Decision Trees and the Dynamics of Classification

Alex Blaine

University of California, Davis

anblaine@ucdavis.com

June 5, 2018

Overview

Background

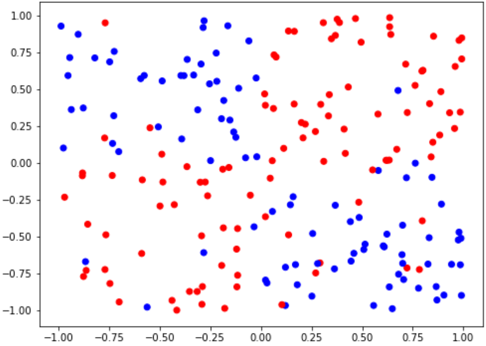
Methods and Preliminary Work

Decision Trees

One of the simplest forms of classification/regression techniques. Relies on making repeated binary cuts on variables to break down the domain into smaller regions. Each region then given a classification or a value.

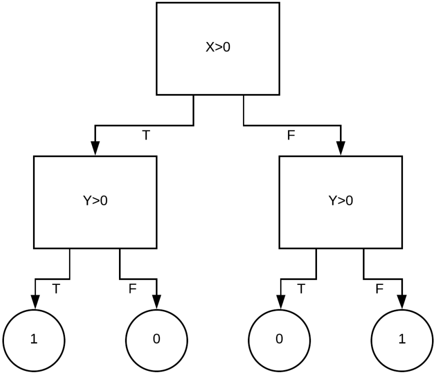
Decision Trees (cont.)

For an example, lets look at this generated data:



Decision Trees (cont.)

Class of each point is close to taking $\text{sgn}(x_i \cdot y_i)$, so we may use the following decision tree to classify new points:



A Look Inside

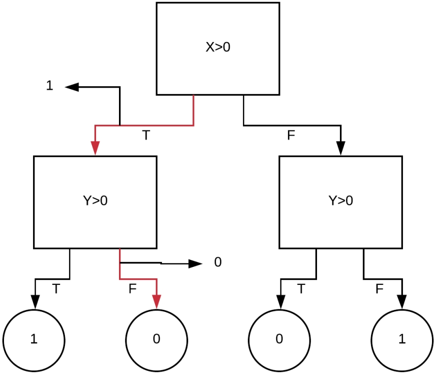
Now, my goal is to analyze the inner workings of a trained decision tree to see how it imposes structure on data. How to accomplish this?

A Look Inside (cont.)

The binary nature of the branching of the decision tree gives a natural way to output a string of bits detailing the inner workings for a particular data point.
Lets look at a sample point $(x, y) = (1, -1)$ to see this.

A Look Inside (cont.)

By examining the decision tree we can see the output for that data point is 10.



Workflow

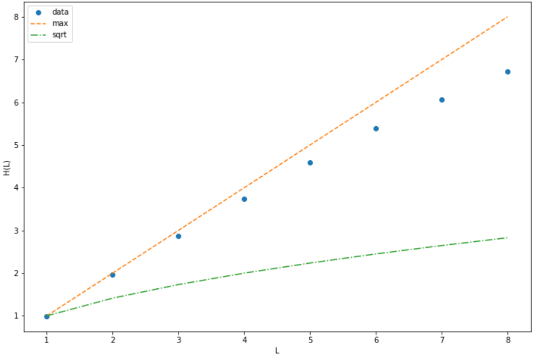
- Train decision trees on much larger data sets. Some of these data sets will be adversarially generated to provide baselines.
- Infer probabilities from output streams and perform block entropy analysis.
- Attempt to reconstruct an ϵ-machine for further insight.

Preliminary Work

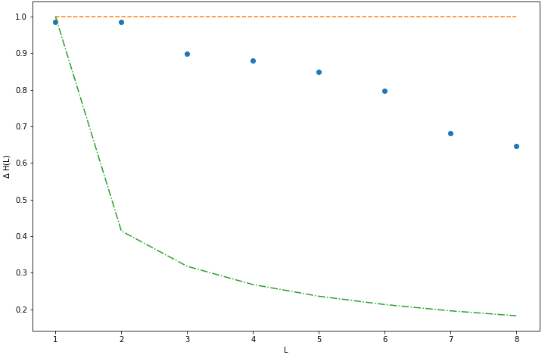
For one baseline I've been considering I'm looking at the block entropy of the binary representation of the data set and comparing it to linear and $\sqrt{L}$ growth.

Preliminary Work (cont.)

Here we can see near linear growth for small L, then rate starts to decrease quickly.



Preliminary Work (cont.)


1.0
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
Δ H(L)
1
2
3
4
5
6
7
8
L