One coordinate at a time

- Adaboost performs gradient descent on exponential loss
- Adds one coordinate ("weak learner") at each iteration.
- Weak learning in binary classification = slightly better than random guessing.
- Weak learning in regression – unclear.
- Uses example-weights to communicate the gradient direction to the weak learner
- Solves a computational problem
What is a good weak learner?

- The set of weak rules (features) should be flexible enough to be (weakly) correlated with most conceivable relations between feature vector and label.
- Small enough to allow exhaustive search for the minimal weighted training error.
- Small enough to avoid over-fitting.
- Should be able to calculate predicted label very efficiently.
- Rules can be “specialists” – predict only on a small subset of the input space and abstain from predicting on the rest (output 0).
Alternating Trees

Joint work with Llew Mason
Decision Trees

Decision Tree:
- **X > 3**
  - **Y > 5**
    - No
      - -1
    - Yes
      - +1
  - No
    - -1
  - Yes
    - +1
Decision tree as a sum

-0.2

X>3

-0.1

Y>5

+0.1

+0.2

sign

-0.3

-1

+1

-1

-1

+1

X

Y
An alternating decision tree

\[
\begin{array}{c}
\text{sign} \\
\begin{align*}
Y < 1 & : 0.0 \quad +0.7 \\
X > 3 & : -0.1 \quad +0.1 \\
Y > 5 & : -0.3 \quad +0.2 \\
\end{align*}
\end{array}
\]
Example: Medical Diagnostics

• **Cleve** dataset from UC Irvine database.
• Heart disease diagnostics (+1=healthy, -1=sick)
• 13 features from tests (real valued and discrete).
• 303 instances.
Adtree for Cleveland heart-disease diagnostics problem
## Cross-validated accuracy

<table>
<thead>
<tr>
<th>Learning algorithm</th>
<th>Number of splits</th>
<th>Average test error</th>
<th>Test error variance</th>
</tr>
</thead>
<tbody>
<tr>
<td>ADtree</td>
<td>6</td>
<td>17.0%</td>
<td>0.6%</td>
</tr>
<tr>
<td>C5.0</td>
<td>27</td>
<td>27.2%</td>
<td>0.5%</td>
</tr>
<tr>
<td>C5.0 + boosting</td>
<td>446</td>
<td>20.2%</td>
<td>0.5%</td>
</tr>
<tr>
<td>Boost Stumps</td>
<td>16</td>
<td>16.5%</td>
<td>0.8%</td>
</tr>
</tbody>
</table>