# 英雄联盟大数据分析平台的构建案例教学指导手册

**教学适用的课程：**《大数据架构与设计》

**教学适用的专业：**电子信息

**教学目的与用途：**了解大数据分析在游戏产业中的应用；掌握大数据预处理、数据分析相关知识点；掌握常用机器学习算法。

**教学内容：**

（1）理论依据：本案例以当下热门游戏“英雄联盟”为蓝本，构建英雄联盟大数据分析平台。使用爬取于英雄联盟比赛网站上的竞赛数据构建数据集，结合数据挖掘、机器学习等相关方法进行数据可视化分析进一步实现游戏胜负预测及阵容推荐等功能。

（2）涉及知识点：分布式爬虫、Hadoop 、Spark、数据可视化、逻辑回归、支持向量机、朴素贝叶斯、决策树。

（3）分析路径：首先明确任务目标：以当下热门游戏“英雄联盟”为蓝本，构建英雄联盟大数据分析平台。使用Hadoop 和Spark搭建服务平台，再利用网络爬虫获取游戏比赛数据进行存储和分析，最终进行数据可视化分析以及使用机器学习算法如逻辑回归、支持向量机、朴素贝叶斯和决策树等方法实现游戏胜负预测与阵容推荐功能。

**启发思考题：**

游戏胜负预测除了使用常用的机器学习算法如逻辑回归、支持向量机、朴素贝叶斯、决策树等之外，还能想到其他的方法吗？（参考答案：深度学习算法如卷积神经网络、循环神经网络等）

**建议课堂计划：**

1. 时间安排：3课时
2. 学生学习准备：了解Hadoop 、Spark、网络爬虫以及常用机器学习算法如逻辑回归、支持向量机、朴素贝叶斯、决策树等。
3. 分组及讨论内容：分组讨论启发思考题。

案例开场白：近些年来游戏产业发展十分迅速，传统的电子竞技市场也正被新兴的IT技术所推动。从调查数据来看，如今的游戏玩家数量相比于几年前的传统电竞时代已增长数倍。随着当前游戏团队合作化模式的不断加深，使得很多游戏元素都出现了竞技化趋势，并且竞技强度和节奏都在逐步提高。一方面，游戏数据分析可以提升游戏竞技人员的技术。另一方面，有利于游戏公司进行游戏平衡性的改进。在挖掘出不平衡的角色之后，可以对单一数值的不同修改幅度进行情景分析。模拟不同情境下该修改对角色平衡性的影响，以此确定最佳的修改范围，作为角色技能修改的参考依据。随着云计算、大数据等技术在生活中的不断渗透，游戏产业也逐渐成为这些新兴技术的“沃土”。因此，我们有必要学习大数据分析在游戏产业中的应用。

1. 结束总结：这节课我们以当下热门游戏“英雄联盟”为蓝本，构建英雄联盟大数据分析平台。使用爬取于英雄联盟比赛网站上的竞赛数据构建数据集，结合数据挖掘、机器学习等相关方法进行数据可视化分析进一步实现游戏胜负预测及阵容推荐等功能。我们通过平时课程所学知识实现了英雄联盟大数据分析平台，课后感兴趣的同学可以搜集相关资料，进一步了解大数据在其他场景的应用，相信你们会收获更多。
2. 案例引导建议：在教师简要分析应用场景后，鼓励学生自主考虑如何使用所学知识解决该场景问题。

**参考文献：**

[1] 周志华. 机器学习[M]. 北京: 清华大学出版社, 2016.

[2] Goodfellow I, Bengio Y, Courville A. Deep learning[M]. MIT press, 2016.

**其他教学支持材料：**

1. 本案例提供配套的PPT、视频、数据集与代码等，发布于Github，链接为：https://github.com/Wanghui-Huang/CQU\_bigdata。

2. 本案例涉及到数据预处理以及多种机器学习算法，建议使用python语言进行编写，推荐的工具包有requests（网络请求库），BeautifulSoup（HTML解析库），scikit-learn（机器学习库）。