Improving Neural Network Performance

Overfitting

* Dropout Layers
* Regularization (L1 and L2)
* Early Stopping

Normalization

* Normalizing inputs
* Batch Normalization
* Normalizing Activations

Vanishing Gradients

* Activation Functions
* Weight Initialization ( Xavier and He Initialization)

Gradient Checking and Clipping

Optimizers