## Mobile Devices and Big Data

## Assignment 4

## (Single Node Hadoop – Matrix Multiplication – Stock Market State Prediction)

1. Download the stock closing price data over the last 1 year for any 10 companies from

<http://www.nasdaq.com/quotes/historical-quotes.aspx>

For every company, do the following {

1. For each of the 52 weeks, find the average closing price – heretofore referred to as closing price.
2. Populate a 3-state (Bull, Bear, Stagnant) transition probability matrix P (3 rows x 3 columns). The matrix should have transition probabilities of Bull to Bull (closing price of a Bull week going even higher), Bull to Bear (closing price of a Bull week going lower), Bull to Stagnant (closing price of a Bull week remaining same within a small threshold wrt to last week’s), Bear to Bull (closing price of a Bear week going higher), Bear to Bear (closing price of a Bear week going lower), Bear to Stagnant (closing price of a Bear week remaining same within a small threshold wrt to last week’s), Stagnant to Bull (closing price of a Stagnant week going higher), Stagnant to Bear (closing price of a Stagnant week going lower), and Stagnant to Stagnant (closing price of a Stagnant week remaining same within a small threshold wrt to last week’s) values. A Bull week occurs when the closing price of the week is higher than that of last week’s; a Bear week occurs when the closing price of the week is lower than that of last week’s; and a Stagnant week occurs when the closing price of the week is same as that of last week’s within a small threshold. The transition probability of an entry in the above matrix is = (number of times that particular transition occurs) / (51), where 51 is the number of weekly transitions in the year)

Steps (1) – (3) above should be implemented using Hadoop (design your own Map-Reduce methods).

1. Evaluate limit (N-> infinity) of P ^ N, where N is the number of matrix multiplications of P with itself. Find the value of N when the above limit converges or shows an oscillation pattern.

Step (4) above should be implemented using Hadoop as well.

}

1. Submit a report explaining the data in the log files for the Hadoop jobs above. The report should contain the values of N’s for the different companies as found in step (4) above.

**Deadline: 29th October 2015.**