**Improving Neural Network Performance**

|  |  |
| --- | --- |
| Problem Name | Solution |
| Vanishing Gradients | * Activation Functions * Weight Initialization |
| Overfitting | * Reduce Complexity/ Increase Data * Dropout Layers * Regularization * Early Stopping |
| Normalization | * Normalizing Inputs * Batch Normalization * Normalizing Activations |
| Gradient Checking and Clipping | |
| Optimizers | * Momentum * Adagrad * RMSprop * Adam |
| Learning Rate and Scheduling | |
| Hyperparameter Tuning | * No of hidden layers * Nodes/layer * Batch Size |