Context-aware Image Tweet Modelling and Recommendation

Tao Chen, Xiangnan He and Min-Yen Kan
School of Computing, National University of Singapore

Introduction

Research Questions
• How to represent the semantics of an image tweet?
• Are visual objects sufficient?
• How to utilize such representations for personalized image tweet recommendation task?

Our Contributions
• We propose a CITING framework to model image tweet by its contextual text
• We develop a feature-aware matrix factorization model to capture user’s personal interest
• We have released code and datasets: https://github.com/kite1988/famf

CITING: Context-aware Image Tweet Modelling

1. Hashtag enhanced text
• Conflate the variants of hashtags
  #icebucket, #ALSiceBucketChallenge
  ice bucket  ALS ice bucket challenge

2. Text in Image
• Overlaid text and screenshot of pure text

Propose filtered rules to fuse the four contextual text
• Text quality: 1>3>2>4
• Save acquisition cost by 18% and improve representation quality

Personalized Image Tweet Recommendation

Experiments

Dataset from Twitter

<table>
<thead>
<tr>
<th></th>
<th>Twitter Users</th>
<th>Retweets</th>
<th>All Tweets</th>
<th>Ratings</th>
</tr>
</thead>
<tbody>
<tr>
<td>Training</td>
<td>174,765</td>
<td>1,316,645</td>
<td>1,592,837</td>
<td></td>
</tr>
<tr>
<td>Test</td>
<td>9,021</td>
<td>77,061</td>
<td>82,743</td>
<td></td>
</tr>
</tbody>
</table>

Evaluation
• For each user, keep the recent 10 retweets as test set and the rest as training set
• Report mean average precision (MAP) and precision at top positions

<table>
<thead>
<tr>
<th>Method</th>
<th>Features</th>
<th>P@1</th>
<th>P@3</th>
<th>P@5</th>
<th>MAP</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 Random</td>
<td></td>
<td>0.115**</td>
<td>0.115</td>
<td>0.115</td>
<td>0.156**</td>
</tr>
<tr>
<td>2 Length</td>
<td>Post’s text</td>
<td>0.176**</td>
<td>0.158</td>
<td>0.150</td>
<td>0.173**</td>
</tr>
<tr>
<td>3 Profiling</td>
<td>Post’s text</td>
<td>0.336**</td>
<td>0.227</td>
<td>0.197</td>
<td>0.202**</td>
</tr>
<tr>
<td>4 FAMF</td>
<td>Visual objects</td>
<td>0.211**</td>
<td>0.205</td>
<td>0.192</td>
<td>0.211**</td>
</tr>
<tr>
<td>5 FAMF</td>
<td>Post’s text</td>
<td>0.359*</td>
<td>0.325</td>
<td>0.287</td>
<td>0.275**</td>
</tr>
<tr>
<td>6 FAMF</td>
<td>Non-filtered context</td>
<td>0.413</td>
<td>0.352</td>
<td>0.319</td>
<td>0.296</td>
</tr>
<tr>
<td>7 FAMF</td>
<td>CITING</td>
<td>0.419</td>
<td>0.355</td>
<td>0.319</td>
<td>0.298</td>
</tr>
<tr>
<td>8 FAMF</td>
<td>CITING + Visual objects</td>
<td>0.425</td>
<td>0.350</td>
<td>0.313</td>
<td>0.298</td>
</tr>
</tbody>
</table>

• 4-8 vs. 1-3: FAMF is effective in modelling user interest
• 4 vs. 5: Visual objects are not sufficient to model Twitter image’s semantics
• 7 vs. other: Our CITING text significantly outperforms the others
• 7 vs. 6: The filtered fusion rules improve text quality
• 8 vs. 7: The further incorporation of visual objects does consistently improve the performance

Tao Chen: chentaokite@gmail.com

ACM MM 2016, Amsterdam, The Netherlands