Word	Description
Hadoop	Hadoop is an open source distributed processing framework used when we have to deal with enormous data. It allows us to use parallel processing capability to handle big data. Here are some significant benefits of Hadoop: • Hadoop clusters work and keeps multiple copies to ensure reliability of data. A maximum of 4500 machines can be connected together using Hadoop • The whole process is broken down into pieces and executed in parallel, hence saving time. A maximum of 25 Petabyte (1 PB = 1000 TB) data can be processed using Hadoop • In case of a long query, Hadoop builds back up data-sets at every level. It also executes query on duplicate datasets to avoid process loss in case of individual failure. These steps makes Hadoop processing more precise and accurate • Queries in Hadoop are as simple as coding in any language. You just need to change the way of thinking around building a query to enable parallel processing
Hidden Markov Model	Hidden Markov Process is a Markov process in which the states are invisible or hidden, and the model developed to estimate these hidden states is known as the Hidden Markov Model (HMM). However, the output (data) dependent on the hidden states is visible. This output data generated by HMM gives some cue about the sequence of states. HMM are widely used for pattern recognition in speech recognition, part-of-speech tagging, handwriting recognition, and reinforcement learning.
Hierarchical Clustering	Hierarchical clustering, as the name suggests is an algorithm that builds hierarchy of clusters. This algorithm starts with all the data points assigned to a cluster of their own. Then two nearest clusters are merged into the same cluster. In the end, this algorithm terminates when there is only a single cluster left. The results of hierarchical clustering can be shown using dendrogram…