



Selected Papers of #AoIR2018:
The 19th Annual Conference of the
Association of Internet Researchers
Montréal, Canada / 10-13 October 2018

LOOK WHO'S TALKING: USING HUMAN CODING TO ESTABLISH A MACHINE LEARNING APPROACH TO TWITTER EDUCATION CHATS

K. Bret Staudt Willet
Michigan State University

Brooks D. Willet
Birch Wayfinders, LLC

Introduction

Twitter has become a hub for numerous organized conversations related to education. Twitter participants contribute very short posts (i.e., limited to 280 text characters or less) called *tweets* which can be indexed and tracked with *hashtags*—a word or phrase preceded by a hash (i.e., “#”) symbol. There are many different types of educational conversations on Twitter, denoted by hashtags and organized by affinities including geography (e.g., #miched for the U.S. state of Michigan), academic subject (e.g., #sschat for social studies education), and school level (e.g., #elemchat for elementary education). Researchers have used a number of theoretical frameworks to conceptualize what is happening in and through these educational conversations, describing them as *communities of practice* (e.g., Britt & Paulus, 2016) and *professional learning networks* (e.g., Carpenter & Krutka, 2014)—sites for teacher professional development (Xing & Gao, 2018).

Although some phenomena, such as the volume of tweets posted with an educational hashtag, are easy to measure, it is difficult to analyze the characteristics of tweet content that would be expected as evidence of a community of practice or professional development. Toward this end, we studied #Edchat—one of the oldest and busiest Twitter educational hashtags—to examine the content of tweets for evidence of professional purposes.

Suggested Citation (APA): Staudt Willet, K. B. & Willet, B. D. (2018, October). *Looks who's talking: Using human coding to establish a machine learning approach to Twitter education chats*. Paper presented at AoIR 2018: The 19th Annual Conference of the Association of Internet Researchers. Montréal, Canada: AoIR. Retrieved from <http://spir.aoir.org>

Method and Results

In a prior study, the first author used a Twitter Archiving Google Sheet (Hawksey, 2014) to collect tweets containing the text “#edchat” from October 1, 2017 to June 5, 2018, resulting in a dataset of 1,228,506 unique tweets from 196,263 different contributors. Starting with the a priori categories used in Carpenter and Krutka’s (2014) survey study on *how* and *why* educators use Twitter for professional purposes—which included (a) sharing and acquiring resources, (b) collaborating with other educators, (c) networking, (d) giving and receiving emotional support, (e) communicating with students, (f) communicating with parents, (g) providing in-class activities for students, (h) providing out-of-class activities for students, (i) participating in Twitter chats, (j) backchanneling, and (k) everything else (i.e., “other”)—the first author conducted three stages of human-coded content analysis. This qualitative work produced a final codebook, which the first author used to sort a stratified random sample of 1,000 tweets into four emergent, inductive categories: tweets demonstrating evidence of different professional purposes related to (a) self, (b) others, (c) mutual engagement, and (d) everything else. Purposes related to *self* included self-promotion and establishing reputation; purposes related to *others* included increasing the visibility of peers as well as sharing content and tips; purposes related to *mutual engagement* included networking, collaborating, disagreeing, and providing emotional support; and the *everything else* category included just that: tweets whose purposes were unclear, neutral, didactic, or off-topic.

Overall, we found about 65% of the tweets in our #Edchat sample demonstrated purposes related to others, about 25% demonstrated purposes related to self, and less than 4% of tweets demonstrated purposes related to mutual engagement. Thus #Edchat could be considered a good conversation space to establish one’s own reputation or discover resources from others, but not an ideal space for meaningful dialogue. These #Edchat findings are not generalizable to all of Twitter; #Edchat is just one of many educational conversations. Blumengarten, Hamilton, Murray, Evans, and Rochelle (n.d.) maintain a list that, as of August 2018, contained 339 different Twitter education chats. To compare educators’ purposes for contributing to these different conversations, we need a better approach. Our initial method was too time intensive—it would be untenable to collect tweets from 339 hashtags and conduct human-coded content analysis of a random sample from each hashtag. Therefore, we are developing a scalable computational model.

We used the *caret* R package (Kuhn, 2018) to build a multiclass logistic regression classifier to categorize tweets into one of the four categories from our earlier work: purposes related to (a) self, (b) others, (c) mutual engagement, and (d) everything else. During the initial study, we found that certain machine-coded tweet types (e.g., original posts, retweets, self-retweets, modified tweets, “via” tweets, “thanks” tweets, replies, extended posts, directed posts, and self-referential posts) and certain keywords (e.g., “should,” “daily,” “worth,” “how to,” “tips”) tended to be associated with certain professional purposes. These observations provided an initial set of features that we added to an input matrix for our classifier. We also identified features related to tweet content, such as sentiment and word count, as determined by Rinker’s (2018) *sentimentr* R package; and hashtags, hyperlinks, and images. Other features were related to tweet metadata, such as retweets, replies, and likes, as well as keywords in

the tweeter's profile. We divided our 1,000 previously coded tweets into a training set (n = 600), a development set (n = 250), and a test set (n = 150).

Conclusion

The anticipated product of this work—a successful, generalizable machine learning model—would help educators and researchers quickly evaluate Twitter educational hashtags to determine where they might want to engage. For example, pre-service teachers—with specific, contextual questions—might want to participate in a conversation whose purposes are more aligned with *mutual engagement*; in contrast, a mid-career in-service teacher—seeking to establish their reputation and expand their network—might prefer a conversation whose purposes tend more toward *self* or *others*. Participants are already selecting Twitter educational hashtags based on affinities such as subject area and geography; our work will allow them to also factor in the observed purposes of these conversations.

Our machine learning model will improve the understanding and evaluation of the purposes demonstrated in each of the 339 known Twitter education conversations. Although machine learning techniques are still rare in educational research, studies such as Xing and Gao's (2018) work have begun to demonstrate the utility of data mining in our field. We believe this current project will similarly contribute new methods worthy of consideration by educational researchers.

References

- Blumengarten, J., Hamilton, C., Murray, T., Evans, C., & Rochelle, J. (n.d.). Twitter education chats [Web resource]. Retrieved from <https://sites.google.com/site/twittereducationchats/education-chat-official-list>
- Britt, V. G., & Paulus, T. (2016). "Beyond the four walls of my building": A case study of #Edchat as a community of practice. *American Journal of Distance Education*, 30, 48-59.
- Carpenter, J. P., & Krutka, D. G. (2014). How and why educators use Twitter: A survey of the field. *Journal of Research on Technology in Education*, 46, 414-434.
- Hawksey, M. (2014). TAGS: Twitter Archiving Google Sheet (Version 6.1) [Computer software]. Retrieved from <http://tags.hawksey.info>
- Kuhn, M. (2018). caret: Classification and regression training (Version 6.0-80) [R package]. Retrieved from <https://CRAN.R-project.org/package=caret>
- Rinker, T. (2018). sentimentr: Calculate text polarity sentiment (Version 2.3.2) [R package]. Retrieved from <https://CRAN.R-project.org/package=sentimentr>
- Xing, W., & Gao, F. (2018). Exploring the relationship between online discourse and commitment in Twitter professional learning communities. *Computers & Education*, 126, 388-398.