alp_2018_identifying_topical_influencers_on_twitter_based_on_user_behavior_and_network_topology

Year

2018

Author(s)

Zeynep Zengin Alp, Şule Gündüz Öğüdücü

Title

Identifying topical influencers on twitter based on user behavior and network topology

Venue

Knowledge-Based Systems

Topic labeling

Manual

Focus

Secondary

Type of contribution

Established approach

Underlying technique

Manual labeling

Topic labeling parameters

\

Label generation

Three human experts analyzed the output of LDA, which is the word clusters representing each topic.
Afterwards, they identified coherent topics which contain words that are semantically similar.
Finally, they labeled the 6 coherent topics with appropriate topic titles (“Politics / Breaking News”, “Spiritual”, “Religion”, “Social Responsibility”, “Soccer / Sports”, and “TV / TV Shows”).

Motivation

After classifying tweets with topics, topical networks were formed (For each topic, a sub-network is generated where nodes represent the users who post on the specific topic and edges represent the following relationship).
For each user, percentage of topical tweets are calculated and topical labels were assigned to users according to topical contents of their tweets.

Topic modeling

LDA

Topic modeling parameters

Nr of topics (k): 20

Nr. of topics

6 (out of the 20 identified as coherent)

Label

Manually assigned single or multi-word labels identifying the broad categories of the 6 topics

Label selection

\

Label quality evaluation

\

Assessors

\

Domain

Paper: Social media analysis
Dataset: Social media (Twitter)

Problem statement

Since Twitter is a good platform to spread knowledge as a word of mouth approach and it has many more public profiles than protected ones, it is a target media for marketers.
In this paper, we introduce a novel methodology, called Personalized PageRank, that integrates both the information obtained from network topology and the information obtained from user actions and activities in Twitter.
The proposed approach aims to determine the topical influencers who are experts on a specific topic.

Corpus

See: `alp_2019_influence_factorization_for_identifying_authorities_in_twitter`

Document

See: `alp_2019_influence_factorization_for_identifying_authorities_in_twitter`

Pre-processing

- Tokenization
- Stemming
- Removal of stop-words, punctuation, and mentions that have no meaning for topic modeling

---

```
@article{alp_2018_identifying_topical_influencers_on_twitter_based_on_user_behavior_and_network_topology,
  abstract = {Social media web sites have become major media platforms to share personal information, news, photos, videos and more. Users can even share live streams whenever they want to reach out to many other. This prevalent usage of social media attracted companies, data scientists, and researchers who are trying to infer meaningful information from this vast amount of data. Information diffusion and maximizing the spread of words is one of the most important focus for researchers working on social media. This information can serve many purposes such as; user or content recommendation, viral marketing, and user modeling. In this research, finding topical influential/authority users on Twitter is addressed. Since Twitter is a good platform to spread knowledge as a word of mouth approach and it has many more public profiles than protected ones, it is a target media for marketers. In this paper, we introduce a novel methodology, called Personalized PageRank, that integrates both the information obtained from network topology and the information obtained from user actions and activities in Twitter. The proposed approach aims to determine the topical influencers who are experts on a specific topic. Experimental results on a large dataset consisting of Turkish tweets show that using user specific features like topical focus rate, activeness, authenticity and speed of getting reaction on specific topics positively affects identifying influencers and lead to higher information diffusion. Algorithms are implemented on a distributed computing environment which makes high-cost graph processing more efficient.},
  author = {Zeynep {Zengin Alp} and {\c S}ule {G{\"u}nd{\"u}z {\"O}{\u g}{\"u}d{\"u}c{\"u}}},
  date-added = {2023-03-17 14:18:13 +0100},
  date-modified = {2023-03-17 14:18:13 +0100},
  doi = {https://doi.org/10.1016/j.knosys.2017.11.021},
  issn = {0950-7051},
  journal = {Knowledge-Based Systems},
  keywords = {Social influence, Topical influence, Twitter, User modeling, PageRank, Topic modeling, LDA, Evaluation metrics, Viral marketing},
  pages = {211-221},
  title = {Identifying topical influencers on twitter based on user behavior and network topology},
  url = {https://www.sciencedirect.com/science/article/pii/S0950705117305439},
  volume = {141},
  year = {2018}}
```

#Thesis/Papers/Initial