

5-4: Rank Metrics

Introduction to Recommender Systems

Learning Objectives

- Understand the basic idea of measuring rank accuracy vs. decision support or prediction accuracy
- Understand how to apply and interpret common rank accuracy metrics

Introduction

- Previous lectures
 - Evaluation methodology
 - Prediction accuracy metrics
 - Decision support metrics
- This lecture
 - Measuring how good a recommender is at ranking

Introduction to Recommender Systems

Metric Families

- Prediction accuracy: how well does the recommender estimate preference?
- Decision support: how well does the recommender do at finding good things?
- Rank accuracy: how well does the recommender estimate *relative* preference?
 - Putting items in order by preference

Overview

- Mean Reciprocal Rank
- Spearman Rank Correlation
- Discounted Cumulative Gain
- Fraction of Concordant Pairs

Introduction to Recommender Systems

Mean Reciprocal Rank

- Reciprocal rank: $1/i$, where i is the rank of the first 'good' item
- Similar to precision/recall
 - P/R measures how good recommender is at only being relevant (precision) and finding things (recall)
 - RR measures how far you have to go to find something good
- MRR is just average over all test queries

Introduction to Recommender Systems

Spearman Rank Correlation

- Pearson correlation over ranks
- Punishes misplacement

$$\frac{\sum_i (r_1(i) - \mu_1)(r_2(i) - \mu_2)}{\sqrt{\sum_i (r_1(i) - \mu_1)^2} \sqrt{\sum_i (r_2(i) - \mu_2)^2}}$$

Introduction to Recommender Systems

What's wrong with Spearman?

- Punishes all misplacement equally
- However: we don't care as much low-down
 - swapping 1 and 3: bad
 - swapping 11 and 13: not nearly so bad
- Goal: weight things at the top of the list more heavily

Introduction to Recommender Systems

Discounted Cumulative Gain

- Measure *utility* of item at each position in the list
- Discount by position, so things at front are more important
- Normalize by total achievable utility
- Result is Normalized Discounted Cumulative Gain (nDCG)

Introduction to Recommender Systems

Utility

- Rating
- 1/0 (for purchases/clicks/views)
 - Be careful – next lecture

Introduction to Recommender Systems

nDCG: The Formula

$$\text{DCG}(r) = \sum \text{disc}(r(i))u(i)$$

$$\text{nDCG}(r) = \frac{\text{DCG}(r)}{\text{DCG}(r_{\text{perfect}})}$$

- nDCG is in range [0,1]; 1 is perfect ranking
- Measures fraction of potential utility achieved

Introduction to Recommender Systems

Discount

- $1/(\log_2 r(i))$ is common
 - no discount for 2nd item
- Half-life utility (Breese, 1996) has good theoretical basis $2^{\frac{r(i)-1}{\alpha-1}}$

- Exponential decay
- Idea: users are exponentially less likely to click each successive item in list
- Measurement: expected utility

Introduction to Recommender Systems

Fraction of Concordant Pairs

- What fraction of pairs are in the correct relative order?
- Tests pairwise accuracy

Introduction to Recommender Systems

Conclusion

- Several metrics to measure recommender's ability to order items
- nDCG increasingly common; MRR also used
- Next lecture: problems with missing-data evaluations

Introduction to Recommender Systems

5-4: Rank Metrics