**Getting Started with Hadoop 2.x**

**Section 1: Intro to the Hadoop World**

**1.1 The Course Overview**

This video gives an overview of the entire course.

**1.2 Installing Hadoop in Local**

In this video we’ll learn how to install Hadoop on our local system.

* Prerequisites to the Hadoop installation
* Hadoop Installation
* Testing our Hadoop installation

**1.3 Bring Process to Data**

The important part of selecting the Hadoop framework for your own solution is to understand why it is a good fit for your application.

* Understand the Hadoop way of execution
* Fit your own application and see if it will benefit from it

**1.4 NameNode Versus DataNode**

Understand the difference between the two nodes in HDFS; Datanode and Namenode

* Understand data resiliency
* Distributed data source

**1.5 Map and Reduce Operations**

The new term Map-Reduce… what does it mean and how does it solve a problem?

* Understand what map-reduce operation is
* Dig deep with different parts of map-reduce in the Hadoop world
* Devise a way to implement your own problem with a map-reduce operation

**1.6 Order of Execution and Parallel Thinking**

When jumping to parallel programming from serial programming, it is always hard to plan the computation.

* Figure out the pitfalls of parallel data
* See how the Hadoop parallel process works
* Start parallel thinking

**Section 2: File System Overdrive with HDFS**

**2.1 Formatting a HDFS**

Prepare your HDD for with HDFS

* Have the HDFS
* Fit your own application and see if it will benefit with it

**2.2 Formatting a HDFS**

Copy data to/from HDFS.

* Copy data to HDFS
* Copy data from HDFS

**2.3 Some Helpful Commands to Communicate with the HDFS**

Using the HDFS commands in the shell.

* Find the basic difference between generic shell command and HDFS command
* Get used to the different ways of writing the commands

**2.4 HDFS Protocol and Using It in Applications**

How do we access the HDFS files from a java program

* Connect to the file system HDFS protocol
* Fetch data from HDFS
* Put data in HDFS

**Section 3: Let's Run Some Hadoop Jobs**

**3.1 Hadoop Jobs Versus Tasks**

What are Hadoop jobs and tasks?

* Understand how tasks are communicated
* Understand how jobs are run

**3.2 The Hadoop UI for Task Progress**

How to see the process flow and progress of a Hadoop job.

* Open the Hadoop UI
* Assess the memory and process

**3.3 Running a Couple of Example Jobs**

Run Hadoop jobs.

* Run Hadoop jobs directly
* Run Hadoop jobs on yarn

**3.4 Analyze the Work Flow/Data Flow/Process Flow**

In this video, we are going to look at how the map and reduce gets executed

* Get to know about the data flow
* Discover how map gets executed
* Discover how reduce gets executed

**Section 4: It's Show Time**

**4.1 Introduction to the Movie Dataset**

Understand the dataset provided by the grouplens.org

* How the data is structured with movies
* How the data is structured with the ratings

**4.2 Data Transformation and Storing to HDFS**

Prepare the data to be fit for our algorithm

* Split the data to be transformed
* Transform the data using Hadoop
* Merge the data using a basic java application

**4.3 Devise a Simple Algorithm for Recommendation**

Devise a simple algorithm for recommendation

* Create an algorithm to prepare data for recommendation by genre
* Understand the data format for that output

**4.4 Implement the Algorithm in Hadoop Map-Reduce Way and Analyze Performance**

Implement the map-reduce for the transformation of the movie -> genre context

* Create a map-reduce job for this problem
* Take different splits for different performance assessment