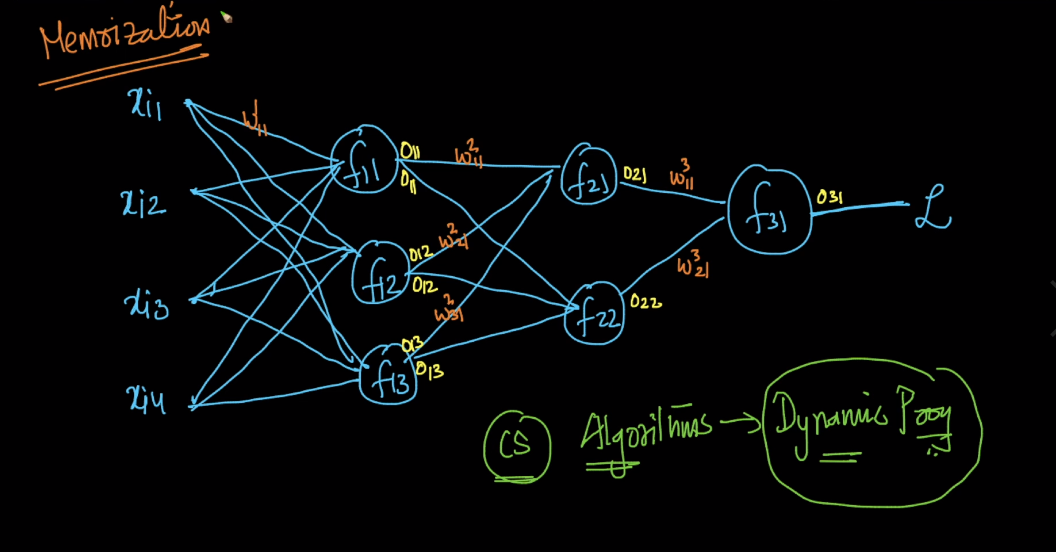
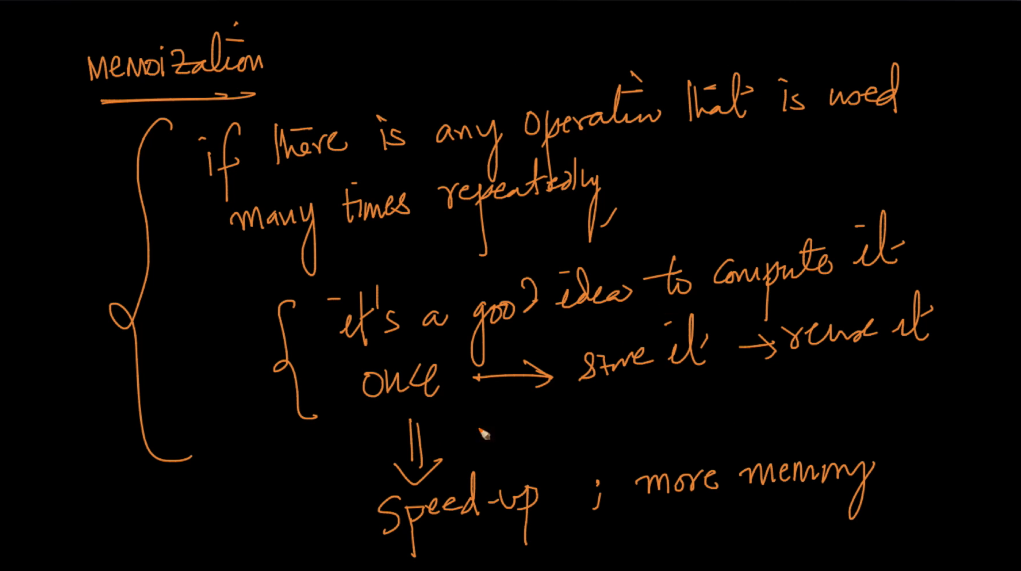
**Training an MLP:Memoization**

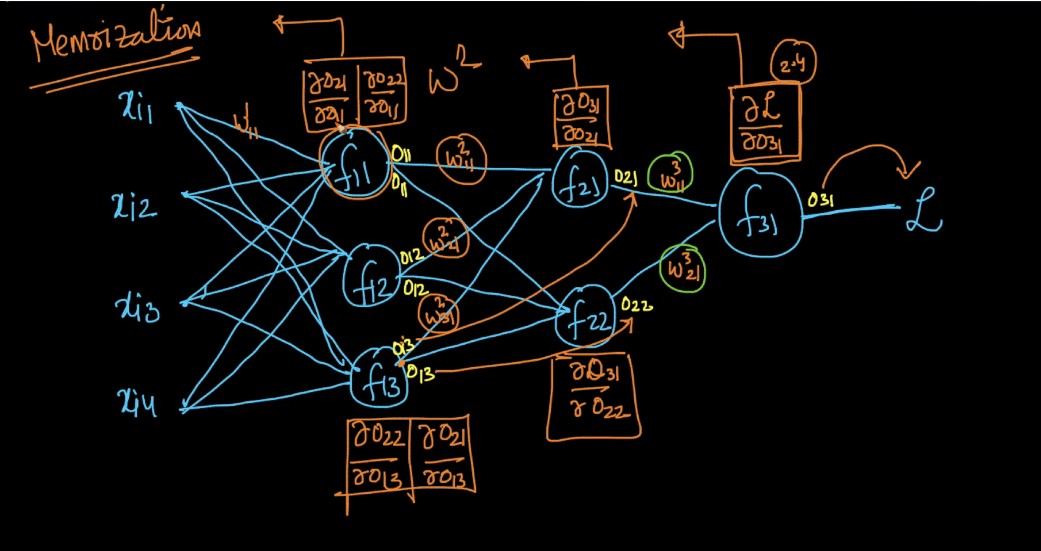
Here we use same memorization trick we use in dynamic programming.



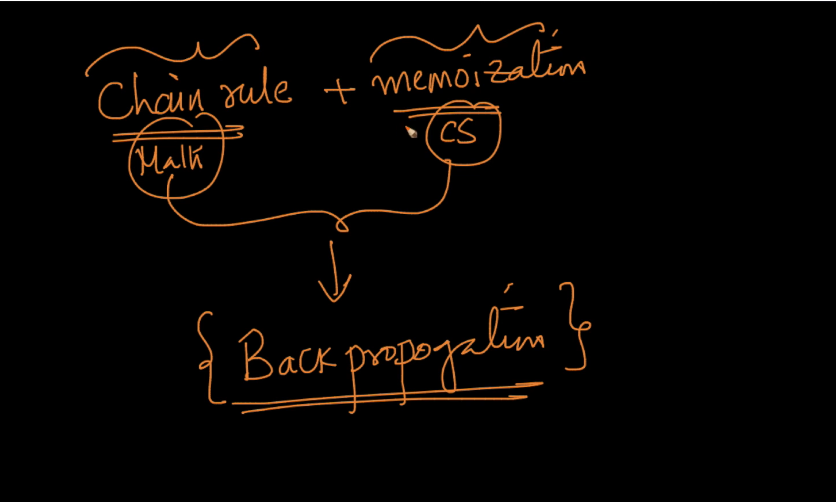


In neural network derivatives are used multiple times therefore we store it and use it again and by this speed-up program by giving slightly more memory.

So we store this derivatives at each node and use it again when it requires again.



So Back propagation is chain rule + memoization



Comments :

