Summary/Review

Deep Learning and Regularization

Technically, a deep Neural Network has 2 or more hidden layers (often, many more). Deep Learning involves Machine Learning with deep Neural Networks. However, the term Deep Learning is often used to broadly describe a subset of Machine Learning approaches that use deep Neural Networks to uncover otherwise-unobservable relationships in the data, often as an alternative to manual feature engineering. Deep Learning approaches are common in Supervised, Unsupervised, and Semi-supervised Machine Learning.

These are some common ways to prevent overfitting and regularize neural networks:

- Regularization penalty in cost function - This option explicitly adds a penalty to the loss function

$$J = \frac{1}{2n}\sum_{i=1}^{n}(\hat{y}_i - y_i)^2 + \lambda \sum_{j=1}^{m} W_i^2$$

- Dropout - This is a mechanism in which at each training iteration (batch) we randomly remove a subset of neurons. This prevents a neural network from relying too much on individual pathways, making it more robust. At test time the weight of the neuron is rescaled to reflect the percentage of the time it was active.

- Early stopping - This is another heuristic approach to regularization that refers to choosing some rules to determine if the training should stop.

  Example:

  Check the validation log-loss every 10 epochs.

  If it is higher than it was last time, stop and use the previous model.

- Optimizers - This approaches are based on the idea of tweaking and improving the weights using other methods instead of gradient descent.