

毕业论文

**题目：基于随机森林分类模型的**

**姓 名** 杨昊

**学 号** 201700301297

**学 院** 软件学院

**专 业** 软件工程

**年 级** 2017级

**指导教师** 蒋志方

**2021 年 5 月 15 日**

**摘要**

儿童及青少年时期的教育对每个人的成长都是十分重要的。最为理想的是在人生初期树立起正确的人生观价值观，然而每个人都是个独特的个体，拥有独一无二的天赋特征，如果能根据每个孩子的天赋特征，采用适合孩子的教育方式，那么孩子将会成长的更加迅速，更加的茁壮。然而，普通父母，老师并没有接受过专业的心理训练，不能像心理咨询师一样能够科学合理的分析出孩子的天赋特征。给每个家庭，每个班级都配备一个专业的心理咨询师是不可能实现的，即使是以学校为单位去分配心理咨询师也是难以实现的。那么让计算机程序去分析孩子的天赋特性，给每所学校的教学用计算机都配置上这样的程序，就能科学的了解每个孩子的天赋特征。而天赋分析具有很大的不确定性，每个地区乃至于每所学校的孩子都有不同的生活环境和心理状态，用过于硬性的规则去判断孩子的天赋是不可取的，结果也将会是不如人意的。同时统计各地区或者各学校的分析结果去规则化将会耗费巨大的人力。那么利用不同地区的统计结果去训练分类模型将之运用于该地区今后的天赋分析中，最终的结果是可信的。基于此，开发有一个基于随机森林分类模型的天赋分析系统是很有必要的。本文将介绍在开发此天赋分析系统时用到的技术和遇到的问题。

**Abstract**

目录

**第一章 绪论1**

键入章标题(第 2 级)2

键入章标题(第 3 级)3

**第二章 4**

键入章标题(第 2 级)5

键入章标题(第 3 级)6

**第三章 1**

键入章标题(第 2 级)2

键入章标题(第 3 级)3

**第四章 UI4**

键入章标题(第 2 级)5

键入章标题(第 3 级)6

**第五章 数据处理1**

键入章标题(第 2 级)2

键入章标题(第 3 级)3

**第六章 分类模型4**

键入章标题(第 2 级)5

键入章标题(第 3 级)6

**第七章 可视化1**

键入章标题(第 2 级)2

键入章标题(第 3 级)3

**第八章 4**

键入章标题(第 2 级)5

键入章标题(第 3 级)6

**致谢1**

**参考文献4**

模块一：GUI

模块二：数据处理

模块三：随机森林

模块四：反馈与可视化

整个应用程序的流程图

数据处理的流程图

随机森林的工作流程图

功能图：

问答

注册

登录

沙盘游戏系统

天赋雷达

分析结果

验证

心理分析

心理分析系统

数据流图：

第一层：

账号，密码 天赋建议

用户

用户

游戏创作内容 天赋雷达图

游戏问答结果

第二层：

流程图：

用户数据处理：

If(i<n)

i=0，n=len(userdata)

游戏数据

F

T

统计每种种类的数量

在数据库中查询其种类X

处理完毕

随机森林各方法的流程图：

决策树的判断过程：

已处理的用户数据

If(当前节点是叶子节点)

F

输出当前叶子节点的类别

Y

If（数据对应属性的值<当前节点的分割属性的值）

Y F

当前节点=右子节点

当前节点=左子节点

选择最优分裂属性和分裂阈值：

数据集xiang’yong

Feature in data.columns

|  |  |  |
| --- | --- | --- |
| MID | MN | category |
| 001 | 警察 | A |
| 002 | 消防员 | C |

模型种类表（主题一：我难忘的一堂课）

模型种类表（主题二：我的一次生日派对）

|  |  |  |
| --- | --- | --- |
| MID | MN | category |
| 001 | 警察 | B |
| 002 | 消防员 | A |

模型种类表（主题三）……

用户表

|  |  |  |
| --- | --- | --- |
| UID | Useraccount | password |
| 001 | 2800 | 108 |
|  |  |  |

类图：

**致谢**

首先要感谢我指导老师蒋志方老师，蒋志方老师每周都会了解我的工作进展，和同学们一起讨论系统的需求，为我们指明设计方向。还要感谢所有参考文献的作者，没有这些参考文献系统将难以实现**。**

**参考文献**

<https://blog.csdn.net/qq1483661204/article/details/77587881> DataFrame .ix 既可以根据行列标签，又可以根据行列数取数据。

<https://www.liaoxuefeng.com/wiki/1016959663602400/> Python使用手册（staticmethod不强制要求参数，可以不用实例调用）

<https://zhuanlan.zhihu.com/p/31417084> Vscode建立Python环境

<https://www.liaoxuefeng.com/wiki/896043488029600> git使用手册

<https://scikit-learn.org/stable/modules/classes.html#module-sklearn.metrics> sklearn阅览

<https://blog.csdn.net/abc_12366/article/details/80447228> 查看collections（集合及集合的方法）底层方法

<https://pandas.pydata.org/pandas-docs/stable/reference/api/pandas.set_option.html> pd文档阅读

<https://www.icode9.com/content-4-550560.html> 随机数种子

<https://blog.csdn.net/qq_36750158/article/details/82417666> checkbutton获取状态

<https://blog.csdn.net/lzw2016/article/details/80472649> 布尔索引

<https://www.cnblogs.com/fish-101/p/11339909.html> 乱序

<https://www.jianshu.com/p/fb6ee6cc5c1c> requests

<https://blog.csdn.net/qq_41556318/article/details/85108303> check