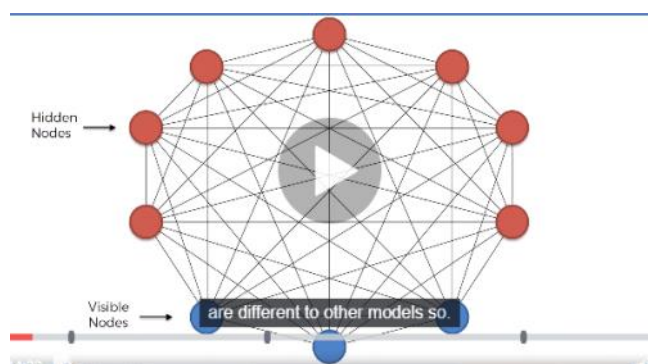


## Boltzmann Machine

- Used for recommendations Systems
- Boltzmann Machine are **undirected models** unlike ANN,CNN,RNN,SOM. They don't have direction so they don't have arrows in their connections and connections can go both ways.



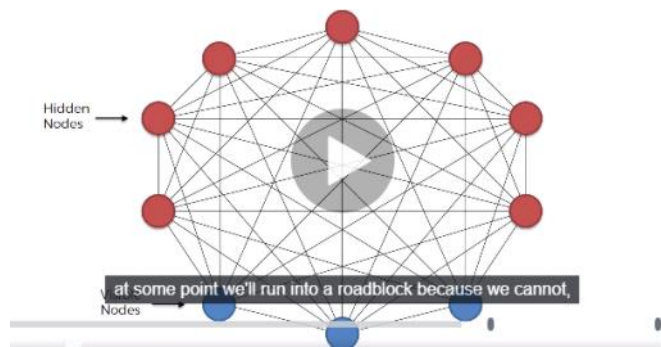
- There is **no output nodes** in Boltzmann machine as we don't output anything in Boltzmann machine. Through Boltzmann Machine we are creating a model that describes our system.
- **Visible nodes are connected to each other.** Boltzmann Machines are fundamentally different from other algorithms as they generate data, they generate information in all the nodes. Boltzmann machine generate its own parameters, it does not need any input. Or we can say it generates different state of the machines.
- Boltzmann Machine can be viewed as the representation of a certain system in which visible nodes are things which we can or do measure and hidden nodes are things we can't or don't measure.
- Boltzmann Machine is a generative deep learning model or stochastic deep learning model. It is not deterministic deep learning model.

## Energy-Based Model

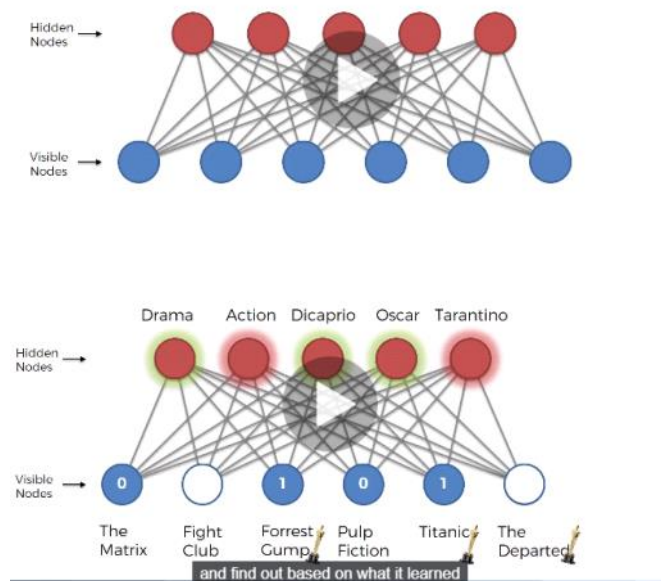
$$p_i = \frac{e^{-\epsilon_i/kT}}{\sum_{j=1}^M e^{-\epsilon_j/kT}}$$

The probability of system being in state 'i' is inversely proportional to the energy of that particular state. Boltzmann Machine works in similar way, the weights of neurons try to be in the lowest energy state in the system.

## Restricted Boltzmann Machine



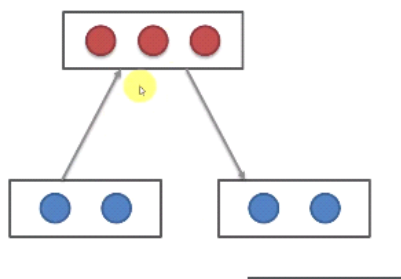
In Boltzmann Machine , every node is interconnected to every other node. But in Restricted Boltzmann Machine , hidden nodes are not connected to each other and visible nodes are also not connected to each other.



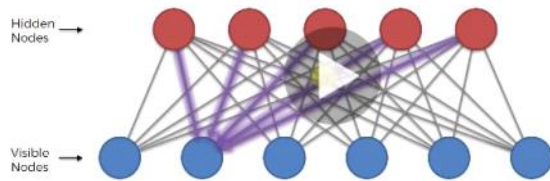
## Contrastive Divergence

This is the algorithm which allows Restricted Boltzmann Machine to learn.

Since RBM is an undirected model , question arises how the weights will be adjusted. Since in ANN and others, the stochastic gradient back propagates the error and weights are adjusted.

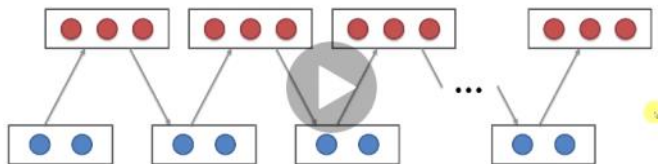


In RBM , the input nodes construct the hidden nodes and the hidden nodes reconstruct the input node. But the reconstructed input node is not similar to the initial input node.

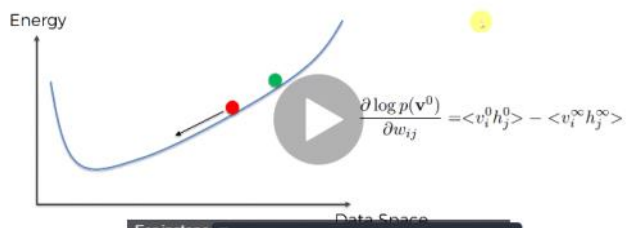


The reconstructed input nodes construct again the hidden nodes and the process continues. This process is called Gibbs Sampling.

This process continues till the time when the hidden nodes reconstructs back the same input node. At this time, the model converges.



In Contrastive Divergence, we are trying to adjust the energy curve by modifying the weights so that the system is in minimal energy state.



Using the Hinton's shortcut, we get to know the energy curve just by 2 passes of reconstructing the input nodes (CD 1).

## Deep Belief Networks

DBN is the stacked Restricted Boltzmann machine (RBM).

