A2 REPORT
Rahul Kizhakkepachilamakol rahulk@ccs.neu.edu
Rakesh Vishawanathan   viswanathan.r@husky.neu.edu

1) Executed job on EMR with variable number of core machines with sampling Rate as 20

Used 1.medium machines in EC2 (1 cpu 3.75 gb memory)

No	No of machines	Average Time to complete job in minutes
1	2	21
2	4	10
3	8	6
4	12	5

2) Running v5 with different sample rates on Mac i5 2 cores 8 GB RAM SSD

No	Sampling Rate %	Average Time	Average Error/Category %
1	50	4 mins 51 sec	0.088931
2	20	4mins, 14sec	0.721437
3	15	3mins, 54sec	2.29389
4	10	3mins, 51sec	18.9384

Below is the graph depicting the decrease in error with increase in sample rate.

The Y-axis is the Error and X-axis shows the sample rate.