CS297

When Can I Trust an Average Rating on Amazon?

By

Pushkar Umaranikar

Supervised by : Dr. Pollett

Outline…

- Online Review System
- Aggregation of reviews and challenges
- Bayesian ranking
- Amazon review system
- Amazon review summary
- Averaging estimators

Online Review System

- We generally pay attention to number of stars for each product.

- We should also consider number of reviews behind averaged number of stars.

- Each review consists of
  - Rating
  - Review
  - Review of review

Online Review System (2)

- People don't read through all reviews.
- Summary of review needed to aggregate individual review.
- Factors considered
  - Some level of accuracy
  - Review population size
  - Usefulness

- Then ,what is proper aggregation ?

Galton's experiment

- Game of guessing the weight of an ox.
- 787 people, guessing independently.
- Average was 1197 pounds, ox weighted 1198 pounds.
- Why it worked well ?
  - An objective answer with clear numerical meaning
  - Unbiased and independent estimates.
  - Enough people participation.

Challenges in aggregation

- Tradeoff between review population and average rating score.

- How to view aggregated ratings ?

- Defining trustworthy average rating.

- Review consideration for finding out average rating.

- Actual challenge is turning vectors into scalars.

Bayesian ranking

- Weighting the raw rating scores with the population sizes.

- R: rating of all the products.

- N : total number of reviews for all the brands.

$$\tilde{r}_i = \frac{NR + n_i r_i}{N + n_i}.$$

What does Amazon do ?

- Rating based on formula in addition to Bayesian ranking.

- Includes :

  – Bayesian adjustment

  – Recency of reviews

  – Reputation score of reviewer

- Exact formula is not known outside Amazon.

Guidelines on Amazon ranking

- N is chosen between max and avg value of N.

- Products having low recency of helpful reviews are at lower locations.

- Products having high quality, positive reviews are up in ranking.

- Products having major issue are demoted in ranking.

Averaging sequentially-trained estimators

- To make average estimator more accurate.

- Focuses on points that we do not know very well.

- With N estimators and M datasets, each estimator y_i gets trained by the datasets.

- Some datasets are well handled while others are less.

- We should adapt accordingly, and give challenging datasets, heavier weight w in the next estimator's parameter training.