Boosting Decision Stumps
Decision Stumps

Let \( x = (x_1, x_2, \ldots, x_n) \)

**Decision Stump** \( h_{i,t} \)

If \( x_i \geq t \)

then \( \text{class} = 1 \)

else \( \text{class} = -1 \)
Training Decision Stumps

Given data of the form $\mathbf{x} = (x_1, x_2, \ldots, x_n)$, one run of the training procedure determines the best $h_{i,t}$.

Algorithm:

- For each $x_i$ in the training set:
  - Sort values. Remove duplicates.
  - Construct candidate thresholds $t$ below min value, above max value, and midway between successive features
  - For each $h_{i,t}$, compute error on training set.

- Return $h_{i,t}$ that maximizes $\left| \frac{1}{2} - error(h_{i,t}) \right|$. 
Example of Training a Decision Stump
Adaboost on Decision Stumps

Run Adaboost for $T$ iterations, with $L$ being the decision-stump learning algorithm just described.

Decision stump $h_t$ is learned using training data selected from current distribution at time $t$.

Coefficient $\alpha_t$ is calculated by running $h_t$ on all training data.

If a decision stump using feature $x_i$ is chosen on an iteration, remove it from pool of features for next iteration.

After $T$ iterations, run ensemble classifier $H$ on test data.