Outline

- Ensemble Methods in Machine Learning
- Boosting

Different Classifiers (1)

- Different Classifiers
  - Conduct classification on a same set of class labels
  - May use different input or have different parameters
  - May produce different output for a certain example

- Learning Different Classifiers
  - Use different training examples
  - Use different features
Different Classifiers (2)

- **Performance**
  - Each of the classifiers is not perfect
- **Complementary**
  - Examples which are not correctly classified by one classifier may be correctly classified by the other classifiers

- **Potential Improvements?**
  - Utilize the complementary property

Ensembles of Classifiers

- **Idea**
  - Combine the classifiers to improve the performance

- **Ensembles of Classifiers**
  - Combine the classification results from different classifiers to produce the final output
    - Unweighted voting
    - Weighted voting
Example: Weather Forecast

<table>
<thead>
<tr>
<th>Reality</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>Combine</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td></td>
</tr>
<tr>
<td></td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td></td>
</tr>
<tr>
<td></td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td></td>
</tr>
<tr>
<td></td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td></td>
</tr>
<tr>
<td></td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td></td>
</tr>
</tbody>
</table>

Machine Learning Basics: 3. Ensemble Learning

Ensemble Learning

- Ensemble Learning
  - Relatively new field in machine learning
  - Achieve state-of-the-art performance
- Central Issues in Ensemble Learning
  - How to create classifiers with complementary performances
  - How to conduct voting

Machine Learning Basics: 3. Ensemble Learning
**Strong and Weak Learners**

- **Strong Learner**
  - Take labeled data for training
  - Produce a classifier which can be arbitrarily accurate
  - Objective of machine learning

- **Weak Learner**
  - Take labeled data for training
  - Produce a classifier which is more accurate than random guessing

**Boosting**

- **Learners**
  - Strong learners are very difficult to construct
  - Constructing weaker Learners is relatively easy

- **Strategy**
  - Derive strong learner from weak learner
  - Boost weak classifiers to a strong learner
Construct Weak Classifiers

- **Using Different Data Distribution**
  - Start with uniform weighting
  - During each step of learning
    - Increase weights of the examples which are not correctly learned by the weak learner
    - Decrease weights of the examples which are correctly learned by the weak learner

- **Idea**
  - Focus on difficult examples which are not correctly classified in the previous steps

Combine Weak Classifiers

- **Weighted Voting**
  - Construct strong classifier by weighted voting of the weak classifiers

- **Idea**
  - Better weak classifier gets a larger weight
  - Iteratively add weak classifiers
    - Increase accuracy of the combined classifier through minimization of a cost function
Example

Training

Combined classifier

Performance

- **Data Set**
  - 27 data sets from UCI ML Repository

- **Methods for Comparison**
  - Decision tree classifier: C4.5
  - Boosting: AdaBoost using C4.5 as the weak learner
Results (Freund and Schapire 1996)

Machine Learning Basics: 3. Ensemble Learning