暑期企业实训个人总结

--林岫铭 0171123396

时间飞逝，光阴如梭，短短的三周过去了。通过这次大数据实训，我学到了很多有关大数据的知识。在这开始的十几天中，老师讲了很多关于大数据技术的理论知识。在老师的教学与文档的帮助下，我先后安装好了CentOS虚拟机以及hadoop，hive等工具。同时也学会了MapReduce工作的基本流程，包括从HDFS下载文件、向HDFS上传文件、利用hadoop进行简单的文本处理和数据分析等操作。

之后我们开始进行大数据实例开发，我们组选择了分析1950-2014年历届世界杯比赛数据。我负责了数据收集和清洗以及MapReduce的编写进行数据分析，通过MapReduce工具，成功分析出了一系列数据。同时，我们组共同学习可视化编程，最终也将分析过的结果成功的可视化，达到了预期的可视化目的，并且得到了分析后的结论。

通过三周短暂的实训，我对大数据开发的基本流程有了更加清晰的认识，了解了进行一个完整的大数据项目开发需要用到的工具以及需要做的工作，第一次接触了Hadoop，Map Reduce等工具，学习了前端开发等相关技术的知识并加以使用，对以后的大数据学习打下了坚实的基础，也树立了更加明确的目标。

在团队合作方面，我深刻的感受到了团队紧密合作的重要性，平日里多关心组员、团结一体、互相学习。工作中我们是彼此的老师互相学习。面对问题我们共同解决。从我个人而言，在工作的过程中受益匪浅。虽然这次实训只是一个小项目，但对于第一次接触大数据的我们来说，想要在一周多的时间里完成是不可能的，正是因为有了组员之间紧密的协作、相互的沟通，才能使我们更好的、更快的、更高质量的完成这次实训项目！在今后的学习中，提高认识，努力学习新知识，增强自身的技术水平。

这次实训中，我收获了很多，学习了的若干的新知识：

1、掌握了大数据计算平台相关系统的安装部署，包括CentOS7虚拟机、Hadoop安装、Hive的安装。

2、 了解并掌握了数据分析之前的数据清洗的主要目的及基本流程。

3、理解了大数据MapReduce计算模型，掌握了简单的MapReduce程序开发。

4、学习前端页面的开发，了解Echars、D3的工作流程。

这次开发过程中，也发现了一些不足之处：

1.开发过程中没有对数据进行太过复杂的分析，也没有实现预测功能。

2.分析数据过程中没有将收集的数据全部利用。

3.审美不行，导致可视化不够美观。

4.技术能力不足无法达到预想的结果。