Part III Individual Project Brief

Enhanced Content Analysis and Information Retrieval Using Twitter Hashtags

Author:
Jamie Davies
jagd1g11@ecs.soton.ac.uk

Supervisor:
Dr. Nicholas Gibbins
nmg@ecs.soton.ac.uk

Microblogging services such as Twitter[1] have become very popular in recent years. One of the key characterists of such services is the use of 'hashtags' — topical/categorical annotations provided by the authors of the posts (tweets) themselves. These hashtags have been proven to be useful and information-rich in previous microblogging research projects; from using them to crowd-source real-time event detection [2], to using them to train sentiment classifiers [1]. However due to their nature, there is a vast and diverse collection of hashtags for users to choose from, resulting in the information provided by hashtags not being as accurate or as complete as it could be.

References

[1] D. Davidov, O. Tsur, and A. Rappoport. Enhanced sentiment learning using twitter hashtags and smileys. In *Proceedings of the 23rd International Conference on Computational Linguistics: Posters*, COLING '10, pages 241–249, Stroudsburg, PA, USA, 2010. Association for Computational Linguistics.

[2] T. Sakaki, M. Okazaki, and Y. Matsuo. Earthquake shakes twitter users: real-time event detection by social sensors. In *Proceedings of the 19th international conference on World wide web*, WWW '10, pages 851–860, New York, NY, USA, 2010. ACM.

[1] http://www.twitter.com