Active learning
Co-training
Summary

- **Boosting** is a method for learning an accurate classifiers by combining many weak classifiers.
- Boosting is **resistant to over-fitting**.
- **Margins** quantify prediction confidence.
- **High noise** is a serious problem for learning classifiers—can’t be solved by minimizing convex functions.
- **Robustboost** can solve some high noise problems. Exact characterization still unclear.
- **Jboost** - an implementation of ADTrees and various boosting algorithms in java.
- **Book** on boosting coming this spring.

- Thank you, questions?
Pedestrian detection - typical segment
Current best results
Image Features

“Rectangle filters”

Similar to Haar wavelets

Papageorgiou, et al.

$$h_t(x_i) = \begin{cases} 1 & \text{if } f_t(x_i) > \theta_t \\ 0 & \text{otherwise} \end{cases}$$

Very fast to compute using “integral image”.

Combined using adaboost
Yotam’s features

\[\max (p1,p2) < \min(q1,q2,q3,q4)\]

Faster to calculate than Viola and Jones
Search for a good feature based on genetic programming
Definition

• Feature works in one of 3 resolutions: full, half, quarter

• Two sets of up to 6 points each

• Each point is an individual pixel

• Feature says yes if all white points have higher values than all black points, or vice versa
Advantages

• Deal better with the variation in illumination, no need to normalize.

• Highly efficient (3-4 image access operations). 2 times faster than Viola&Jones

• 20% of the memory
Steps of batch learning

- Collect labeled examples
- Run learning algorithm to generate classification rule
- Test classification rule on new data.
Labeling process

- Collected 6 Hrs of video $\rightarrow$ 540,000 frames
- 170,000 boxes per frame
- 20 seconds for marking a box around a pedestrian.
- 3 seconds for deciding if box is pedestrian or not.
- How to choose “hard” negative examples?

1500 pedestrians
Steps of active learning

- Collect labeled examples
- Run learning algorithm to generate classification rule
- Apply classifier on new data and label informative examples.
SEVILLE screen shot 1
SEVILLE screen shot 2
Margins

Consider the following:

An example: \( <x, y> \) e.g. \( <+, +1> \)

Normalized score:

\[
-1 \leq \frac{\sum_{t=1}^{T} \alpha_t h_t(x)}{\sum_{t=1}^{T} |\alpha_t|} \leq 1
\]

The margin is:

\[
y \frac{\sum_{t=1}^{T} \alpha_t h_t(x)}{\sum_{t=1}^{T} |\alpha_t|}
\]

margin \( > 0 \) means correct classification
Display the rectangles inside the margins.
large margins $\Rightarrow$ reliable predictions

Validation

Learning

5/17/06 UCLA
Margin Distributions
## Summary of Training effort

<table>
<thead>
<tr>
<th>Step</th>
<th>Total candidates</th>
<th>Presented</th>
<th>Labeled</th>
<th>Human labor</th>
<th>Positive</th>
<th>Negative</th>
<th>Training time</th>
<th>Weak rules</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>510 K</td>
<td>-</td>
<td>16</td>
<td>3m</td>
<td>6</td>
<td>10</td>
<td>2s</td>
<td>1</td>
</tr>
<tr>
<td>2</td>
<td>680 K</td>
<td>364</td>
<td>403</td>
<td>3m</td>
<td>36</td>
<td>374</td>
<td>6s</td>
<td>3</td>
</tr>
<tr>
<td>3</td>
<td>3,400 K</td>
<td>153</td>
<td>156</td>
<td>4m</td>
<td>46</td>
<td>520</td>
<td>22s</td>
<td>7</td>
</tr>
<tr>
<td>4</td>
<td>66,470 K</td>
<td>805</td>
<td>852</td>
<td>10m</td>
<td>86</td>
<td>1332</td>
<td>1m30s</td>
<td>30</td>
</tr>
<tr>
<td>5</td>
<td>37,910 K</td>
<td>1350</td>
<td>1439</td>
<td>10m</td>
<td>182</td>
<td>2675</td>
<td>8m</td>
<td>59</td>
</tr>
<tr>
<td>6</td>
<td>116,960 K</td>
<td>5150</td>
<td>5364</td>
<td>1h30m</td>
<td>417</td>
<td>7804</td>
<td>1h10m</td>
<td>270</td>
</tr>
<tr>
<td>7</td>
<td>24,140 K</td>
<td>1320</td>
<td>863</td>
<td>3h</td>
<td>848</td>
<td>8236</td>
<td>7h30m</td>
<td>893</td>
</tr>
<tr>
<td>8</td>
<td>189,550 K</td>
<td>8690</td>
<td>8707</td>
<td>3h</td>
<td>1178</td>
<td>16613</td>
<td>17h</td>
<td>1500</td>
</tr>
<tr>
<td>9</td>
<td>209,610 K</td>
<td>2933</td>
<td>2933</td>
<td>3h</td>
<td>1486</td>
<td>19238</td>
<td>30h</td>
<td>2034</td>
</tr>
<tr>
<td>10</td>
<td>274,210 K</td>
<td>3861</td>
<td>3861</td>
<td>4h</td>
<td>2046</td>
<td>22533</td>
<td>30h</td>
<td>3150</td>
</tr>
</tbody>
</table>

5/17/06 UCLA
Summary of Training

Only examples whose score is in this range are hand-labeled

<table>
<thead>
<tr>
<th>Step</th>
<th>total candidates</th>
<th>(\mu^-)</th>
<th>(\mu^+)</th>
<th>presented</th>
<th>labeled</th>
<th>human labor</th>
<th>positive</th>
<th>negative</th>
<th>training time</th>
<th>Weak rules</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>510 K</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>16</td>
<td>3m</td>
<td>6</td>
<td>10</td>
<td>2s</td>
<td>1</td>
</tr>
<tr>
<td>2</td>
<td>680 K</td>
<td>0</td>
<td>1</td>
<td>364</td>
<td>403</td>
<td>3m</td>
<td>36</td>
<td>374</td>
<td>6s</td>
<td>3</td>
</tr>
<tr>
<td>3</td>
<td>3,400 K</td>
<td>0.6</td>
<td>1</td>
<td>153</td>
<td>156</td>
<td>4m</td>
<td>46</td>
<td>520</td>
<td>22s</td>
<td>7</td>
</tr>
<tr>
<td>4</td>
<td>66,470 K</td>
<td>0.4</td>
<td>1</td>
<td>805</td>
<td>852</td>
<td>10m</td>
<td>86</td>
<td>1332</td>
<td>1m30s</td>
<td>30</td>
</tr>
<tr>
<td>5</td>
<td>37,910 K</td>
<td>0.1</td>
<td>0.8</td>
<td>1350</td>
<td>1439</td>
<td>10m</td>
<td>182</td>
<td>2675</td>
<td>8m</td>
<td>59</td>
</tr>
<tr>
<td>6</td>
<td>116,960 K</td>
<td>0</td>
<td>0.6</td>
<td>5150</td>
<td>5364</td>
<td>1h30m</td>
<td>417</td>
<td>7804</td>
<td>1h10m</td>
<td>270</td>
</tr>
<tr>
<td>7</td>
<td>24,140 K</td>
<td>-0.02</td>
<td>0.5</td>
<td>1320</td>
<td>863</td>
<td>3h</td>
<td>848</td>
<td>8236</td>
<td>7h30m</td>
<td>893</td>
</tr>
<tr>
<td>8</td>
<td>189,550 K</td>
<td>-0.02</td>
<td>0.5</td>
<td>8690</td>
<td>8707</td>
<td>3h</td>
<td>1178</td>
<td>16613</td>
<td>17h</td>
<td>1500</td>
</tr>
<tr>
<td>9</td>
<td>209,610 K</td>
<td>-0.02</td>
<td>0.5</td>
<td>2933</td>
<td>2933</td>
<td>3h</td>
<td>1486</td>
<td>19238</td>
<td>30h</td>
<td>2034</td>
</tr>
<tr>
<td>10</td>
<td>274,210 K</td>
<td>-0.02</td>
<td>0.5</td>
<td>3861</td>
<td>3861</td>
<td>4h</td>
<td>2046</td>
<td>22533</td>
<td>30h</td>
<td>3150</td>
</tr>
</tbody>
</table>
Few training examples
After re-labeling feedback
Final detector
Examples - easy

Positive

Negative
Examples - medium

Positive

Negative
Examples - hard
And the figure in the gown is..
Seville cycles
Summary

• Boosting and SVM control over-fitting using margins.
• Margins measure the **stability** of the prediction, not conditional probability.
• Margins are useful for co-training and for active-learning.