

Oct 18th

Ensemble methods

- construct a set of classifiers from training data
- predict class label of prev. unseen

* If all classifiers are independent, for 25 classifiers each w/ .35 error \rightarrow ensemble is 0.06 error.
 \hookrightarrow to keep independent, use bagging (sample w/ replacement)

Boosting

- iterative procedure to adaptively distribute. TRAIN ON WHAT YOU ARE NOT GOOD AT, SO KEEP GOING MORE.
- Adaboost (weighted)

Recommender System

| | | |
|--------------|----------------------------|--|
| given: users | goal: rec. movies to users | challenges |
| movies | | SCALE [MILLIS OF USERS, MILL OF MOVIES] |
| Rating | | COLD START [FIRST USER, NEW MOVIE] |
| | | SPARSE DATA [NOT MANY USERS RATE MOVIES] |

use rating predic. as proxy for rec.

sim measures

b/w users

b/w movies.

Feature Extraction

(content based)