数据仓库与大数据工程课程实验一

**实验报告提交时间要求： 截止日期：4月13日**

1.1 实验目的

在大数据时代，存在很多开源的分布式数据采集、计算、存储技术，本实验将熟悉并搭建几种常用的大数据采集、处理分析技术环境。

1.2 实验环境

学院为每名研究生提供的公用计算与存储资源等。

1.3 实验内容

1. 熟悉并搭建Hadoop（HDFS、MapReduce、HBase、Hive）或Cassendra、Spark、Storm、Kafka、MySQL、ZooKeeper等开发环境，撰写各开发环境的搭建技术文档报告；
2. 了解Hadoop（HDFS、MapReduce、HBase、Hive）、Cassendra、Spark、Storm、Kafka、MySQL、ZooKeeper等技术的基本原理与架构，同时尝试对其进行简单操作并撰写操作文档报告。