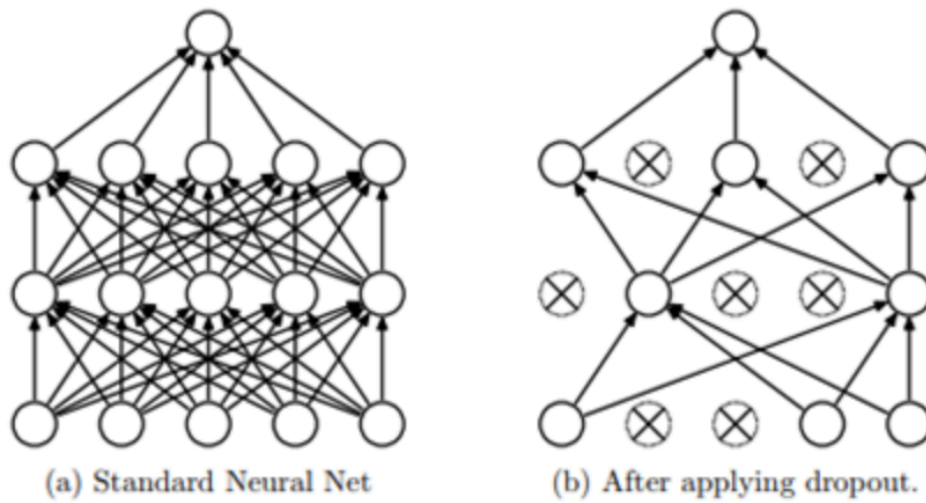


# Dropout

## 1. Concept

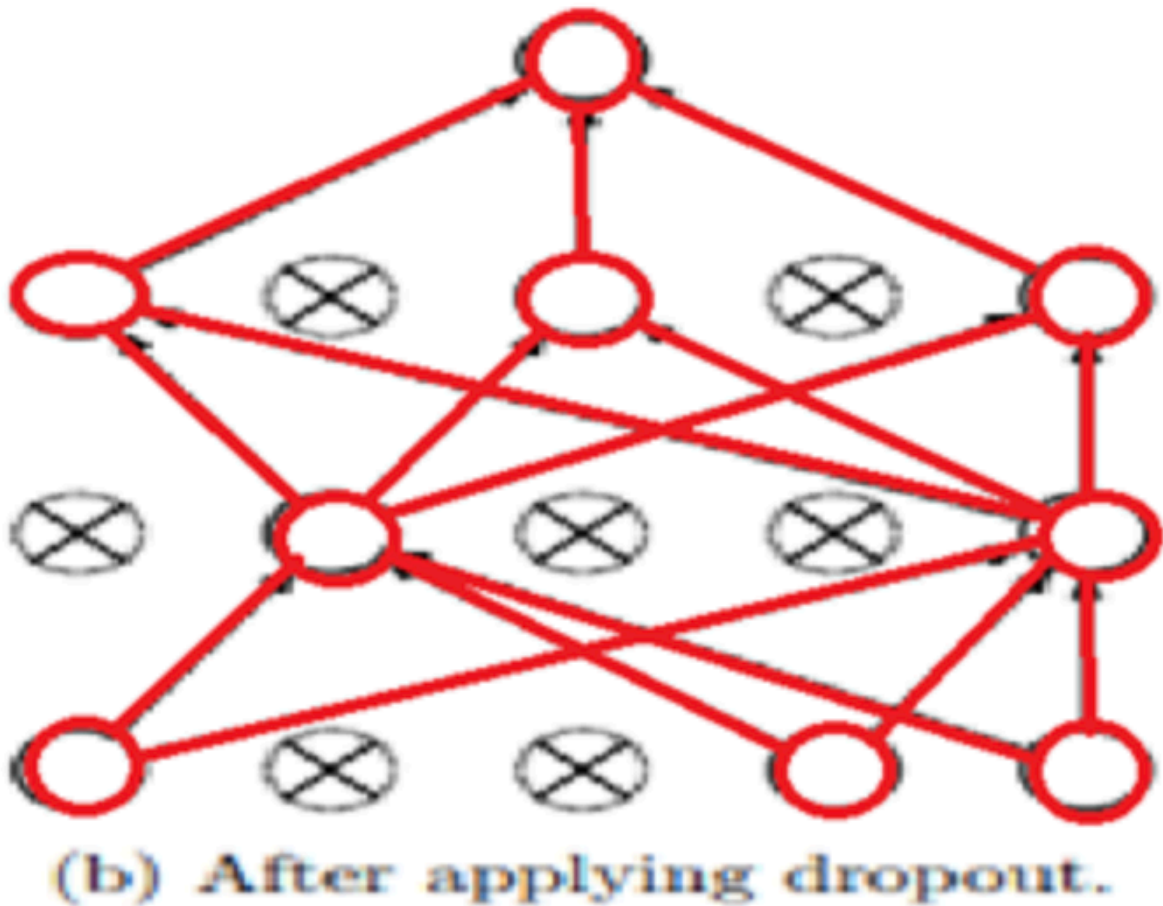


**Figure 1: Dropout Neural Net Model. Left:** A standard neural net with 2 hidden layers. **Right:** An example of a thinned net produced by applying dropout to the network on the left. Crossed units have been dropped.

기본 신경망의 구조는 (a)와 같이 각 레이어가 노드로 연결되어 있다.

하지만, 위 그림을 보면 모델이 깊어짐에 따라 선들이 매우 많아지게 된다. 즉, 과하게 학습하게 되며 이는 **Overfitting**을 야기한다.

**Dropout**은 과적합 방지를 위해 인간처럼 기억을 까먹을 수 있게 한 것이다.



(b)를 명확하게 나타내기 위한 그림이다.

Dropout은 선택적으로 노드(Neuron)를 Drop하는 것이다. 위 그림을 보면, Dropout의 결과 몇 개의 랜덤하게 선택받은 노드만을 가지고 학습을 하는 것으로, 모델이 훨씬 단순해질 것이다.

*Neuron: 그림에서의 동그라미*

이 방법은 Training Data에서는 학습이 덜 될 수도 있지만 일반화(Regularization)능력을 키워서 Test Data에 대한 예측률을 높이는 방법이다.

게다가 Dropout을 하였을 때, Dropout을 하지 않았을 때보다 오류가 더 낮아 성능이 좋다.

## 2. Code

```
W1 = tf.get_variable("W1", shape = [784, 512],
                    initializer = tf.contrib.layers.xavier_initializer())
b1 = tf.Variable(tf.random_normal([512]))

#Define Activation Function
L1 = tf.nn.relu(tf.matmul(X, W1) + b1)

#Dropout
L1 = tf.nn.dropout(L1, keep_prob = keep_prob)
```

*keep\_prob* : 노드를 선택할 확률을 정함. 예를들어 *keep\_prob*가 1이면 모든 노드를 다 쓰는 것이고, 0.5이면 노드를 랜덤하게 반만 선택한다. (0.6이 최적이나, 단순한 계산을 위해서라면 0.5를 사용해도 무방하다)

### 3. Dropout & Ensemble

*ensemble* : 여러 모델을 종합적으로 고려하여 최적의 결과를 찾아내는 것

예를 들어, 수능 수학을 풀 때에 미적분 전문가, 확통 전문가, 기백 전문가가 모여서 문제를 푸는 것. 이 세 분야의 전문가가 모여서 풀면 미적분 전문가 혼자서 모든 문제를 다 풀 때보다 정확도가 올라감. 이것이 앙상블이다.

Dropout은 앙상블과 유사하다고 할 수 있는데, 그 이유는 다음과 같다.

Dropout은 학습 때마다 랜덤으로 노드를 죽여서 학습했다. 쉽게 설명하자면, 첫 번째 학습 때에는 랜덤으로 미적분 전문가, 확통 전문가가 문제를 풀어보고 두 번째는 확통 전문가와 기백 전문가가 문제를 풀고.. 를 반복한다. 즉 서로 다른 모델을 학습시킨다. 이러한 측면에서 Dropout과 앙상블은 유사하다고 할 수 있다.