1. How can each of these parameters be fine-tuned? • Number of hidden layers

• Network architecture (network depth)

• Each layer's number of neurons (layer width)

• Form of activation

• Optimization and learning

• Learning rate and decay schedule

• Mini batch size

• Algorithms for optimization

• The number of epochs (and early stopping criteria)

• Overfitting that be avoided by using regularization techniques.

• L2 normalization

• Drop out layers

• Data augmentation