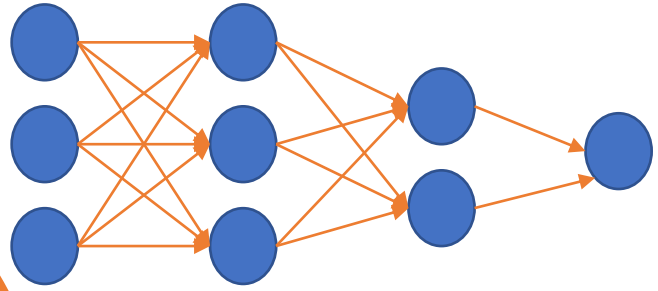
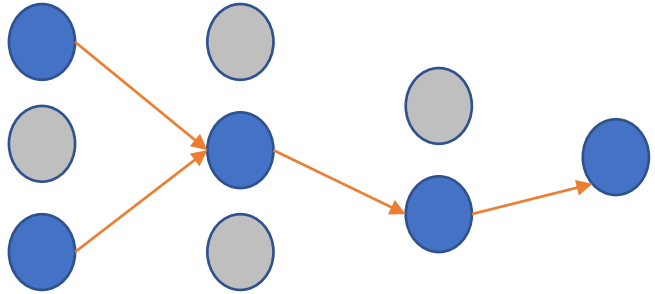
Dropout



Regular Feedforward Network



After applying Dropout

- ✓ Arguably the most effective regularization strategy against overfitting
- ✓ During training randomly turn-off a certain percentage of neurons
  - The percentage is a hyperparameter
  - "Turn-off" means that weights of those random neurons are set to 0
- ✓ The "turned-off" neurons do not have any output and hence they do not contribute to error
- ✓ Weights of dropped out neurons won't get updated during backpropagation
- ✓ Implies that more units/neurons will be active during test time
- ✓ Dropout strategy is not applied during test
  - Layer output values are scaled down by the same factor as dropout value
  - To balance for the fact that more neurons are active during test
  - Keras handles this internally
- ✓ Typically dropout values vary between 20% - 50%
- ✓ What's makes dropout an effective regularization strategy?
  - This strategy of turning off neurons randomly introduces noise in network output and hence prevents network from learning undesired patterns
- ✓ Higher dropout rate would prevent the network from learning
  - It is not common to see values higher than 50%