

Predicting Pinterest: Organising the World's Images with Human-machine Collaboration

Nishanth Sastry
King's College London, UK
nishanth.sastry@kcl.ac.uk

ABSTRACT

The user generated content revolution has created a glut of multi-media content online – from Flickr to Facebook, new images are being made available for public consumption everyday. In this talk, we will first explore how, on sites such as Pinterest, users are bringing order to this burgeoning collection by manually curating collections of images in ways that are highly personalised and relevant to their own use. We will then discuss the phenomenon of social bootstrapping, whereby existing mature social networks such as Facebook are helping bootstrap engaged communities of content curators on external sites such as Pinterest. Finally, we will demonstrate how the manual effort involved in curation can be amplified using a unique human-machine collaboration: By treating the curation efforts of a subset of users on Pinterest as a distributed human computation over a low-dimensional approximation of the content corpus, we derive simple yet powerful signals, which, when combined with image-related features drawn from state-of-the-art deep learning techniques, allow us to automatically and accurately populate the personalised curated collections of all other users.

Categories and Subject Descriptors

H.3.5 [Online Information Services]: Commercial Services, Data Sharing, Web-based services; I.4.9 [Image Processing and Computer Vision]: Applications; J.4 [Social and Behavioral Sciences]: Sociology

Keywords

Content Curation; Pinterest; Deep learning; User behaviours; Image analysis; Supervised learning; Crowdsourcing; Social Bootstrapping; Friend Finder Tools; Community Design; Social Property; Social Interaction; Copied Networks

1. ACKNOWLEDGEMENTS

This research was partially supported by the UK Engineering and Physical Sciences Research Council, the Economic and Social Sciences Research Council, and by the Henry Lester Trust.

2. REFERENCES

- [1] CAPPELLETTI, R., AND SASTRY, N. IARank: Ranking users on Twitter in near real-time, based on their information amplification potential. In *Proceedings of The 2012 ASE International Conference on Social Informatics* (Washington D.C., USA., 2012).
- [2] SASTRY, N. How to tell head from tail in user-generated content corpora. In *Proceedings of The 6th International AAAI Conference on Weblogs and Social Media (ICWSM)* (Dublin, Ireland, June 2012).
- [3] SASTRY, N. Crowdsourcing and social networks. In *Encyclopedia of Social Network Analysis and Mining*. Springer-Verlag, 2014, pp. 316–318.
- [4] ZHONG, C., KARAMSHUK, D., AND SASTRY, N. Predicting pinterest: Automating a distributed human computation. In *Proceedings of the 24th international conference on World wide web* (Florence, Italy, May 2015), International World Wide Web Conferences Steering Committee.
- [5] ZHONG, C., SALEHI, M., SHAH, S., COBZARENCO, M., SASTRY, N., AND CHA, M. Social bootstrapping: how pinterest and last. fm social communities benefit by borrowing links from facebook. In *Proceedings of the 23rd international conference on World wide web* (2014), International World Wide Web Conferences Steering Committee, pp. 305–314.
- [6] ZHONG, C., SHAH, S., SUNDARAVADIVELAN, K., AND SASTRY, N. Sharing the loves: Understanding the how and why of online content curation. In *Proceedings of The 6th International AAAI Conference on Weblogs and Social Media (ICWSM)* (Boston, USA, Jul 2013).

Copyright is held by the International World Wide Web Conference Committee (IW3C2). IW3C2 reserves the right to provide a hyperlink to the author's site if the Material is used in electronic media.
WWW 2015 Companion, May 18–22, 2015, Florence, Italy.
ACM 978-1-4503-3473-0/15/05.
<http://dx.doi.org/10.1145/2740908.2744719>.