CSE518A
David Huang
Assignment 2 Writeup
Discussed with: John Kirchenbauer, Reese Wills

Part 1-4

Brief description of each algorithm's implementation:

- **Majority Vote:** for each task, count the votes for "1", if the count is greater than k/2 then predict 1, otherwise 0
- **EM:** use Binomial CDF to calculate likelihood of each worker answering correctly (latent parameter), then adjust predictions by doing D*theta (where D is matrix of worker responses for each task, and theta is worker quality vector) and normalizing, repeat until convergence
- **SVD:** following the approach from *Who Moderates the Moderators?* We "cheat" and find a trustworthy worker (highest accuracy), perform SVD on matrix of worker responses, take sign of top eigenvector, v, from left singular matrix and correct best worker's responses based off v

Below are the results from each model.

k	MV	EM	SVD
1	0.2375	0.2375	0.9425
2	0.21625	0.17875	0.8975
3	0.215	0.17	0.845
4	0.15125	0.13625	0.78625
5	0.1475	0.12375	0.73375
6	0.1225	0.1125	0.6775
7	0.12625	0.0925	0.68375
8	0.095	0.10125	0.6425
9	0.11875	0.0925	0.5375
10	0.08125	0.08875	0.505

Table 1: Average error rate

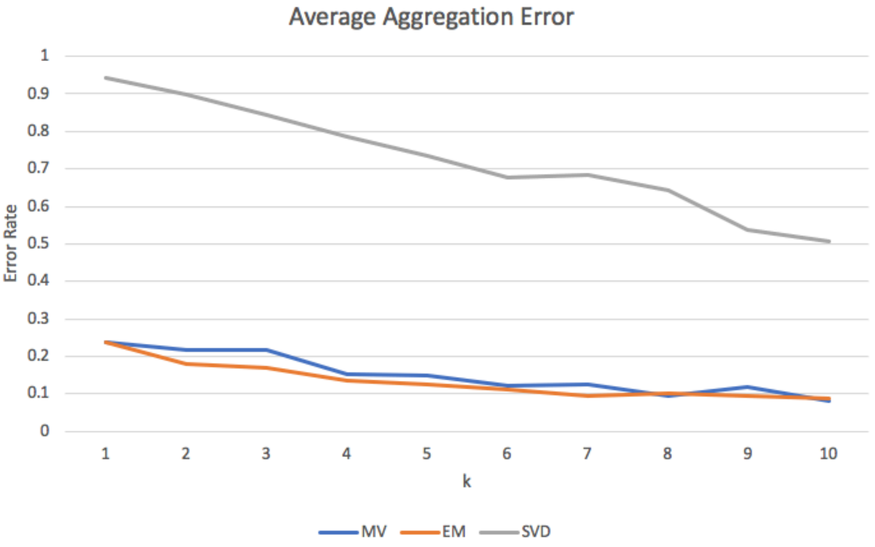

Figure 2: (**Part 5**) Error Rate Trend

Part 6

The clear trend is prediction accuracy increases as more workers are provided to the algorithms. Out of all the algorithms MV and EM perform the best most likely because they are less sensitive to sparse data. The SVD algorithm is only able to do just as good as random guessing since most tasks are missing workers. EM and MV follow each other closely and don't seem to diverge much. For this particular dataset this behavior might mean that all the workers are reliable (answer correctly w/ high likelihood), so assigning weights to a worker is effectively the same as doing MV (where weight is technically 1).

An interesting characteristic of EM I noticed on this dataset was that it seemed to converge quite quickly. I decided to test this, so I varied the max number of iterations between 10 and 20 but found little to no difference in the performance of the algorithm.