Instructions:

- Please do not edit this template, instead use “File -> Make a Copy” to save it to your workspace
- You are required to submit the your slides in pdf on gradescope by **Tuesday Dec 10, 1:00 pm**.

UC San Diego

Data Mining:
Item Recommendation Systems
Evaluating Amazon Data from SVD to BPR and T4Rec

Group 1
Joel Polizzi, Antariksha Ray, Lora Khatib, Jemin Vagadia
12/03/2024

Sec 1: Introduction

Overview of the field/topic

- Introduce the topic here

Importance of the topic

- Why this topic is important and why do people care?
- You can discuss applications etc.

Sec 2: Problem Formulation

#1 Problem formulation

- What is the problem that you are solving? Include both high level idea and mathematical formulation

#2 Relation to Numerical Linear Algebra

- How your topic is related to Numerical Linear Algebra
  - E.g. why this topic is related to numerical linear algebra? Why NLA can provide a solution?

#3 Approach of Numerical Linear Algebra (NLA)

- Details of the method in numerical linear algebra to solve the problem
  - E.g. how ppl are using Numerical Linear Algebra to solve your topic? What are the techniques could be useful?

BPR: Bayesian personalized Ranking

Bayesian Personalized Ranking

- Conditional Probability
- Bayes Theorem
- Matrix Factorization for Latent Factors & the Parameter Vector

$$P(A|B) = \frac{P(A \cap B)}{P(B)}$$

$$P(A|B) = \frac{P(A)P(B|A)}{P(B)}$$

$$p(\Theta| >_u) \propto p(>_u |\Theta)p(\Theta)$$

Bayesian Personalized Ranking

Pros

- Good for implicit feedback
- Ranking
- Scalability
- Effective for sparse data

Cons

- Noisey sampling
- Latent Factor Interpretability
- Overfitting

#4 experiment setup and result

- Make 2-3 slides on implementing NLA techniques to solve this problem
  - What is the dataset, what is the metric to evaluate success/failure
  - What are the hyper-parameters, experiment settings
  - What are the results, and how to interpret?

Sec 3: State of the Art (SOTA)

Bridge to SOTA: what are the approaches in the literature?

- Make 3-5 slides on what are the SOTA & comparisons:
  - How do people solve this problem now? Are they still using NLA approach or are they using a new approach?
  - Summarize/introduce what are the new (SOTA) approaches to solve this problem with math details
  - What are the pros and cons of each approach, and what are the remaining challenges?
  - At the end of this section, please make comparison tables between different approaches, and remaining challenges

Experiment: SOTA approach details

- Make 2-3 slides on implementing SOTA techniques to solve this problem
  - What are the hyper-parameters, experiment settings
  - What are the results, and how to interpret?
  - How does it compare to the NLA approach? Make a graph/table to compare

Sec 4: Concluding remarks

Conclusions

- Summarize your findings and thoughts
  - E.g. are newer approaches needed to solve the problem?
  - E.g. What are the limitations and un-resolving challenges?
  - Use your creativity and imagination!