# Theoretical Perspectives and New Applications

- Three Questions
  - Name one open theoretical question in ensemble learning
  - Describe one application for which ensemble learning is NOT APPROPRIATE and explain why
  - Describe a new kind of application where ensemble methods have not been applied but where they should work

# Panel Members

- Tin Kam Ho
- Ludmila Kuncheva
- Nathan Intrator
- Fabio Roli
- Tom Dietterich

# Open Theoretical Questions

- We need a theory of WHEN ensemble methods work
  - Relate measurable properties of the data to expected performance of ensemble methods
  - Give guidance to the design of ensemble methods

# Example: Bias-Variance Theory

- High bias: Need to increase expressive power of classifiers
  - Adaboost
  - Intrator & Cohen (mix model types)
  - Tetko: bias correction
- High variance: Need to smooth multiple classifiers
  - Bagging, random forests
  - Tune classifiers to have low bias/high variance, then combine

# Shortcomings of Bias-Variance Theory

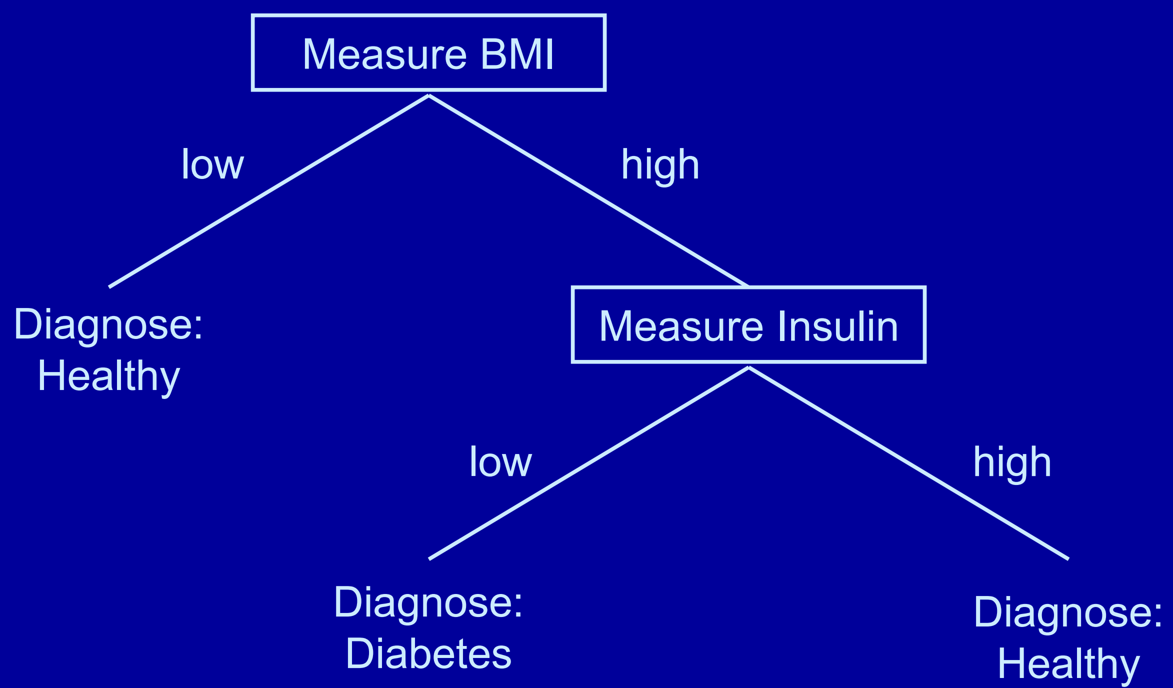
- Predictions not always correct
  - 1-NN has high variance but bagging does not work
- Problems measuring bias and variance
  - Experimental measurements underestimate variance
  - Experimental measurements combine bias and noise (but theory combines variance and noise)

# Application where ensembles cannot be applied

- Cost-sensitive diagnosis
- Repeat
  - Choice:
    - Halt and make a diagnosis: y
    - Choose a medical test x_i, execute it, obtain result

# Ensemble Methods Don't Work

- We must choose exactly one attribute to measure at each time t
- Equivalent to a single decision tree



# A Possibility

- Apply an ensemble in each leaf to estimate P(y|x) given the x values observed so far

# A New Application for Ensembles

- Markov Random Fields for image processing and remote sensing
  - $P(Y \mid X) = \exp(\Sigma_\alpha\, w_\alpha\, \psi_\alpha(y_a, y_b, x_a, x_b))/Z$
  - Hard to train
  - Hard to evaluate
    - Search for best y's
    - Evaluating Z

$$
\begin{array}{ccccccc}
y_{1,1} & - & y_{1,2} & - & y_{1,3} & - & y_{1,4} \\
| & & | & & | & & | \\
y_{2,1} & - & y_{2,2} & - & y_{2,3} & - & y_{2,4} \\
| & & | & & | & & | \\
y_{3,1} & - & y_{3,2} & - & y_{3,3} & - & y_{3,4} \\
| & & | & & | & & | \\
y_{4,1} & - & y_{4,2} & - & y_{4,3} & - & y_{4,4}
\end{array}
$$

# Ensemble of Tree-Structured Random Fields

- Each tree structure can be trained via fast dynamic programming algorithm
- Z can be evaluated by dynamic programming
- Run-time application is still expensive

# More Questions?