

在线学习行为分析数据可视化快速开发框架设计与实践

洪丹丹,李 飞,姚 磊,徐 墨,锁志海

(西安交通大学 网络信息中心,陕西 西安 710049)

摘 要: 为了充分挖掘高校在线学习平台中的用户学习行为,发挥各类用户行为数据的价值并增强各类在线统计数据的可读性,以统计学理论为基础,结合 JFinal 轻量级框架及 ECharts 前端数据可视化技术研究设计了一款可扩展、通用性强的轻量级数据分析可视化快速开发框架.并以典型的在线学习平台中用户最喜爱课程分析实例为背景,对其部分数据库模型、算法指标模型、前端数据可视化及后端数据处理等关键技术进行了详细研究及实现.以框架在西安交通大学在线学习平台中的实践结果表明,其可以满足数据分析快速可视化需求且支持数据关联分析和数据逐级下钻可视化.最后,对未来大数据环境下的高校学习平台用户学习行为分析提出建设意见.

关键词: JFinal;ECharts;JSON;RESTful;数据可视化;在线学习;大数据;行为分析

中图分类号: TP399

文献标识码: A

文章编号: 1000-7180(2018)07-0006-07

Design and Implementation of A Fast Data Visualization Framework For E-learning Behavior Analysis

HONG Dan-dan, LI Fei, YAO Lei, XU Mo, SUO Zhi-hai

(Network and Information Center of Xi'an Jiaotong University, Xi'an 710049, China)

Abstract: In order to analyze the user learning behaviors in university campus, a fast and extensible development framework is designed and implemented that combined with JFinal lightweight framework and ECharts visualization technology. Based on the university e-learning platform, a typical analysis theme was chosen as the development background, the key technologies such as algorithms model, front-end data visualization, database model and back-end data processing were studied and implemented in detail. The result of the implementation shows that the framework developed in this paper can support various visualization requirements of the data analysis in the e-learning platform. Meanwhile, the future constructions were put forward for the e-learning behavior analysis in big data environment.

Key words: JFinal; ECharts; JSON; RESTful; data visualization; e-learning; big data; behavior analysis

1 引言

大数据环境下,越来越多的高校开始构建在线学习平台,为校内外用户提供各类线上学习资源并对用户的学习行为进行多指标诊断、对学习资源进行多指标评价^[1].如基于用户的网络学习登录行为与资源浏览行为,分析影响学生学习的各种因素,从而对教学内容和教学模式进行优化升级;基于多维度与多层次模型对学习者的行为进行智能化行为评价^[2];通过 MOOC 及社交平台挖掘学生感兴趣的信

息^[3]等.

与此类似的研究工作还有很多,大部分都集中在大数据算法及数据分析处理算法方面.但是从系统开发者角度,研究一款具有快速构建数据可视化模块、轻量级后端数据交互、支持数据多级下钻的开发框架涉及较少.同时便捷快速的数据可视化开发框架可以帮助统计分析系统开发者仅通过少量代码配置即可实现网页前端的统计分析数据集可视化,增强分析数据的可读性,为学校管理者提供直观、易理解的辅助决策支持.

因此,本文以西安交通大学在线学习平台的用户学习行为分析为实例,以构建数据分析快速可视化开发框架为研究对象,设计并实现一种可快速开发、支持轻量级数据库交互、支持多维度且数据可下钻的数据可视化开发框架。

2 框架设计及工作机制

2.1 框架前端 ECharts 工作机制

ECharts 是百度公司的一个开源数据可视化工具^[4],其纯 Javascript 的图表库,可以流畅的运行在 PC 和移动设备上,兼容当前绝大部分浏览器(IE8/9/10/11,Chrome,Firefox,Safari 等),底层依赖轻量级的 Canvas 类库 ZRender,提供直观生动、可交互、可高度个性化定制的可视化图表^[5]。同时支持深度交互式数据探索,大数据量展现、多维数据可视化关联、动态数据实时图形化更新等(实例见文章第 4 部分)。

ECharts 工作机制如图 1 所示。用户通过前端浏览器向后端服务器发出访问请求时,EChartsJS 文件依据配置发起资源访问请求。后台服务器接受请求,计算后输出 ECharts 统计图的 JSON 数据,之后 EChartsJS 调用 ECharts 脚本进行图形绘制。

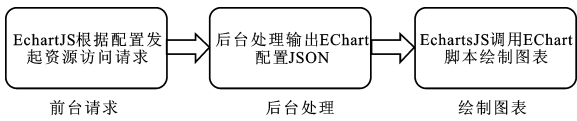


图 1 ECharts 工作机制

2.2 后端 JFinal 工作机制

JFinal 框架的逻辑架构主要采用典型的 MVC 架构^[6],是基于 Java 的极速 Web 开发框架,开发迅速、代码量少、框架轻量级且易扩展并支持 RESTful 数据格式,其工作机制如图 2 所示。

JFinal 以插件(Plugin)为基础,依据开发需求逐步配置各类处理器(handler)、拦截器(inteceptor);通过控制器(controller)和渲染器(Render)向框架上层提供数据展示,提升开发框架的可扩展性配置、保证系统功能模块可配置化、简便化、提升程序员开发效率高(简洁化开发见文章第 3 部分)。

2.3 RESTful 原则

RESTful 是符合 REST 原则(Representational State Transfer,表现层状态转化)的架构^[7]。其关键是定义可表示流程元素与资源的对象,且每个对象都是通过 URL 表示。此种方式的优点是可以充分利用 http 协议本身语义调用各类资源(接口、访问)

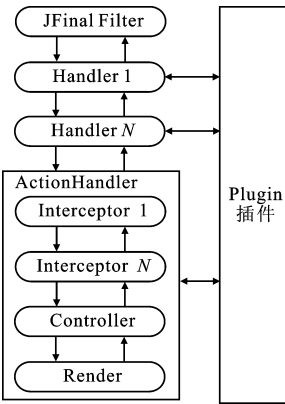


图 2 JFinal 工作机制

等,且不需要考虑上下文,降低资源访问复杂度。同时又与各类中间件无关,扩展性好和可移植性好。本文采用所有实验数据和演示数据,均遵循 RESTful 原则,且以 JSON 格式进行封装,供前端 ECharts 调用。

3 用户最喜爱课程三维指标分析

以高校在线学习平台中的用户最喜爱课程分析为例,其分析结果的可视化过程分为三个部分,即:原始数据抽取与计算、计算结果入库待展示、数据格式化后经前端数据调用,最后进行页面图形构建。

因此,用来验证快速可视化开发框架的实例—用户最喜爱课程分析—的多维指标计算方法在 2.1 节中进行了阐述;算法依赖的数据集在 2.2 节中进行了数据库建模;其最终利用快速开发框架进行可视化实现的过程在文章第三部分进行重点阐述与实现。

3.1 最喜爱课程指标模型设计

本文用来验证快速开发框架数据可视化工作的实例支持多维度指标联合分析。其基础指标计算模型设计如下:

$$T_i = \sum_{j=1}^n \alpha_j T_{ij} \tag{1}$$

模型中的 T_i 为某一门课程所有指标的总得分。例如 T_{i1} 为该课程学习热度总得分, T_{i2} 为该课程点赞热度总得分, T_{i3} 为该课程大众好评总分数, T_{ij} 为该课程按其他评价指标所获得的总分数。

模型中 α 为评价体系指标,例如 α_1 为某课程获得的学习热度得分权重, α_2 为该课程获得的点赞热度得分权重, α_3 为该课程获得的大众好评得分权重, α_i 为其他评价指标得分权重,且 $\sum_{i=1}^n \alpha = 1$ 。

```
public void configPlugin(Plugins me) {
    //配置 C3p0 数据库连接池插件
    DruidPlugin druidPlugin = createDruidPlugin();
    me.add(druidPlugin);

    //配置 ActiveRecord 插件
    ActiveRecordPlugin arp = new ActiveRecordPlugin(
(druidPlugin);

    //所有映射在 MappingKit 中自动化处理
    _MappingKit.mapping(arp);
    me.add(arp);
}

/* * * 配置全局拦截器 */
public void configInterceptor(Interceptors me) {}
```

```
/* * * 配置处理器 */
public void configHandler(Handlers me) {
    me.add(new ContextPathHandler("BASE_PATH"));
}
```

4.2 JFinal 数据资源 JSON 格式化

基于 JFinal 的数据资源 JSON 格式化其工作目的是将来自于后端服务器的各类数据分析统计子集,格式化符合 RESTful 规则的数据串,保证前端负责数据可视化的 ECharts 插件进行数据调用。

以用户最喜爱的课程资源分析为例(图 4),该模块的数据可视化由三个部分组成,第一部分是课程资源名称数据集的格式化,包含课程名、任课老师的姓名等(其实现见 3.2.1 节);第二部分是统计指标的数据集可视化,包含每个课程资源的点击量分数、好评量分数、点赞量分数等(其实现见 3.2.2 节);第三部分是前端 ECharts 插件的 JSON 数据调用与统计图的构图展示,保证数据快速可视化,提升数据分析结果的可读性(实现过程见 3.3 节)。

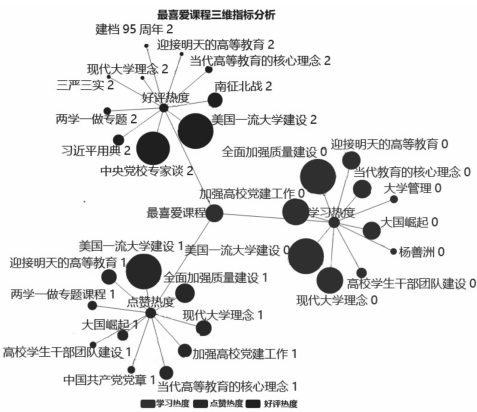


图 4 用户最喜爱课程三维指标数据可视化效果图

4.2.1 课程资源数据格式化

本实验的可视化实例其课程资源信息有课程名称及任课教师姓名等。该部分数据集均来自于在线学习平台后端基础数据库,因此需要从数据库中把此部分数据集读取出来,格式化前端 ECharts 可读的 JSON 数据。因为 JFinal 支持轻量级数据库交互,因此只需要开发人员在代码中简单配置数据集的表格信息(SQL 语句),然后调用 JsonKit.toJson 将结果进行格式化输出即可。实现过程片段如下所示。

```
public void fomartJson() {
    String sql="select a. *,b. courseTeacher from "
    +"favoritelesson a, wky_course b where "
    +" a. ccID=b. ccID and a. qishu=1' order by "
```

```
+" total desc limit 50";
List<FavoriteLesson> flist = FavoriteLesson.ft.dao().find(sql);
int n=flist.size();
this.setAttr("flist", flist);

String[] legend=new String[10];
String[] leg=new String[n];

List<Map> mplist=new ArrayList<Map>();
for(int i=0;i<10;i++){
    legend[i]=flist.get(i).getCcName();//课程名称
    Map<String, Object> map = new HashMap<String, Object>();
    List<Object> list = new ArrayList<Object>();
    List<Integer> valueList = new ArrayList<Integer>();

    valueList.add(flist.get(i).getDjcount());
    valueList.add(flist.get(i).getHpcount());
    valueList.add(flist.get(i).getDzcount());

    map.put("name", flist.get(i).getCcName());
    map.put("value", valueList);
    mplist.add(map);
}

//课程名称 json 格式化
this.setAttr("legend", JsonKit.toJson(legend));
//数据值 json 格式化
this.setAttr("data", JsonKit.toJson(mplist));
}
```

4.2.2 计算指标 JSON 格式化实现

本文实验的计算实例是用户最喜爱课程分析,其计算指标有点击量分数,好评量分数及点赞量分数。为支持前端 ECharts 构图,需将各课程资源的点击量分数数据集、好评量分数数据集、点赞量分数数据集进行 JSON 格式化,保证可视化图形的快速可变性及实时性。例如把点击量分数数据集暂存入 map3,点赞量分数数据集暂存入 map4,好评量暂存入 map5,之后将其他所需的数据集一起存入 List 中,调用 JsonKit.toJson 进行 JSON 格式化,其代码片段如下所示。

```
//实现点击量分数所有数据格式化
Map<String, Object> map3 = new HashMap<String, Object>();
map3.put("name", "学习热度");
map3.put("symbolSize", degree3);
map3.put("category", "学习热度");
```

```
map3.put("draggable", "true");
map3.put("value", deg3);
mplist.add(map3);
//开始点赞量分数数据格式化
Map<String, Object> map4 = new HashMap<String, Object>();
map4.put("name", "点赞热度");
map4.put("symbolSize", "degree4");
map4.put("category", "点赞热度");
map4.put("draggable", "true");
map4.put("value", deg4);
mplist.add(map4);
//开始好评量分数数据格式化
Map<String, Object> map5 = new HashMap<String, Object>();
map5.put("name", "好评热度");
map5.put("symbolSize", "degree5");
map5.put("category", "好评热度");
map5.put("draggable", "true");
map5.put("value", deg5);
mplist.add(map5);
//中间代码省略
this.setAttr("djson", JsonKit.toJson(mplist));//
this.setAttr("links", JsonKit.toJson(links));//
```

4.2.3 ECharts 前端数据可视化构图

基于 3.1 和 3.2 的数据集 JSON 格式化,当系统管理员在网页前端触发用户最喜爱课程分析动作时,ECharts 接收系统后端传输过来的 JSON 数据串,然后进行统计结果构图.其结果集的构图过程如下所示:

(1)系统前端绘图时先引入 Echarts 插件,并在页面中定义图像 DOM 容器大小:

```
<script src="#"(BASE_PATH)/js/echarts.min.js"></script>
<script src="#"(BASE_PATH)/js/echarts.js"></script>
<div id="main" style="width: 500px;height:400px; align="center">
</div>
```

(2)在容器中定义图像实体,准备好 dom 并初始化 ECharts 实例,其省略细节后的主体代码如下:

此外,前端可视化图形的数据逐级下钻展示方面,基于 JFinal 与 ECharts 组合构建的数据可视化快速开发框架,对此亦有较好地支持.例如分析全校参与在线学习的用户群体分布、学院分布、学习积分

```
<script type="text/javascript">
//于准备 dom,初始化 echarts 实例
var myChart =
echarts.init(document.getElementById('main'));
option = {//此处代码省略};
//使用指定的配置项和数据显示图表
myChart.setOption(option);
</script>
(3)加载课程资源及其统计分析数据集(JSON 格式),其实现过程的部分代码如下:
series: [{
name: '最喜爱课程',
type: 'graph',
layout: 'force',
force: { repulsion: 300 },
data: # (djson), links: # (links),
categories: [{ 'name': '学习热度' },
{ 'name': '点赞热度' },
{ 'name': '好评热度' },
],
//其他构图配置信息省略
}]
```

经过以上 JFinal 数据库配置,用户最喜爱课程统计数据集在 JFinal 完成 JSON 格式化后,ECharts 插件于开发框架前端完成最喜爱课程三维指标可视化图形构图(如图 4 所示,《美国一流大学建设与高等教育改革的思考与借鉴》是最受用户喜爱的课程之一,且其在学习热度、好评热度及点赞热度方面均排在榜首).

5 应用效果

本文设计实现的基于 JFinal+ECharts 的数据分析可视化快速开发框架已经在我校在线学习平台中进行使用.如图 4 所示,该图基于 ECharts 关系散点图,直观展示了用户最喜爱课程的三维指标分析(学习热度分数、好评热度分数、点赞热度分数),且数据分析结果依据数据指标进行分类展示.

(圆点面积越大,表明该课程资源在此指标上的数值也越大).当后端数据分析子集发生变化时,前端可视化图形可依据后端数据子集而随之变化.

与人数联合分析,其数据可视化和数据下钻可视化效果如图 5 所示.

图 5(a)实例中,一级指标分析并显示了全体学习用户在五类群体中的分布(男教工、男学生、女教

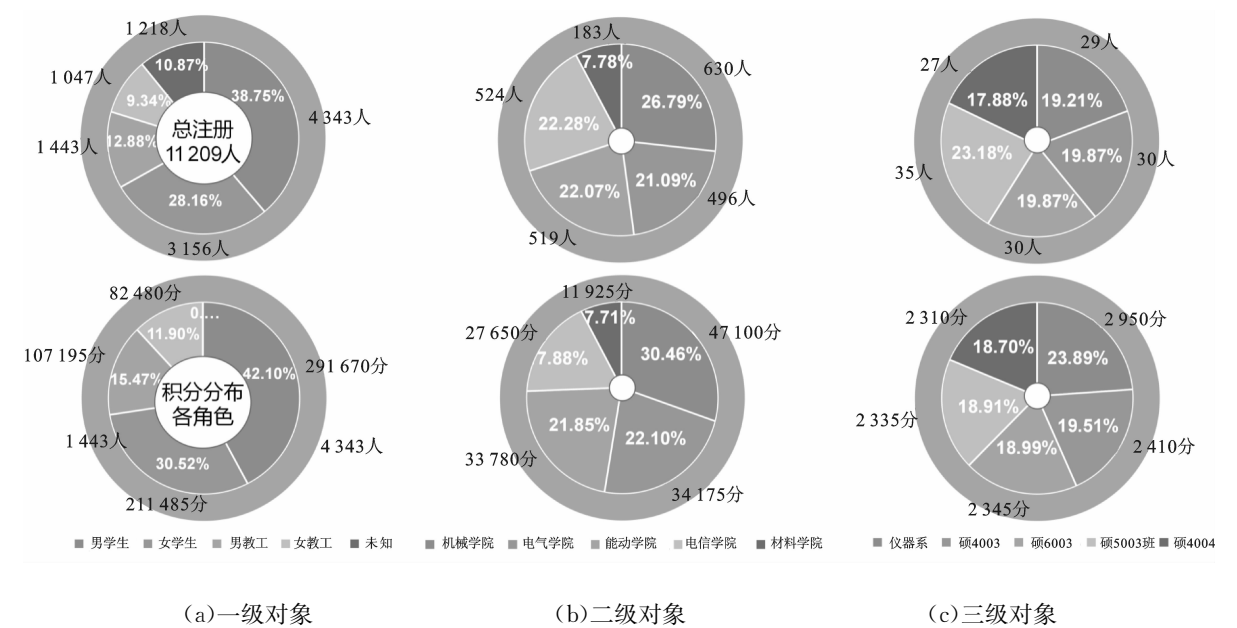


图 5 用户群体学习行为数据逐级下钻可视化过程

工、女学生、未知)及各群体的总体学习情况分析(学习人数及获得的学习积分联合分析).图 5(a)上面饼图中心点为全体学员人数,其外层彩色比例环形为各用户体群参与此次学习的人数比例,最外层比例环为各用户群体参与本次在线学习的学员人数.下面饼图内部第一层比例环显示每个用户群体学习后获取的学习积分百分比,外环是各用户群体参与学习的具体人数.且所有两个饼图可实时联动,激活其中某一个用户群体的图例后所有饼图同时变化.

图 5(b)实例显示了激活图 5(a)中的男学生群体图例后,该用户群体统计数据开始下钻,跳转并展示二级对象(男学生群体活跃度 Top5 的学院分布),包含此 Top5 学院男学生群体学习人数及其所获得的积分展示.上面饼图显示的是 TOP5 学院男学员学习人数及其比例,下面饼图显示的是 TOP5 学院男学员学习积分总数及比例.

图 5(c)实例显示了激活图 5(b)中的 Top1 学院男学生群体后,数据对象进一步细化与下钻,展示该学院男学生学习活跃度 Top5 的班级分布(含各班级男学生学习人数及其获取的学习积分).上下饼图含义与图 5(b)相同.

实现过程证明了采用 JFinal 与 ECharts 进行组合构建并搭建数据可视化快速开发框架可以较好地支持高校在线学习平台中的数据关联可视化分析及数据对象逐级下钻关联分析.所展示的可视化数据可以从一级对象(全体学员)逐级下钻至二级对象(某用户群体),再下钻至三级对象(该类用户群体所

在单位).只要后台计算数据库的统计数据子集粒度足够精细,其统计数据的可展示精度就越高,且各指标之间可关联构图.

6 结束语

经过以上的设计、开发及应用验证表明,采用 JFinal 与 ECharts 组合构建的在线学习行为分析数据可视化快速开发框架在轻量级快捷开发方面具有较好的可配置性、可扩展性,可同步提升程序员的开发效率.基于 JFinal 插件技术的后端数据库交互操作减少了前后端数据库交互开发成本;基于 ECharts 的前端统计数据可视化增强了统计数据的可读性、直观性,为高校在线学习平台的用户行为分析提供数据可视化开发框架,对辅助决策具有直观的展示作用.

由于各类数据统计分析系统的建设核心是后端用户行为的集成、清洗、交换及基于各类数据挖掘算法的实时处理技术.因此在线学习行为数据分析可视化研究工作的未来重点将集中于后端数据算法的开发与实现,提升系统核心统计分析价值.例如学习模型研究、行为预测研究等^[10].

参考文献:

- [1] Wang J, Lv H Y, Cao B, et al. Application of educational data mining on analysis of students' online learning behavior [C]// International Conference on Image, Vision and Computing, China, Chengdu, IEEE, 2017.
- [2] 彭文辉,杨宗凯,黄克斌.网络学习行为分析及其模

- 型研究[J]. 中国电化教育, 2006, 26(10): 31-35.
- [3] Wang J, Lv H Y, Cao B, et al. Application of educational data mining on analysis of students' online learning behavior [C]// International Conference on Image, Vision and Computing. China, Chengdu, IEEE, 2017.
- [4] 郑幸源, 洪亲, 蔡坚勇, 等. 基于 AJAX 异步传输技术与 Echarts3 技术的动态数据绘图实现[J]. 软件导刊, 2017, 16(3): 143-145.
- [5] Echarts [EB/OL]. [2017-10-09]. <http://echarts.baidu.com/feature.html>.
- [6] 韩玉焱. Java Web 的 JFinal 开发框架研究[J]. 电子世界, 2017, 38(12): 153-153.
- [7] Safaripour R, Khendek F, Glitho R, et al. A REST-full architecture for enabling rapid development and deployment of companion robot applications [C]// International Conference on Computing, Networking and Communications. Hawaii, USA, 2014: 971-976.
- [8] 耿秀丽, 马万元. 考虑未知属性权重的区间直觉模糊 VIKOR 方法 [J]. 计算机工程与应用, 2016, 52(12): 1-5.
- [9] 董雪琦, 耿秀丽, 徐士东. 灰色关联分析与云模型集成的方案评价方法 [EB/OL]. [2017-07-21]. <http://www.aocmag.com/article/02-2018-08-001.html>.
- [10] Kinnebrew J, Segedy J, Biswas G. Integrating model-driven and data-driven techniques for analyzing learning behaviors in open-ended learning environments [J]. IEEE Transactions on Learning Technologies, 2017, 10(2): 140-153.

作者简介:

洪丹丹 女, (1981-), 硕士, 工程师. 研究方向为移动互联网技术、数据分析与共享技术、校园信息化关键技术.

E-mail: dddhong@xjtu.edu.cn.

李飞 男, (1981-), 工程师. 研究方向为大数据分析、系统架构设计等.

姚磊 男, (1982-), 硕士, 工程师. 研究方向为教育大数据分析 & 校园信息化关键技术.

徐墨 女, (1978-), 硕士. 研究方向为校园信息化.

锁志海 男, (1971-), 博士研究生. 研究方向为数据库应用技术、高校教育信息化.

(上接第 5 页)

- [8] Yang X S, Deb S. Cuckoo search via levy flights [C]// Proceedings of World Congress nature & Biologically Inspired Computing. India, IEEE Publications, 2009: 210-214.
- [9] 吴虎胜, 张凤鸣, 吴庐山. 一种新的群体智能算法—狼群算法 [J]. 系统工程与电子技术, 2013, 34(11): 2430-2438.
- [10] 吴虎胜, 张凤鸣, 战仁军, 等. 求解 0-1 背包问题的二进制狼群算法 [J]. 系统工程与电子技术, 2014, 35(8): 1660-1667.
- [11] 陈汉武, 李科, 赵生妹. 基于相位匹配的量子行走搜索算法及电路实现 [J]. 物理学报, 2015, 82(24): 29-39.
- [12] 李国亮. 狼群算法的研究与应用 [D]. 江西: 东华理工大学, 2016.
- [13] 李娟, 方平, 周明. 一种求解背包问题的混合遗传算法 [J]. 南昌航空工业学院学报, 1998, 11(3): 35-39.

作者简介:

严雅榕 女, (1994-), 硕士研究生. 研究方向为管理信息与决策支持. E-mail: ljf653483717@163.com.

项华春 男, (1980-), 博士, 副教授. 研究方向为装备可靠性与系统工程.

聂飞 男, (1970-), 博士, 副教授. 研究方向为管理信息与决策支持.

李京峰 男, (1993-), 硕士研究生. 研究方向为装备发展战略与管理决策.