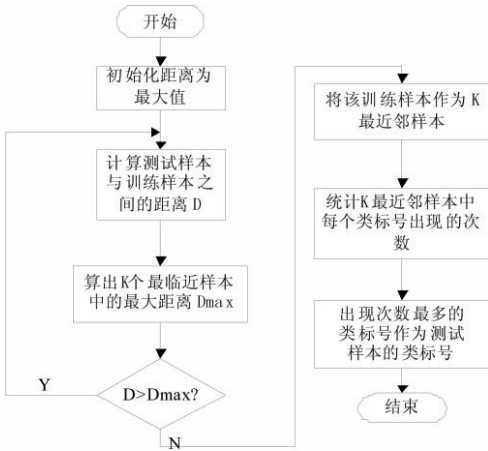


图7 KNN算法流程图

程序编写KNN算法主体类时,定义comparator作为优先级队列的比较函数,得到的距离越大,其优先级越高,具体程序如下所示。

```
public class KNN
{
    private Comparator<KNNNode> comparator =
new Comparator<KNNNode>()
    {
        public int compare(KNNNode x1, KNNNode x2)
        {
            if(x1.getDistance() >= x2.getDistance())
            { return -1;
            }
        }
    }
}
```

```
else
{ return 1;
}
};

在计算测试样本与训练样本之间的距离时,设D1为测试样本,D2为训练样本,distance为距离返回值。
public double calDistance(List<Double> D1,
List<Double> D2)
{
    double distance = 0.00;
    for (int i = 0; i < D1.size(); i++)
    {
        distance += (D1.get(i) - D2.get(i)) * (D1.get(i) -
D2.get(i));
    }
    return distance;
}
```

### 4 测试结果分析

系统完成后,首先启动tomcat服务器之后,然后在浏览器输入127.0.0.1:8080/com.study即可进入自主学习交流平台的首页(图8)。



图8 自主学习交流平台首页

自主学习交流平台的所有功能在用户成功注册账户之后即可免费使用,点击首页的注册按钮之后,按要求填写对应的信息并点击遵守平台规则,再点击注册按钮即可成功注册账号,但所填信息不符合平台规则,对乱填信息注册平台的用户,可实现无效拦截过滤(图9)。

管理员进入登录界面后,可实时查看系统统计出的年龄结构、性别结构等,通过年龄因素预测哪个年龄段对平台比较关注,便于后期更好的定向推广(图10)。

将KNN算法用到数据库中,在后台管理平台上,系统收集最新话题,并根据点赞数、浏览数、回复数等参数预测出最热话题,再进行归类并预测出





图 9 注册平台信息拦截过滤



年龄结构分析表

年龄	人数	比例
0-10	1	0.25%
10-20	4	0.25%
20-30	7	43.75%
30-40	1	0.25%
40-50	1	0.25%
50-60	1	0.25%
大于60	1	0.25%

性别比例分析表

性别	人数	比例
男	3	57.14%
女	6	42.86%

图 10 年龄分类结构图

热门话题,供用户学习,也给平台创建学习兴趣小组提供依据(图 11)。

单击问题管理,可看到注册者浏览的所有问题。点击右上角的违规问题统计按钮,系统根据 KNN 算法可展示所有判定为违规的问题,对违反平台规则的问题,点击对应问题后的删除按钮即可删除(图 12)。

点击主题管理按钮,可看到供用户选择的主题(图 13)。根据 KNN 算法,平台根据用户对各个主

参考文献:

- [1] 殷杭华,陈丽萍,滕姿,等.基于 HRCT 模型的自主学习交流平台的设计与开发[J].嘉兴学院学报,2010,22(3):127-131.
- [2] 曹文渊.JAVA 语言在计算机软件开发中的应用[J].电子技术与软件工程,2019(2):53-54.
- [3] 岳青玲.Java 面向对象编程的三大特性[J].电子技术与软件工程 2019(24):239-240.
- [4] 张继东.MySQL 数据库基于 JSP 的访问技术[J].电子技术与软件工程,2017(15):169.
- [5] 刘彩利.C/S 和 B/S 混合体系结构的开发与应用[J].电子设计工程,2015,23(14):26-28.
- [6] 汪汇泉.基于 B\_S 架构的招生管理平台设计与实现[J].襄阳职业技术学院学报,2019(9):77-79.
- [7] 肖斌.分布式 KNN 算法在微信公众号分类中的应用[J].计算机应用,2017(S1):233.

(上接第 35 页)

- [14] MURALIDHARAN V, SUGUMARAN V.Selection of discrete wavelets for fault diagnosis of monoblock centrifugal pump using the J48 algorithm[J].Applied Artificial Intelligence,2013,27(1):1-19.
- [15] ISVORAN A,PITULICE L,CREASCU C T,et al.Fractal aspects of calcium binding protein structure[J].Chaos,Solutions & Fractals,2008,35(5):960-966.
- [16] 刘其和,李云明.LabVIEW 虚拟仪器程序设计与应用[M].北京:化学工业出版社,2011.
- [17] 孙楠楠.大型旋转机械振动监测与故障诊断知识体系的研究与实现[D].重庆:重庆大学,2006.
- [18] 田宁.单向阔声发射信号的故障诊断研究[D].昆明:昆明理工大学,2013.
- [19] PIOTR P,RAFAL D,TOMASZ O.Signatures verification based on PNN classifier optimised by PSO algorithm[J].Pattern Recognition,2016,12(60):998-1014.



图 11 热门话题预测节目



图 12 问题管理界面



图 13 主题管理界面

题的学习资源的评价进行分析,对表现较好的主题和用户给予更优质的服务与优惠,从而更好地提高平台的质量。如果某个主题不是很受用户的欢迎,管理员可单击该主题后的删除按钮进行删除。