Bagging and Boosting

CS194-10 Fall 2011 Lecture 9
Boosting

\[ x_1 \]
\[ x_2 \]
\[ x_3 \]
\[ x_4 \]

\[ h_1 \]
Boosting

\[ h_1 = \begin{array}{c}
  x_1 \\
  x_2 \\
  x_3 \\
  x_4 \\
\end{array} \begin{array}{c}
\checkmark \\
\times \\
\times \\
\checkmark \\
\end{array} \]

\[ h_1 = \text{node} \]
Boosting

\[ h_1 = h_2 \]

\( x_1 \)
\( x_2 \)
\( x_3 \)
\( x_4 \)

\( h_1 \)
\( h_2 \)
Boosting

\[ h_1 = h_2 = \]

\[ x_1 \quad \checkmark \quad \checkmark \]
\[ x_2 \quad \times \quad \checkmark \]
\[ x_3 \quad \times \quad \checkmark \]
\[ x_4 \quad \checkmark \quad \times \]

\[ h_1 = \quad \text{Diagram 1} \quad \checkmark \quad \text{Diagram 2} \]
\[ h_2 = \quad \text{Diagram 3} \quad \checkmark \quad \text{Diagram 4} \]
Boosting

\[ h_1 = h_2 = h_3 \]

\[ x_1 \]
\[ x_2 \]
\[ x_3 \]
\[ x_4 \]

\[ h_1 = \]
\[ h_2 = \]
\[ h_3 \]
Boosting

\[ h_1 = h_2 = h_3 = \]

\[ x_1 \]  
\[ x_2 \]  
\[ x_3 \]  
\[ x_4 \]  

\[ h_1 = \]  
\[ h_2 = \]  
\[ h_3 = \]
Boosting

\[ h_1 = h_2 = h_3 = h_4 \]

\[ x_1 \]
\[ x_2 \]
\[ x_3 \]
\[ x_4 \]

\[ h_1 = \]
\[ h_2 = \]
\[ h_3 = \]
\[ h_4 = \]
Boosting

\[ h_1 = h_2 = h_3 = h_4 = \]

\[ x_1 \]

\[ x_2 \]

\[ x_3 \]

\[ x_4 \]
Boosting

\[ h_1 = h \]
\[ h_2 = h \]
\[ h_3 = h \]
\[ h_4 = h \]
Boosted decision stumps: Restaurant data

The graph plots the proportion of correct predictions on the test set against the training set size. The x-axis represents the training set size, while the y-axis shows the proportion correct on the test set. Two methods are compared: Boosted decision stumps and Decision stump. The graph shows that the Boosted decision stumps method generally outperforms the Decision stump method, especially as the training set size increases.
Effect of adding more predictors

![Graph showing the effect of adding more predictors on training and test accuracy. The x-axis represents the number of hypotheses M, ranging from 0 to 200. The y-axis represents training and test accuracy, ranging from 0.6 to 1. The graph shows the training error (red line) and test error (green dotted line) as M increases.]