Time–aware Collaborative Topic Regression: Towards Higher Relevance in Textual Items Recommendation

BIRNDL 2018, Ann Arbor

Anas Alzogbi
Motivation

- More than 100K Paper in computer science are published yearly
- 3 times more papers in 2010 than in 2000

Which papers are relevant to me!

[Onur, 2015]
Problem Definition

- Set of $n$ users $\mathcal{U}$
- Set of $m$ papers $\mathcal{I}$
- Users interactions with papers with timestamps
  - Ratings Matrix with ones at the positive interactions and zeros elsewhere

![Ratings Matrix]

- Predict the unknown Ratings

Is a positive Rating (1)

Is an unknown rating (0)
Collaborative Topic Modelling (CTR)

[Wang, 2011]

- Hybrid recommendation
- Builds on Matrix Factorization and extends it to benefit from items’ textual content
- Jointly learns items and users latent factors from the rating matrix and the document matrix

- All ratings have the same importance!
Concept drift in user interest

- User interest might change over time

- Not all ratings represent the actual user interest in the same extent
- Rating’s time should be considered when learning the user latent model!
Time in Recommender Systems

- **timeSVD++** [Koren, 2010]
  - MF approach, learns time-based biases along learning users and items latent factors
  - Predictions can be computed only for time intervals which are already seen

- **Time series model** [Liu, 2015], [Lu, 2016], [Gao, 2017]
  - Apply MF for each interval individually
  - Learn an auto-regressive model that finds the linear correlation between the intervals models
  - Extra complexity and requires a rich and long history of ratings

- **Forgetting mechanism**
  - Old ratings are either discarded or down weighted
  - A forgetting factor regulates the weights calculation
Time–aware Collaborative Topic Regression (T–CTR)

- Confidence scores in implicit feedback

\[
\arg\min_{U,V} \sum_{u \in \mathcal{U}, i \in \mathcal{I}} C_{ui}(R_{ui} - U_u^TV_i)^2 + \lambda_u \sum_{u \in \mathcal{U}} ||U_u||^2 + \lambda_v \sum_{i \in \mathcal{I}} ||V_i||^2
\]

Confidence score for \( R_{ui} \)

\[
C_{ui} = \begin{cases} 
  a & \text{if } R_{ui} = 1 \\
  b & \text{otherwise}
\end{cases} ; a \gg b > 0
\]

- Associate each rating with a confidence weight that controls the rating’s importance
Time-aware Collaborative Topic Regression (T–CTR)

- How to set the ratings confidence weights?
  - Old ratings should have less influence
  - But Users might have different dynamics, some users tend to stick longer to the same topic.

The paper age is not enough!

\[ i_2 \text{ and } i_4 \text{ should not have the same confidence weight} \]
User concept-drift score

- An individual concept drift score for each user
- How heterogeneous are the items in user ratings
- Pairwise similarity between each successive items

User Concept-drift score

\[ S_u = 1 - \text{Average Similarity} \]
Ratings confidence weights

- Ratings confidence weights are decided by
  - Rating age
  - User’s concept drift score

- Time decay function:
  \[
  W_{ui} = \frac{2}{1 + e^{S_u(T - t_{Rui})}}; \\
  T: \text{is the current time, } t_{Rui} \text{ time of rating } R_{ui}
  \]

- Forgetting factor is the user confidence score
T-CTR Model Learning and prediction

- Set the confidence Scores
  \[ C_{ui} = \begin{cases} 
  \max(W_{ui}, b) & \text{if } R_{ui} = 1 \\
  b & \text{otherwise} 
\end{cases} \]

- Learn the model parameters that maximize the log likelihood:
  \[ \mathcal{L} = - \sum_{u \in \mathcal{U}, i \in \mathcal{I}} \frac{C_{ui}}{2} (R_{ui} - U_u^T V_i)^2 - \frac{\lambda_u}{2} \sum_{u \in \mathcal{U}} \|U_u\|^2 - \frac{\lambda_v}{2} \sum_{i \in \mathcal{I}} \|(V_i - \theta_i)\|^2 \]

- Predictions:
  \[ \hat{R}_{ui} = U_u^T V_i \]
Evaluation Experiments

Dataset from citeulike

- ~ 3 K Users
- ~ 210 K Papers with titles, abstracts and keywords
- ~ 285 K Ratings
- From Nov 2004 to Dec 2007
Time-aware vs time-ignorant evaluation

Training/Test data split

<table>
<thead>
<tr>
<th>Fold</th>
<th>Split date</th>
<th>#Users</th>
<th>#Papers</th>
<th>#Ratings</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Training</td>
<td>Test</td>
<td>Training</td>
</tr>
<tr>
<td>1</td>
<td>Sep 2005</td>
<td>484</td>
<td>310</td>
<td>18,988</td>
</tr>
<tr>
<td>2</td>
<td>Mar 2006</td>
<td>978</td>
<td>564</td>
<td>46,585</td>
</tr>
<tr>
<td>3</td>
<td>Sep 2006</td>
<td>1,518</td>
<td>793</td>
<td>83,217</td>
</tr>
<tr>
<td>4</td>
<td>Mar 2007</td>
<td>2,105</td>
<td>1,023</td>
<td>122,596</td>
</tr>
<tr>
<td>5</td>
<td>Sep 2007</td>
<td>2,716</td>
<td>970</td>
<td>176,977</td>
</tr>
</tbody>
</table>
Time-aware vs time-ignorant evaluation

- CTR performance measured on two setups:
  - Time-aware evaluation
  - Time-ignorant evaluation

- Time-aware evaluation are worst but realistic
T–CTR performance

- Baselines
  - Collaborative Topic Regression (CTR) [Wang, 2011]
  - Collaborative Evolution For User Profiling (CE) [Lu, 2016]
  - Collaborative Filtering for Implicit Feedback (CF) [Hu 2008]

- Results
User–specific Concept–drift score importance

- Compare the following setups
  - Common concept–drift score (CTR–0.1, CTR–0.5, CTR–1)
  - User–specific Concept–drift score (T–CTR)

- Results
Conclusion & Future Work

**Conclusion**
- Time-aware hybrid recommender system
- Dynamically adapt to the user dynamics in interest drift
- Study on a real-world dataset
- Time-aware vs time-ignorant offline evaluations

**Future Work**
- Develop a probabilistic model that learns the users concept drift scores instead of the heuristic approach

[http://www.superscholar.org]
Thanks for your attention

Questions & Comments?
alzogha@informatik.uni-Freiburg.de
References

[Onur, 2015]

[Wang, 2011]

[Koren, 2010]

[Liu, 2015]

[Gao, 2017]

[Hu, 2008]

[Lu, 2016]