CSE 3521 FINAL REVIEW

BRENDAN WHITAKER

Classification - which category DISCRETE - use logistic regression
Regression - predicting a continuous number CONTINUOUS - use linear regression
Clustering - does not have an output!!!

Two parts of a model: Structure, parameters, parameters come out of the model structure.
If I give you data, are they linearly separable. "How does the structure of the model I'm choosing reflect the structure of the data?"
How complicated is the shape of your data, and thus how complicated do we need the model to be?
Different data distributions require different models.

Feature function selection: Given a bunch, tell which are useful.
On-policy gives you a margin of safety.
Neural Net consists of linear transformations and nonlinear functions. You just alternate those. Two scalar values x_1, x_2 multiplied by weights w_1, w_2 enter a linear transformation, adding a constant for the function called w_0 so $f(x, y) = w_1 x_1 + w_2 x_2 + w_0$. The linear transformation is $f(x, y)$ since the weights are the linear coefficients. We can take 2 inputs x_1, x_2 and put them each into 2 nonlinear functions by crossing over, and then run the same process on the output of each of the two nonlinear functions. Discrete output looks like a threshold function with 90 degree angles, this is classification, and a continuous curve gives us linear regression, which is regression (these are the types of nonlinear functions, in the logistic case, it's a piecewise function).
Bayesian has a prior and posterior, Frequentist doesn't, no prior beliefs, just count and divide.