Machine Learning for Medical Applications

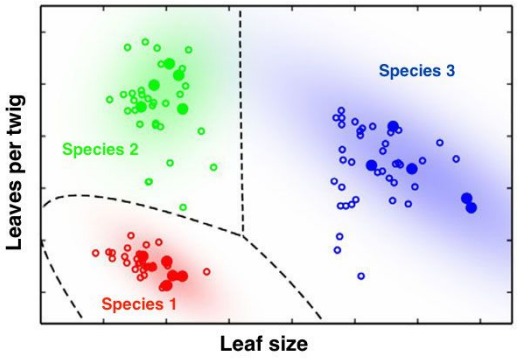
DSC at W&M, March 2021

What is machine learning?

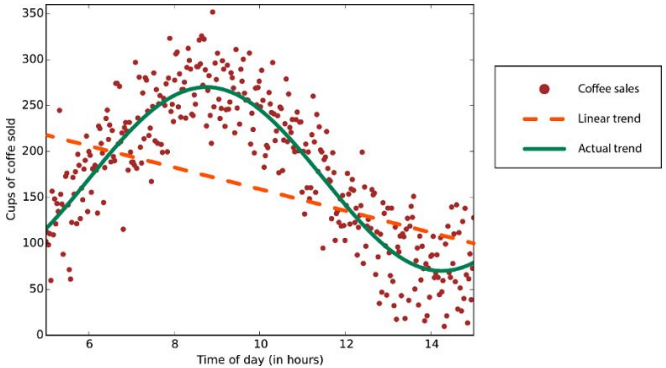
- An algorithm that “learns” from data to build a complex, non-linear model that can make predictions by identifying patterns and categories.

Supervised (labeled data)

Classification

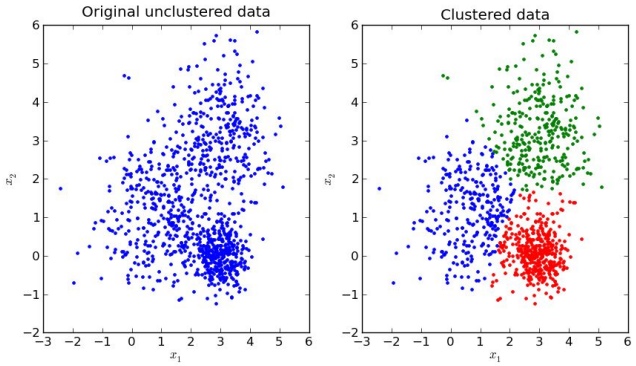


Regression



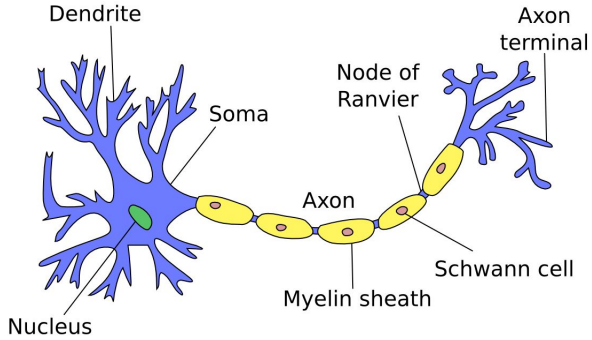
Unsupervised (unlabeled data)

Clustering



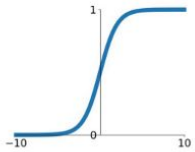
Neural Networks

- Input layer, hidden layer(s), output layer
- Weights
- Activation functions



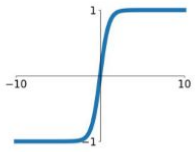
Sigmoid

$$\sigma(x) = \frac{1}{1+e^{-x}}$$



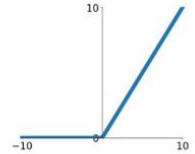
tanh

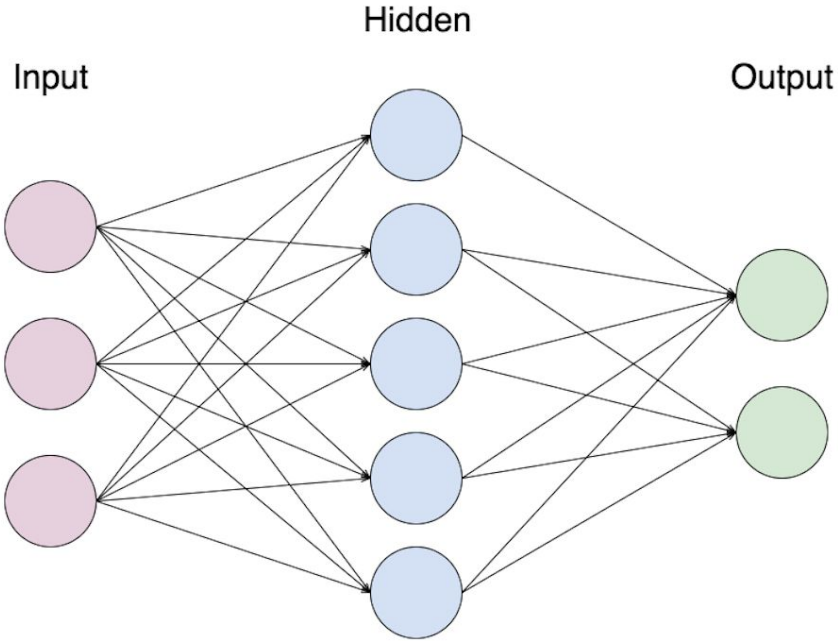
$$\tanh(x)$$



ReLU

$$\max(0, x)$$



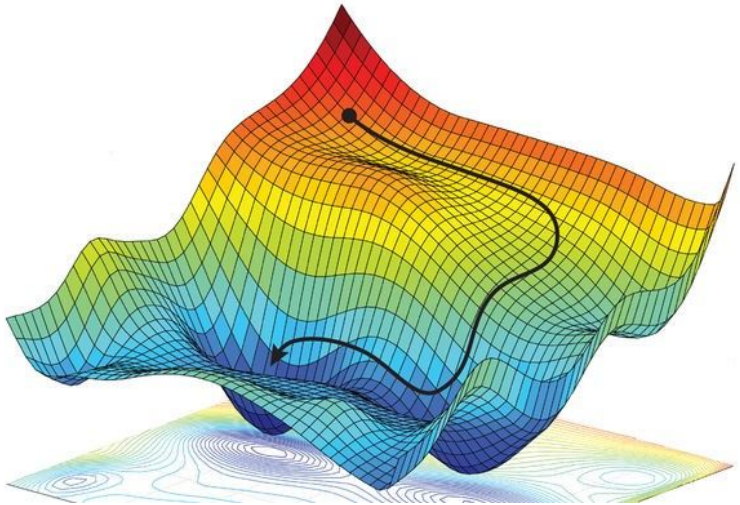


Techniques

- Data pre-processing
  - Data reduction
  - Normalizing values
- Test/training data
  - Split data into two sets (about 70-30 split)
  - Train the model on the training set, evaluate its performance on “unseen” data with the test set

Gradient Descent and Hyperparameters

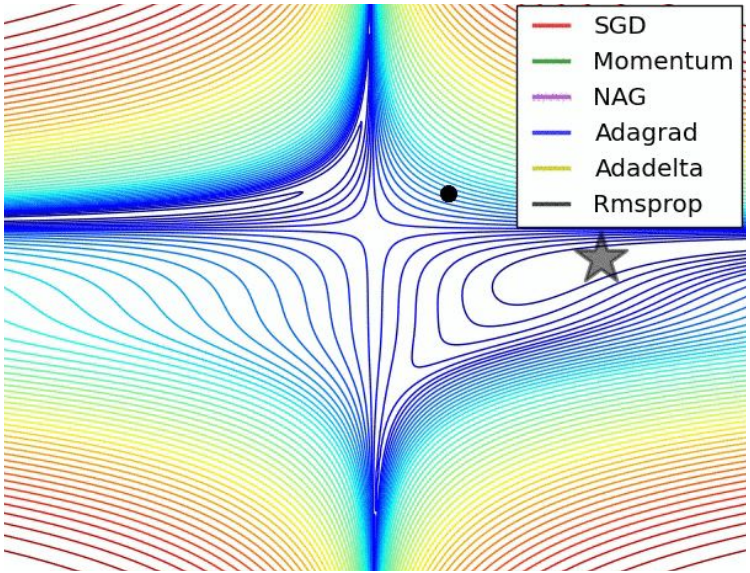
- Gradient descent
  - The algorithm chooses an initial random value for the weights
  - It then calculates the *direction of steepest descent* (the **gradient**) of the **loss function** (error) and takes a small step in that direction



Gradient Descent and Hyperparameters

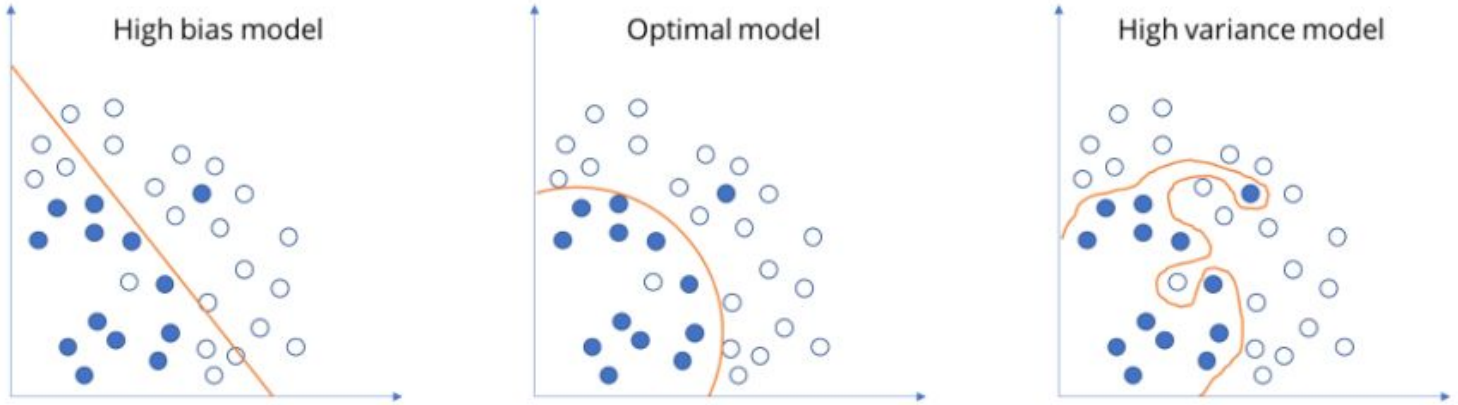
- Hyperparameters
  - **Learning rate**: the size of each step
  - **Batch size**: the amount of training data used to calculate the direction of steepest descent each step
  - **Number of epochs**: the number of times we feed the training data through the network
  - **Momentum**: a fraction of the previous update quantity for a weight is added to the current update to avoid getting stuck in a local minimum

Different Optimizers at Work

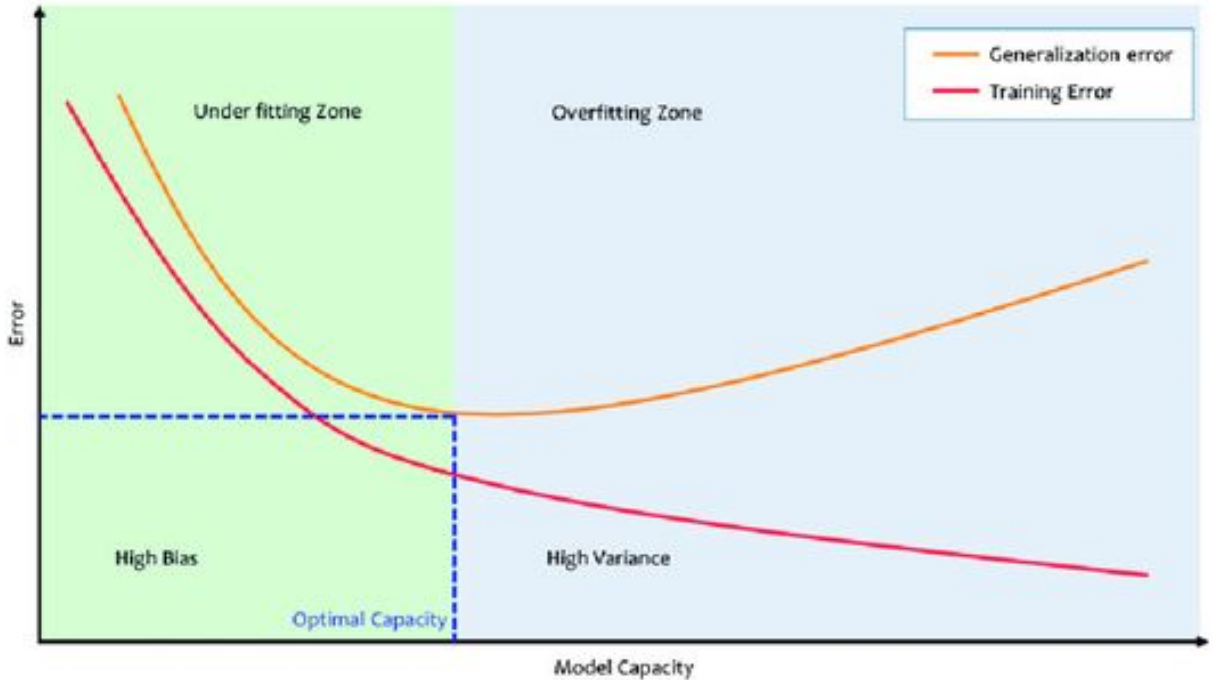


Overfitting and Generalization

- Bias
  - Underfitting: The model is too simple to fit the training data *or* the test data well.
- Variance
  - Overfitting: The model fits the training data very well but is too specific to generalize to test data.

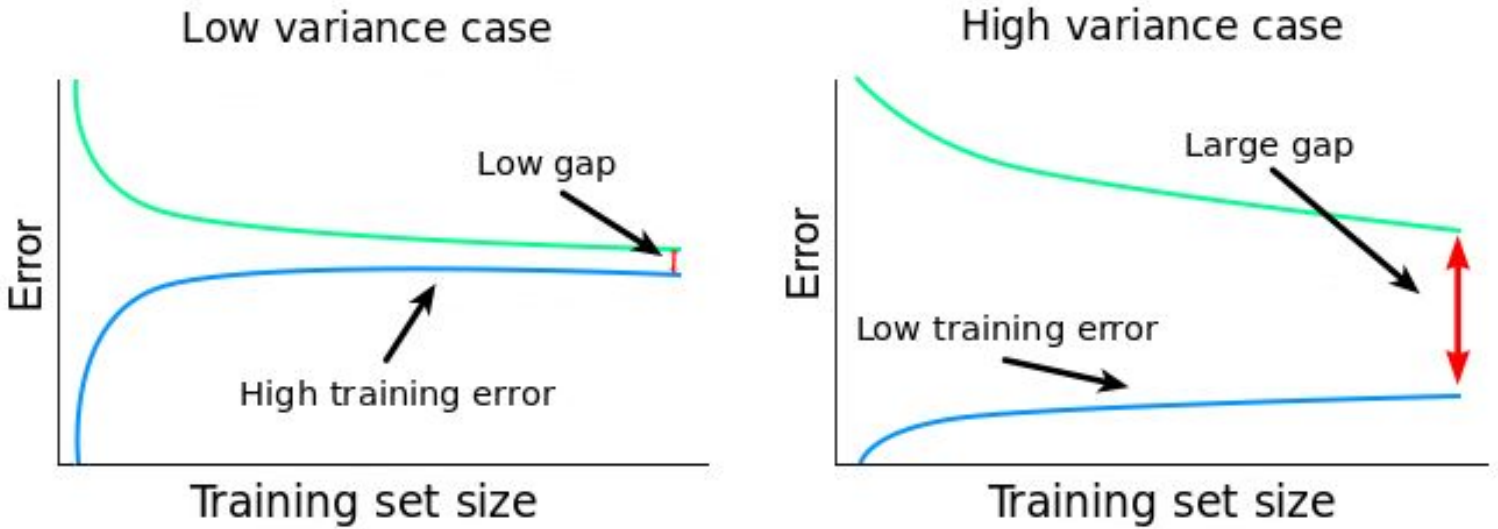


Overfitting and Generalization



Learning curves

Use learning curves to diagnose bias or variance and adjust model complexity accordingly.



Additional Resources

- A high-level intro to machine learning algorithms
- A more mathematical treatment of machine learning
- Keras documentation (what we'll use to build our networks)
- Machine learning as applied to medicine