大数据学习规划

第一阶段：掌握数据挖掘的基本概念和方法。先对数据挖掘有一个概念的认识，并掌握基本的算法，如分类算法、聚类算法、协同过滤算法等。  
 参考书：《数据挖掘概念和技术》（第三版）范明，孟小峰 译著。  
 第二阶段：掌握大数据时代下的数据挖掘和分布式处理算法。现在已经进入大数据时代，传统的数据挖掘算法已经不适用于  
 参考书：《大数据：互联网大规模数据挖掘和分布式处理》 王斌 译著。  
 第三阶段：使用Hadoop进行大数据挖掘。Hadoop里面有一个Mahout组件，几乎包括了所有的数据挖掘算法，包括分类、聚类、关联规则等。  
 参考书：Hadoop实战（第二版）.陆嘉恒 著。  
 另外，数据挖掘是数据库技术、人工智能技术、机器学习技术、统计学习理论、数据可视化等一系列技术的综合，所以，要想学好数据挖掘，这些技术也得懂的呀。  
 推荐入门时先看浙江大学王灿老师的数据挖掘课程，网上搜下。