|  |
| --- |
| **Online Learning** |
| **Definition**: continuous stream of data coming in |
| Train an algorithm with a training set  When a new example came up, update theta with that single example  Discard that example, never use it again  So that the algorithm adapt to a change of user preferences |

|  |
| --- |
| **Map Reduce & Data Parallelism – open source lib for map reduce: Hadoop** |
| Map Reduce: split data into j set to feed into different computers  Each computer train on the data assigned to compute derivative of J  Set to a master machine to use the cost to update theta |
| If a learning algorithm can be expressed as computing sums of functions over the training set. |