Word	Description
Machine Learning	Machine Learning refers to the techniques involved in dealing with vast data in the most intelligent fashion (by developing algorithms) to derive actionable insights. In these techniques, we expect the algorithms to learn by itself wiithout being explicitly programmed.
Mahout	Mahout is an open source project from Apache that is used for creating scalable machine learning algorithms. It implements popular machine learning techniques such as recommendation, classification, clustering. Features of Mahout: • Mahout offers a framework for doing data mining tasks on large volumes of data • Mahout lets applications to analyze large sets of data effectively and in quick time • It also offers distributed fitness function capabilities for evolutionary programming • It includes several MapReduce enabled clustering implementations such as k-means, fuzzy k-means, Dirichlet, and Mean-Shift
MapReduce	Hadoop MapReduce is a software framework for easily writing applications which process vast amounts of data (multi-terabyte data-sets) in-parallel on large clusters (thousands of nodes) of commodity hardware in a reliable, fault-tolerant manner. A MapReduce framework is usually composed of three operations: 1. **Map:** each worker node applies the map function to the local data, and writes the output to a temporary storage. A master node ensures that only one copy of redundant input data is processed. 2. **Shuffle:** worker nodes redistribute data based on the output keys (produced by the map function), such that all data belonging to one key is located on the same worker node. 3. **Reduce:** worker nodes now process each group of output data, per key, in parallel.