"A NLP pipeline to automatically predict human behavioral responses to viral epidemics like Ebola"

[1,2] *Donath E*, [1,3]*Hebert-Dufresne L*, [1,3]*Cheney N.*

[1] *University of Vermont*, [2] *University of Vermont Medical Center*, [3] *University of Vermont Dept of Comp Science*

Background.

During the 2014-15 West African Ebola outbreak, communities were confronted with a variety of challenges relating how to manage disease transmission. In Sierra Leone, the Social Mobilization Action Consortium (SMAC) trained and supported a network of 2500 community mobilisers that used a set of structured tools to facilitate community engagement. There would be an initial visit, termed a triggering event, to encourage communities to take their own actions to prevent disease transmission. They would identify community "champions" and support communities in their development of plans focused on curbing Ebola transmission by normalizing risk-reduction behaviors. After this initial triggering event, subsequent visits were made to follow up on progress of these plans and provide additional support. We're interested in examining free-text unstructured responses provided by champions to a specific set of questions posed by mobilisers. These responses can be cross-examined with metrics that provide a window into how a given region managed the Ebola crisis. Little is known about the extent to which advanced natural language processing (NLP)-based techniques can be used to depict, in real-time, a given region's handling of pandemics. We'd like to investigate whether there is some pattern, examined with a variety of these techniques, embedded within text data that may be correlated with a given region's response to and handling of Ebola. Our aim is to elucidate whether they can be used to examine a community's behavioral tendencies when confronted with a public health challenge.

Methods.

Data were obtained from reports made by SMAC community mobilisers to about 12 000 communities in Sierra Leone. Mobilizers underwent training in data collection and submission via standardized forms. The reports included open-ended responses to questions about the community's Ebola-related concerns at the time of visit (as part of a broader questionnaire that also included non-text based responses). Information collected in the reports also included geographical identifiers, numbers of individuals participating in the community visits, numbers of seriously sick and deceased individuals in the past week, and numbers of prompt referrals to treatment and safe burials in the past week. For each report, three community-evaluated metrics (i.e. mortality rate, referral rate and safe burial rate) form the basis of how we evaluate a given region's performance in handling the Ebola crisis. Geographical identification occurred at four separate hierarchical levels (from the top down, these were "district", "chiefdom", "community" and "section") - this dataset contained responses and metrics from 190 chiefdoms in Sierra Leone. A variety of text classification strategies were deployed including more traditional machine learning-based techniques as well as more advanced techniques incorporating deep learning strategies and transformer models. The models built around this corpus of unstructured text would be used to predict how well a given region performed. Outcomes of interest were evaluated as both a binary outcome as a classification task (separating the regions by a median into those above and below the 50% threshold) and then as a continuous outcome. All models were built on a training set and then outcomes were evaluated on the test set in a 80:20 distribution. Outcome evaluation metrics included F1 score and AUC-ROC (area-under-the-curve of receiver-operating curve). All analyses were done using Python. Traditional machine learning models included a bag of words (BOW)/term-frequency-inverse-document frequency (TFIDF) approach followed by multivariate logistic regression, random forests and support vector machines, whereas more advanced transformer-based techniques revolved around the BERT (bidirectional encoder representations from transformers) architecture.

Results.

There were 7,898 free text responses obtained during the trigger period that formed the basis of this analysis. These text responses were used as predictors to examine which chiefdoms successfully managed the epidemic (measured by percentage change in safe burial rate, mortality rate and referral rate during "trigger" period compared to follow-up). AUC-ROC (examined across several text responses) ranged from 0.743-0.787 for safe burial rate, 0.734-0.811 for mortality rate and 0.732-0.801 for referral rate. In general, model performance was improved by focusing particularly on the question related to obstacles and levels of concern with carrying out their "action plan". Model improvements were achieved by concatenating all the response together (AUC-ROC improved significantly in this scenario, ranging from 0.858-0.881 across the three metrics of interest). Furthermore, by isolating the most and least "successful" regions (and ignoring those in the middle 25-75%), we were able to maximize performance and achieve an AUC-ROC of 0.917 on the mortality rate, 0.922 on the referral rate and 0.932 on the safe burial rate. Performance was similar across both transformer-based and traditional ML-based TFIDF/BOW models. Word combinations that included "veronica buckets", "washing material" and "hand washing" were consistently related to higher-performing regions while "shaking hands", and "gathering of strangers" were seen as correlated with poorer performing regions.

Conclusion.

This research suggests there may be some predictive value in examining unstructured free text data from publicly accessible data streams gathered from communities confronting a public health crisis.