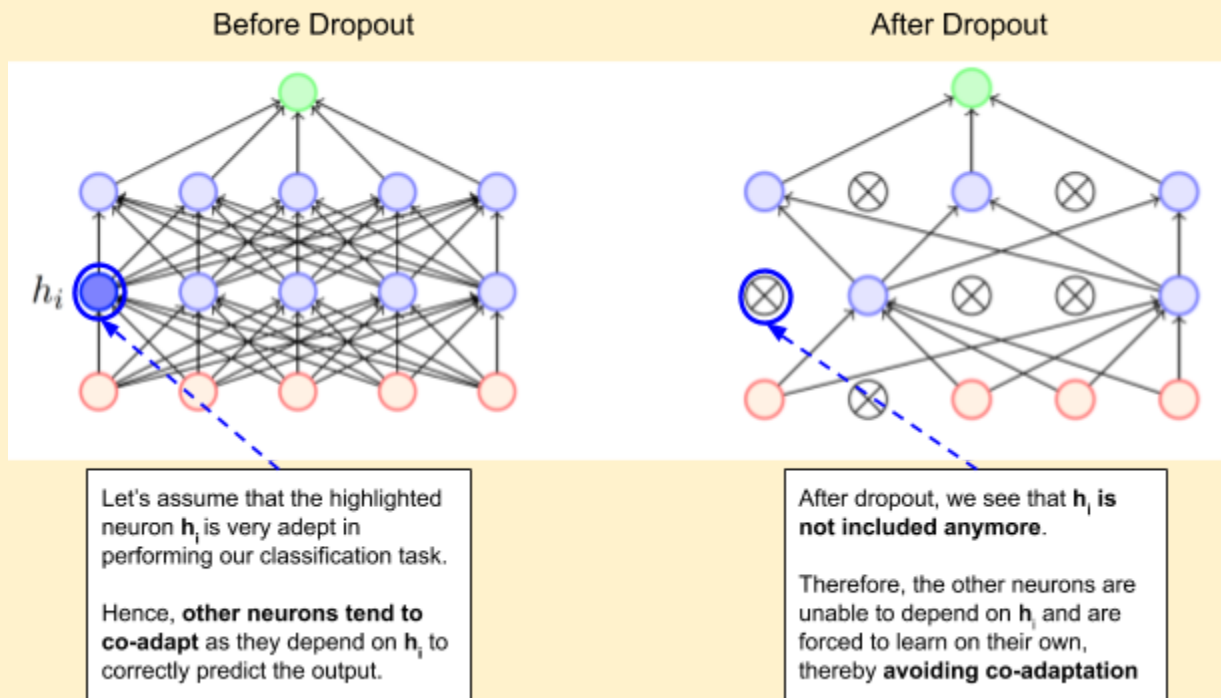


How does dropout act as a regularizer?

Why does dropout help?

1. Let us first talk about co-adaptation of neurons. It is the phenomenon where neurons adapt to each others' performance and begin individual specialization.
2. The following example will illustrate how we regularize the network using dropout



3. From the above diagram, we can see how other neurons are forced to learn to compensate for h_i being dropped from the network, thereby preventing co-adaptation.
4. Dropout acts as a regularizer by corrupting the input given to the subsequent layers. As certain neurons are dropped, this results in a less complete output being transferred to the next layer. This can also be viewed as adding noise to the input.
5. These processes add a degree of robustness to the network.