# Notes on Machine Learning

* Using rating alone can skew the data, does not take popularity into account – tends to favor movies with extremely high ratings from a small number of users.
* Can use a weighted rating (WR) = (v/(v+m) \* R) + (m/(v+m) \* C)
* V = vote count, R = vote average
* Use a percentile or total count cutoff to eliminate data that has too few votes
* Use cross-validation, instead of shotgun approach?