(21st March, Week 4)

1 Ensembles

Some of the best performing classification methods to date come from the idea; aggregating other classifiers... We will look at

- Averaging
- Boosting
- Bootstrapping
- Bagging
- Cascading
- Random Forests
- Stacking

Ofcourse, we wouldn't do this unless **ensemble methods often have a higher accuracy than the individual input classifiers**

2 Finding

We recall the bias, variance tradeoff. E_{train} and E_{approx}. We can make E_{train} "really small" at the cost of E_{approx} and vice versa.

We find that ensemble methods can do "much better" on one of these, than each individual classifier, whilst not being "only somewhat worse" on the other.

- **Boosting**; improves E_{train} in cases where E_{train} is high, and,
- **Averaging**; improves E_{approx} in cases where E_{approx} is high.

And now we begin sections for each of these methods in the following lecture...