ATinstagram

Mason Barden
Steve Cho
Nicholas Halstead
Tashi Jeshong and
Zubin Joseph

CS 4624 Multimedia
Hypertext & Information Access
Edward Fox
Virginia Tech, Blacksburg VA 24061
May 9, 2022
Outline

● Main Goal/Deliverables
● Timeline
● Project Results
  ○ Sentiment Analysis
  ○ Geo Locations
  ○ Visualizations
● Lessons Learned/Future Plans
● Acknowledgements
● References
Main Goal/Deliverables

- Visualization(s) of analysis on Leave No Trace (LNT) related posts
- Quantitative analysis of distribution of hashtags
## Timeline

<table>
<thead>
<tr>
<th>January</th>
<th>February</th>
<th>March</th>
<th>April</th>
</tr>
</thead>
<tbody>
<tr>
<td>• Initiated data scraping and cleaning</td>
<td>• Completed data scraping and cleaning</td>
<td>• Collected geo-locations of posts</td>
<td>• Completed sentiment analysis</td>
</tr>
<tr>
<td>• Gathered requirements from sponsors</td>
<td>• ~120k posts collected</td>
<td>• Started working on sentiment analysis</td>
<td>• Created visualizations based off collected data</td>
</tr>
</tbody>
</table>
Given sentence: This guy is extremely smart and handsome, but not funny.

Words in lexicon list and their valence scores are:
- smart: 1.7
- handsome: 2.2
- funny: 1.9

<Words with ALL CAPS will have their intensity increased by adjusting their scalar>

Negation words used:
- not

<These negation words are considered when calculating valence scores>

Booster words used:
- extremely

<These booster words are considered when calculating overall sentiment score>

Counting all exclamation points and question marks...
The total value of intensity scale applied from emphasis amplifiers: 0.000
Sentiment Analysis Format (2)

1. Each word in the sentence score with necessary scalar and intensity applied: The booster and negation words are not considered for valence scores.
   - This: 0.000
   - guy: 0.000
   - is: 0.000
   - extremely: 0.000
   - smart: 0.996
   - and: 0.000
   - handsome: 1.232
   - but: 0.000
   - not: 0.000
   - funny: -2.109

2. Normalizing the total score...
3. Computing positive, negative, and neutral scores...
   - neg score was: 0.217
   - neu score was: 0.488
   - pos score was: 0.295
   - compound score was: 0.031

The sentence overall had neutral sentiment with compounding score 0.031

Sentiment scores of each word:
1. Positive, negative, and neutral scores
2. Total score
Sentiment Analysis Results

<table>
<thead>
<tr>
<th></th>
<th>#Int OR #leavenotrace</th>
<th>#Int OR #leavenotrace AND #leavenothingbutfootprints</th>
</tr>
</thead>
<tbody>
<tr>
<td>Positive Posts</td>
<td>27016</td>
<td>691</td>
</tr>
<tr>
<td>Slightly Positive Posts</td>
<td>6503</td>
<td>137</td>
</tr>
<tr>
<td>Neutral Posts</td>
<td>14209</td>
<td>411</td>
</tr>
<tr>
<td>Slightly Negative Posts</td>
<td>1719</td>
<td>53</td>
</tr>
<tr>
<td>Negative Posts</td>
<td>1386</td>
<td>68</td>
</tr>
</tbody>
</table>
## Sentiment Analysis Results

<table>
<thead>
<tr>
<th></th>
<th>#atsobo2021</th>
<th>#atnobo2021</th>
<th>#atclassof2021</th>
<th>#appalachiantrail</th>
</tr>
</thead>
<tbody>
<tr>
<td>Positive Posts</td>
<td>48</td>
<td>73</td>
<td>89</td>
<td>796</td>
</tr>
<tr>
<td>Slightly Positive Posts</td>
<td>15</td>
<td>15</td>
<td>24</td>
<td>163</td>
</tr>
<tr>
<td>Neutral Posts</td>
<td>35</td>
<td>185</td>
<td>43</td>
<td>594</td>
</tr>
<tr>
<td>Slightly Negative Posts</td>
<td>8</td>
<td>1</td>
<td>3</td>
<td>63</td>
</tr>
<tr>
<td>Negative Posts</td>
<td>6</td>
<td>6</td>
<td>2</td>
<td>45</td>
</tr>
</tbody>
</table>
Geo-Location of Posts
Zoomed in on Appalachian Trail
Visualization Results

Top 100 most frequently occurring tag combinations from all of the data collected.
Visualization Results (2)

Top 150 most frequently occurring tag combinations from all of the data collected.
Lessons Learned

- Maintaining time management skills and communication
- Planning
- Add time for debugging
- Comment code for future users
Future Plans

- Use something other than Selenium to scrape Instagram data
- Use a different sentiment analysis technique and compare results
Acknowledgements

CLIENTS

Primary Contact:

- Computer Science Ph.D. student Morva Saaty, morvasaaty@vt.edu

Other Project Members:

- Computer Science Professor, Scott McCrickard
- School of Urban & International Affairs, Kris Wernstedt
- School of Urban & International Affairs, Shalini Misra
- USGS & Department of Forest Resources, Jeff Marion