Position paper for Dr. Andrew Pilny

The purpose of this position paper is to describe some of my primary research interests, particularly to as it relates to the current workshop theme related to "*Analyzing Social Perception and Amplification using Social Media and Big Data in Human Dynamics* ". In particular, two of my ongoing projects have relevance towards the subtopic theme of "The linguistic analysis and geospatial visualization of public health data or social media data related to disaster events or health disparities." Although the topics discussed below are not currently related to crises and disasters, the techniques described can be easily amended to focus around these important issues.

Hybrid approach to automated content analysis

According to Krippendorf (2019), content analysis, defined as a "research technique for making replicable and valid inferences from texts (or other meaningful matter) to the contexts of their use" (p. 24), is one of the backbones of social science research given its emphasis on both information *and* meaning. One of the most obvious drawbacks of content analysis is the time, effort, and money needed complete a reliable and valid study. The goal of this project is to develop a hybrid procedure for ACA. By hybrid, we simply mean a methodology that includes both humans and computers in the mix.

The first step of developing a hybrid ACA is to describe the inference, or in other words, the topic or interpretation of various texts. For this project, we focused on *relational uncertainty* and mined data from social media (i.e., blog posts) to see if we can develop a coding scheme to verify if the post was describing either a (1) uncertain relationship, (2) certain relationship, or (3) no relationship specified.

Two coders were trained to code social media posts in line with the three above categories. Two social media sites, where individuals disclose aspects of their relationship, were randomly sampled. The first sample is the *training set*, where we primarily use it to train both the two humans and the computer. These 100 articles were tested to establish reliability for the two human coders. For these two coders, reliability was easily established ($\alpha = .97$). The text from each social media post was analyzed using Linguistic Inquiry and Word Count (LIWC). The data from LIWC was used to see if a machine can be trained to associate patterns from LIWC with the ground truth of relationship uncertainty (i.e., the three codes).

Another 100 random posts were coded by two human coders. This is the *validation set* to see if the computer can learn from the training set to provide predictions on the validation set. These predicted classes from the machine (we used a random forest machine learning algorithm) were used as a "third" coder alongside the two human coders. When all three entities (two humans, one machine) coded the validation set, established reliability was still achieved ($\alpha = .84$).

Finally, we needed to establish some sort of validity test to see if the machine can code from a different set of text and still provide valid predictions. To test this, we created an online survey via MTurk where we had participants write about their relationship in text

and complete a relational uncertainty measurement scale (n = 212). The machine analyzed these text using the rules it established from the first training set. Using the predicted probabilities of uncertainty, we then correlated the two. There were significant correlations between three dimensions of relational uncertainty: mutuality (r = -.18, p = .009), definition (r = -.15, p = .033), and future (r = -.21, p = .002). Moreover, when the machine prediction high levels of certainty from the text, the correlations were slightly higher: norms (r = .20, p = .003), mutuality (r = .30, p = .000), definition (r = .23, p = .001), and future (r = .29, p = .000).

Our results suggest that a hybrid automated content analysis procedure is very much possible. We found evidence of reliability and validity with the current example of relational uncertainty. To us, this suggests that a hybrid procedure is certainly possible across multiple contexts, including crises and disasters.

Predicting fake news

We used a similar approach to predicting fake news articles to demonstrate that deconstructing text with LIWC can provide promising data that can be used in predictive models across different contexts. Training, validation, and test sets were created too.

The training set derived from To gather articles resembling closer to fake, rather than real news, we used a dataset compiled by an early version of the B.S. Detector - a Chrome browser extension application developed by Daniel Sieradski. The application tagged 244 websites as “bullshit” from October 25, 2016 until November 26, 2016. A total of 12,999 articles and metadata were extracted from those websites. According to the source, there were no “genuine, reliable, or trustworthy news sources represented in this dataset (so far)” (https://github.com/demidovakatya/competitions/tree/master/fake-news). We randomly selected 2000 articles from the dates October 25, 2016 through November 2, 2016 for our training set. Our validation set contained a random sample of 750 articles from November 3, 2016 until November 7, 2016. To compare the fake news dataset with a real news dataset, we took advantage of traditional news outlets from the Lexis Nexis archive, extracting an identical training set of 2000 articles (October 25, 2016 - November 2, 2016) about politics from the Washington Post (n = 888), New York Times (n = 229), National Public Radio (n = 42), CNN.com (n = 411), USA Today (n = 38), Los Angeles Times (n = 97), and the BBC (n = 295). A validation set with a similar distribution of 750 articles from November 3, 2016 until November 7, 2016 was extracted as well. Overall, the machine learning algorithm was 81% accurate in predicting fake versus real news articles.

A test set was developed to see if the same algorithm can predict post-election fake news articles. Here, we crawled articles from Snopes.com, which documents fake news from January 1, 2017 to January 1, 2018. 178 validated fake news stories by Snopes were analyzed along with 178 randomly sampled political news stories from LexisNexis. Similarly, the machine learning algorithm was 79% accurate from using the same rules in the training set. Overall, the results from both ongoing studies show that using text data from social media and online content can successfully be used to both predict and better understand human dynamics.