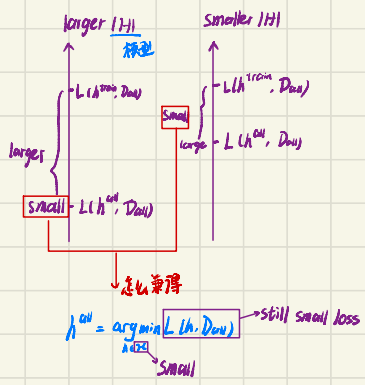
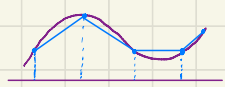




Why deep learning

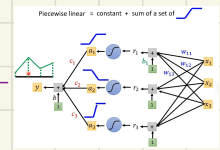


Review: why hidden layer?

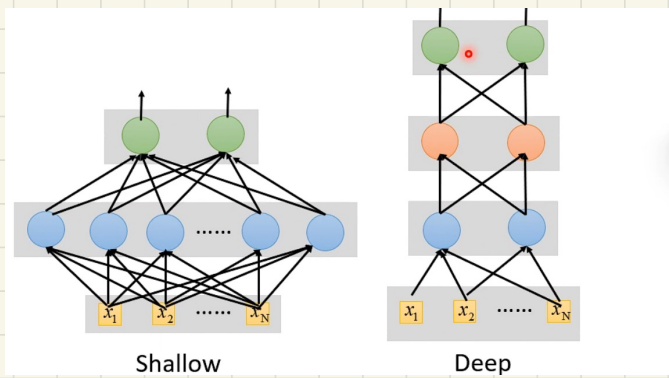


Constant + sum of a set of

- ① Sigmoid function
- ② ReLU

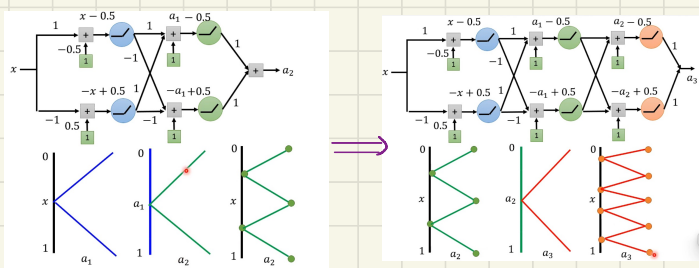


Fat + short v.s. thin + tall

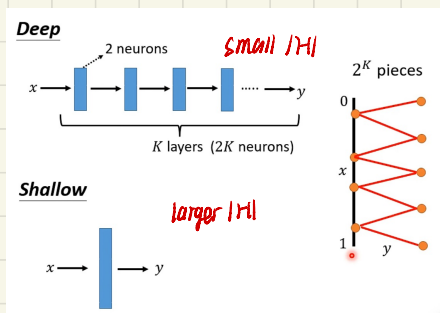


- One hidden layer can represent any function.
- using deep structure is more effective

eg.



Conclusion



所以 deep learning 有 small $|H|$ 和 small L , 达到双赢.