**Improving Neural Network Performance**

1. **Vanishing Gradients**
   * Activation Functions
   * Weight Initialization
2. **Overfitting**
   * Reduce Complexity/Increase Data
   * Dropout Layers
   * Regularization (L1 & L2)
   * Early Stopping
3. **Normalization**
   * Normalizing inputs
   * Batch Normalization
   * Normalizing Activations
4. **Gradient Checking and Clipping**
5. **Optimizers**
   * Momentum
   * Adagrad
   * RMSprop
   * Adam
6. **Learning rate scheduling**
7. **Hyperparameter Tuning**
   * No. of hidden layers
   * nodes/layer
   * Batch size