Catching Zika Fever: Application of Crowdsourcing and Machine Learning for Tracking Health Misinformation on Twitter

Amira Ghenai
University of Waterloo
Waterloo, Canada
Email: aghenai@uwaterloo.ca

Yelena Mejova
Qatar Computing Research Institute
HBKU, Doha, Qatar
Email: ymejova@hbku.edu.qa

***Abstract*—In February 2016, World Health Organization declared the Zika outbreak a Public Health Emergency of International Concern. With developing evidence it can cause birth defects, and the Summer Olympics coming up in the worst affected country, Brazil, the virus caught fire on social media. In this work, we use Zika as a case study in building a tool for tracking the misinformation around health concerns on Twitter. We collect more than 13 million tweets regarding the Zika outbreak and track rumors outlined by the World Health Organization and Snopes fact checking website. The tool pipeline, which incorporates health professionals, crowdsourcing, and machine learning, allows us to capture health-related rumors around the world, as well as clarification campaigns by reputable health organizations. We discover an extremely bursty behavior of rumor-related topics, and show that, once the questionable topic is detected, it is possible to identify rumor-bearing tweets using automated techniques.**

I. Introduction

The information overload poses serious challenges to public health, especially with regards to infectious diseases. Similar to people's increased mobility, the availability and ubiquity of information facilitates the transmission of misinformation that can hamper the efforts to tackle a major public health crisis. With a continuous threat of digital "wildfires" of misinformation [1], health rumors are a worldwide serious problem [2].

The complexity of dealing with communication during a health crisis is growing, as social media is playing a more prominent role. Social media, compared with traditional media, is harder to monitor, track and analyze. Public health institutions such as the World Health Organization (WHO) include social media as a crucial part in monitoring a health crisis [3]. However, guidelines and tools on best approaches to tackle this are not yet available.

This paper proposes a suite of tools for tracking health-related misinformation, and describes a case study of tracking a health crisis, as discussed on Twitter. We provide a methodology for uncovering the streams of tweets spreading rumors about the 2016 Zika outbreak. In particular, the suspected link between Zika and serious brain malformation (microcephaly) in newborns, as well as the approaching Olympic Games in Rio introduced urgency in effective communication about the crisis. In particular, we track rumors outlined by the WHO (along with Snopes.com[1]) in the stream of nearly 13 million tweets. We employ high-precision expert-led Information Retrieval approach to identifying the relevant tweets. Using crowdsourcing, we distinguish between rumor and clarification tweets, which we then use to build automatic classifiers. Here, we present in-depth temporal analysis of the found rumors, their origins, and interactions with informational sources.

II. Related Work

WHO's white paper on Risk communication in the context of Zika virus urges to develop communication resources that quickly transform information into usable and easily understood format [4]. Although earlier works have found misinformation to be part of the Twitter chatter related to flu [5], little work thus far concentrated on detecting and tracking health-related rumors. Recently, Kostkova et al. [6] created the "VAC Medi+ board" online interactive visualization framework integrating heterogeneous real-time data streams with Twitter data. They track the spread of vaccine related information on Twitter and the sources of information spread. A potential framework to engaging expert knowledge in a real-time crisis, including health-related, situation is described in Imran et al. [7], where content is selected to be annotated via crowdsourcing into predefined classes. These can then be used to train a classifier, and update it as necessary with active learning data selection. Work most relevant to the current study is by Dredze et al. [8], who analyzed the characteristics of nonscientific claims about vaccine misconceptions by the vaccine refusal community. Specifically, the authors analyzed the two most prominent misleading theories about Zika vaccination in Twitter using unspecified "supervised machine learning technique" and observed the effect of vaccine-skeptic communities over other users' vaccination opinion. While [8] look at two Zika vaccine related memes, in this work we propose a more general methodological pipeline to track health-related rumors.

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

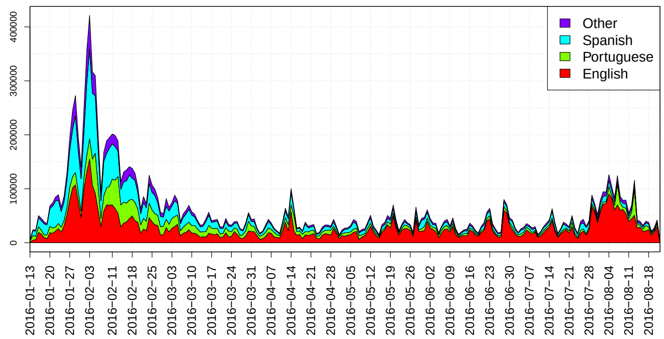

Fig. 1. Zika-related Twitter data volume, separated by language.

III. Data Collection

The data was collected using the Artificial Intelligence for Disaster Response (AIDR)[2] platform, which taps into Twitter Streaming Application Program Interface (API). The keywords contained the following (searched as quotes): zika, microcefalia, microcephaly, #zika, zika virus, *Aedes*, zika fever, *Spondweni* virus, *Aedes albopictus*, *maculopapular* rash. We aimed to cover both everyday wording as well as medical jargon which may be associated with the topic. Further, "zika" word is used by all English, Portuguese, and Spanish (the major languages of populations affected). The resulting collection of 13,728,215 tweets spans January 13 - August 22, 2016 (1), and includes the peak of interest in Zika (early February) and the Olympic Games in Brazil (August 5-21).

Since no language restriction was imposed during data collection (besides some bias English keywords introduced), we captured a plurality of languages, with three dominant ones which represent more than half of the dataset – English (6.2M), Spanish (3.7M) and Portuguese (1.5M tweets). In this work we focus on the English data, which comprised 46% of all tweets, and included 1.3M distinct users. Geo-locating these tweets using World Borders API[3] (for geo-coordinates) and Yahoo Placemaker API[4] (for user Place fields), we succeeded in locating 68% of the tweets, with the top locations being USA, UK, India, Canada, Nigeria, and Brazil, indicating a highly international data.

IV. Rumor Selection

We chose the WHO website as an authority for detecting and verifying rumors about Zika, which provided a listing of major international rumors and misinformation about the virus. At the time of writing, WHO website [9] listed 8 statements debunking ongoing rumors. Out of these, 4 were unsuitable, as they were not topically cohesive. Additionally, we employed Snopes.com, which is an online authority for detecting and verifying rumors in social media, emails and other online networks [10], based on expert sourcing. The final list of rumors, shown in Table I, along with example tweets which propagate it, includes a total of 6 Zika rumor stories (4 from WHO and 2 from Snopes). Note that the selection of these Zika rumor topics was supervised by health experts (acknowledged below) in order to insure the coverage of the most important and influential topics related to the Zika outbreak.

[2]http://aidr.qcri.org/
[3]http://thematicmapping.org/downloads/world_borders.php
[4]http://www.programmableweb.com/api/yahoo-placemaker

TABLE I
Zika rumor descriptions and example tweets. First four come from WHO and last two from Snopes.

Rumor Description	Example tweets
R1) Zika virus is linked to genetically modified mosquitoes	*BIOWEAPON! #Zika Virus Is Being Spread by #GMO #Mosquitoes Funded by Gates!*
R2) Zika virus symptoms are similar to seasonal flu	*The affects of Zika are same symptoms as the Common Cold. #StopSpreading-GMOMosquitos*
R3) Vaccines cause microcephaly in babies	*Government document confirms tdap vaccine causes microcephaly.. https://t.co/4ZVLbaabbG*
R4) Pyriproxyfen insecticide causes microcephaly	*"Argentine and Brazilian doctors suspect mosquito insecticide as cause of microcephaly"*
R5) Americans are immune to Zika virus	*Yup and Americans R immune to Zika, so why fund a response to it?*
R6) Coffee as mosquito-repellent to protect against Zika	*Bring on the Cuban coffee. Say Goodbye to Zika mosquitoes. Dee Lundy-Charles Fredric Sweeney Joshua Oates Laure... http://fb.me/tArL595b*

V. Rumor Tracking

A. Query Construction

We consider the task of extracting tweets relevant to rumors as a standard Information Retrieval task. After indexing the collected tweets using Indri[5], we submit a set of handcrafted interactively designed search queries (similarly to [11]) over at least 3 iterations of labeling the top 10 returned results. Each query is a boolean string consisting of a list of keywords connected using the AND, OR and NOT operators where a series of possible synonyms and keyword replacements are connected via the OR operator.

Designing the queries to extract the tweets was not a trivial task. One of the challenges is that many medical term synonyms needed to be added to the query to get the highest coverage. Additionally, we added words that distinguish general information tweets from rumors. For example, in R2, to distinguish a rumor from a general information, we need to add (NOT rash) to the query because this is the symptom that differs between Zika symptoms and the seasonal flu ones.

The final retrieval resulted in 89,572 tweets varying greatly by rumor, with a maximum of 73,832 to 202 (Table II). These tweets, however, still may contain false positives, tweets that match the query but are not a rumor. For example, the tweet "*Government document confirms tdap vaccine causes microcephaly.. https://t.co/4ZVLbaabbG*" states

[5]http://www.lemurproject.org/indri.php

TABLE II
RUMOR QUERIES AND THE NUMBER OF TWEETS RETRIEVED.

No	Regular Expression Query	# tweets
R1	genetically | GMO	73,832
R2	(symptom & (flu | cold)) & (not(rash))	469
R3	(tdap | MMR | Measles | Mumps | Rubella) & vaccine & microcephaly) | (vaccine &(cause | link | relate) & microcephaly)	4,329
R4	(montsanto | pesticide | pyriproxyfen | insecticide) & microcephaly	10,389
R5	american & immune	351
R6	((coffee | java | jive) & (repellent | protect)) & (java & jive) & (coffee & mosquito))	202
Total	-	89,572

TABLE III
CROWDFLOWER LABEL STATISTICS OF UNIQUE TWEETS IN EACH CATEGORY (PROPAGATED LABELS TO DUPLICATES IN PARENTHESES).

	Labeled	Rumor	Clarification	Other
R1	1,000 (42,432)	253 (11,773)	50 (1,912)	697 (28,747)
R2	302 (469)	217 (348)	71 (100)	14 (21)
R3	796 (4,329)	478 (2,853)	88 (846)	230 (630)
R4	1,000 (8,085)	749 (5,586)	221 (2,338)	30 (161)
R5	131 (351)	17 (22)	99 (17)	15 (312)
R6	114 (202)	72 (129)	5 (25)	37 (48)

that Zika vaccine causes microcephaly **(rumor)**. However, the tweet "*Anti-vaccination extremists falsely claim that Tdap #vaccine causes microcephaly suspected to be caused by.. https://t.co/yvfHlAFKhw*" clarifies that there is no evidence suggesting Zika vaccine causes microcephaly **(clarification)**. Similarly the tweet "*No cure, no vaccine for a virus that scientists believe to cause microcephaly! #microcephaly #ZikaVirus https://t.co/EuG9b1AJVw*" does not mention anything about the relationship between the vaccine and microcephaly **(other)**.

B. Crowdsourced Annotation

To annotate the tweets as to whether they are indeed rumors, we employ the crowdsourcing platform Crowdflower.com. Previous studies have shown that using crowds (anonymous workers) for health-related annotation is an effective way to label large amounts of data without employing experts [12], [13]. Each tweet was labeled as either supporting the rumor (by outright statement or ambiguity), debunking it (by clarification), or doing neither. For each topic we create a set of no fewer than 20 "gold standard" tweets in order to test the quality of annotations throughout the labeling. If an annotator did not pass the threshold of 70% accuracy, he/she would be banned from the task Each tweet was labeled at least 3 times and a majority vote determined its classification.

The tweets were first de-duplicated by stripping tweet-specific elements such as RT (standing for "re-tweet"), special characters, and mentions, such that only one copy of each tweet was to be labeled. A maximum of 1,000 tweets were annotated per rumor. For those which had more than 1,000 unique tweets (R1 and R4), we first selected 700 most re-tweeted tweets, and sampled 300 from the rest. After the labeling of these unique ones, the label was then propagated to the duplicates within the set.

Table III shows the distribution of classes for the six rumors, with the number of tweets with propagated labels in parentheses. Although the queries were hand-crafted to capture rumors, only 51% of final tweets were rumors (an average percentage across topics) and 15% were clarifications.

C. Temporal Tracking

Next, we examine the "paths" these rumors have taken in the story line of Zika in our dataset. Figure 2 illustrates the bursty nature of these rumors. The plots also show Pearson product-moment correlation r between the rumor and clarification volumes. For R4,5,6, the volume of clarification corresponds rather closely to that of the rumor with r of around 0.5. However, R1,2,3 display a mismatch between clarification attempts and the rumors. We define the "origin" tweets for rumors or clarifications as the most prominent tweets at that time for the corresponding class. For space limitations, we explain in details paths of some rumors in Figure 2 as follows:

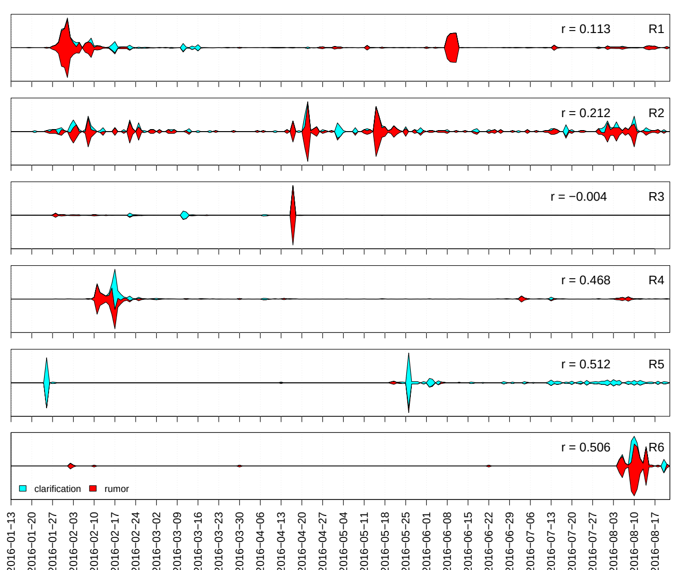

Fig. 2. Volume of the six rumors and their clarifications, along with the Pearson product-moment correlation r between the rumor and clarification volumes.

The killer vaccines: R3's peak originated in April with an article on an advocacy website www.march-against-monsanto.com (which argues that Monsanto, an agricultural biotechnology corporation, threatens the environment and the farmers) titled "1991 Government Document Confirms Tdap

Vaccine Causes Microcephaly"[6]. The article was readily believable to people who already view Monsanto negatively influenced by pharmaceutical companies to sell new Zika virus vaccines, as Dredze et al. [8] suggested in his paper. The post happened after a major WHO campaign in February and March saying "No evidence that vaccines cause microcephaly"[7]. Interestingly, the April spike receded just as quickly without any clarifications from authoritative sources.

Pesticides, immunities and coffee grounds: R4,5,6 however had a strong interaction between the rumor and a quick reaction with clarifications. For instance, the most retweeted stories of R4 are those coming from mainstream media including CNN and WHO stating there is "No link between pesticide and microcephaly". At the top three of R5 are stories on the "crazy and dangerous story [that] Americans are immune to Zika" and links to the debunking website Snopes. Similarly to R2, in R6 is a case of hyperbole and exaggeration of a story saying mosquito larvae do not thrive in coffee-infused water, which was turned into sensationalist tweets claiming "Could Coffee Be the Answer in the Fight Against Zika Mosquitoes?", but which still linked to the original correct information.

Thus, we show the varied nature of the rumors in the Zika stream. Those which were accompanied with mainstream coverage quickly decreased (R4-6), but even those which originated from the websites of various advocacy groups and were not met with official response were also short-lived (R1, R3). The longer-lived one is the one which concerned the daily occurrences (having a flu R2 or, possibly, coffee R6) which propagates in the Twitter lore.

D. Rumor Classification

Next, we turn to the supervised methods which have been proposed in previous work on news in social media that seek to establish the level of credibility of information automatically by observing specific features extracted from the social media. For instance, Castillo et al. [16] and Qazvinian et al. [11] suggested that the best features to assess the credibility of news topics are those that look into the user, message and topic features. Inspired by these works, we build a set of features in order to automatically distinguish rumors from non-rumors.

Gathering all the relevant tweets to the topics in Table 2, results in a total of 56,985 tweets. Later, we filter tweets that are exact duplicates (tweets sharing exact similar information including text, urls, hashtags, and mentions) as the presence of the duplicates might influence the precision and recall values, resulting in a total of 26,728 tweets with human-assigned labels. We group the labels used in Table 2 such that we consider a rumor as the tweet that has been labeled by Crowdflower users as "rumor" (32% - 8,488 tweets) and a non-rumor as the tweet that has either been labeled as "clarification" or "other" (68% - 18,240 tweets).

The feature set consists of 48 features (Table IV) grouped into five categories. The first three categories (Twitter, sentiment and linguistic features) have been previously implemented in news credibility detection, whereas the last two (readability and medical features) are new to this work:

Twitter features As [16] use Twitter features to define credibility in news topics, we build 18 similar features (number of retweets, number of users followers/following, etc.).

Sentiment features We consider five different measures of emotional state: count of positive/negative words, positive/negative smileys and sentiment scores [14].

Linguistic features We also introduce 17 measures to characterize different linguistic styles in the text [16] (count adjectives, adverbs, etc.).

Readability features [15] defined the readability score as a measure of easiness to understand a text. We introduce text readability measures with the intuition that more readable information is more credible. We implement the predefined readability scores: Flech, automated, Flesch_kincaid, Gunning, and SMOG [15]. We also count the number of complex words, average number of syllables per word and number of words not in word2vec news vocabulary [17] which may signal slang language.

Medical/Domain features We suggest medical domain features which focus on the medical lexicon and the reliability of sources shared in tweets. First, we build a medical lexicon[8] which signals how many medical terms are used in the tweet. Prior work [18] showed that Wikipedia is a reliable knowledge base for medical data extraction [19]. We build a medical lexicon by crawling 113 Wikipedia pages under the "Infectious disease" category, resulting in 22,123 words labeled as corpus *M*. Then, we download the 22,123 most frequent words in Wikipedia, labeled as corpus *W*. Later, we compute the probability of every word in *M* and *W* as $mp_w = count_w / \sum_w M$ and $wp_w = count_w / \sum_w W$ (respectively). Next, for every word in *M* and *W*, we compute the differences in probabilities $p_w = mp_w - wp_w$. Intuitively, p_w provides the most descriptive words related to "infectious disease" topic which are *not* as prevalent in the general Wikipedia. Ranking the terms by p_w, we only keep the top 13,300 meaningful words. For example specific words like *syphilis* and *bronchitis* have high ranks by p_w values (0.01 and 0.002 respectively) compared to *treatment* and *life* general words (-4.633 and -34.608 respectively).
Additionally, Wikipedia references are considered trusted citations as Wikipedia increasingly includes references to medical journals with high impact factors [20]. From the 113 Wikipedia pages under "Infectious disease", we collect a total of 2,979 cited URLs from 441 different domains[9] including medical literature databases and news agencies. As most Twitter URLs are shortened, we expanded the URLs and we count the number of URLs within a tweet that are Wikipedia domains. Further, we manually classify tweet URL domains as *advocacy* group

[6] http://www.march-against-monsanto.com/1991-government-document-confirms-tdap-vaccine-causes-microcephaly/

[7] https://twitter.com/WHO/status/708317001366806528

[8] Available at http://bit.ly/2m56t0w

[9] Available at http://bit.ly/2m59wpm

TABLE IV

AUTOMATICALLY EXTRACTED FEATURES OF TWEETS POTENTIALLY BELONGING TO A RUMOR.

Scope	Feature	Description
Twitter	IS RETWEET	Is a retweet; contains RT
	FOLLOWING	The number of people the user is following
	FOLLOWERS	The number of people following the user
	STATUS_COUNT	The number of tweets at posting time
	AGE	The time passed since the author registered his/her account, in days
	HAS MENTIONS	Mentions a user, eg: @CNN
	HAS HASH TAG	Contains hash_tags
	COUNT HASH TAG	Count total number of hash_tags
	DAY WEEKDAY	The day of the week in which the tweet was written
	COUNT URLS	Count total number of URLs in text
	COUNT RT	Count total number of Retweets
	COUNTRY	The country the tweet was originated from
Sentiment	SENTIMENT SCORE	sentiment score value [14]
	POSITIVE/NEGATIVE WORDS	The number of positive/negative words in text
	EMOTICONS POS/NEG	Count total number of positive and negative emoticons in text
Linguistic	QUESTION MARK	Contains question mark '?'
	EXCLAMATION MARK	Contains exclamation mark '!'
	WORDS COUNT	Count total number of words in text
	COUNT SENTENCES	Count number of sentences
	CHAR COUNT	Count total number of characters in text
	UPPER COUNT	Count total number of upper case letters
	PERCENTAGE UPPER	The percentage of upper case characters
	PERCENTAGE UPPER/LOWER	The percentage of upper and lower case characters
	MULTIPLE QUES/EXCL	Contains multiple questions or exclamation marks
	COUNT NOUN	Count total number of nouns in text
	COUNT ADVERB	Count total number of adverbs in text
	COUNT ADJECTIVE	Count total number of adjectives in text
	COUNT VERB	Count total number of verbs in text
	COUNT PRONOUN	Count total number of pronouns in text
	HAS PRONOUN 1	Contains a personal pronoun in 1th person
	HAS PRONOUN 2	Contains a personal pronoun in 2nd person
	HAS PRONOUN 3	Contains a personal pronoun in 3rd person
Readability	COMPLEX WORDS	Count total number of complex words in text
	READABILITY SCORES	Flesch, Automated, Flesch_Kincaid, Gunning, and SMOG [15]
	COUNT NOT WORD2VEC	Count total number of words not in "word2vec" Google News vocabulary
	AVG SYLLABLES	The Average number of syllables per word in text
Medical	MEDICAL_LEXICON	Count number of words in the medical lexicon
	WIKIPEDIA DOMAIN	Count number of URL domains mentioned in the wikipedia web pages
	ADVOCACY	Count number of URLs belonging to advocacy domains
	NEWS	Count number of URLs belonging to news domains
	SOCIAL	Count number of URLs belonging to social media domains
	INFORMATIVE	Count number of URLs belonging to informative/trusted domains

(claiming to be the best in providing information without official ties), *social_media* (social media helper websites that forward or aggregate content such as YouTube), *news* (news sources CNN, Reuters, etc.), *informational* (reliable resources providing medical information: medical companies, Snopes, etc.) or *non-informative* (URLs having no specific domain type), resulting in four features for each domain type.

In order to pick the best features for the classification task, we employ two different automatic feature selection techniques: Information Gain (IG) [21] and Greedy backward elimination technique (GBE) [21]. Table V shows the top features each technique produced. Here, we list the top ten features by information gain value and GBE results selecting the best ten features. Based on both techniques, the most significant features correspond to the medical features (advocacy domains count, Wikipedia domains count) followed by the syntax of the tweet text (question marks, exclamation marks...) and the sentiment features (sentiment score, count positive/negative words) and some Twitter features.

Note that advocacy feature domain type is the strongest feature with high IG value (table V). It is understandable this feature would be useful, given that it requires expert annotation. Further, we find that out of the URLs cited in rumor tweets, 35.0% were from advocacy websites, 0.1% from social media, 39.1% were news and 25.9% were informative domains, compared to 3.1% from advocacy and 0.6% social media, 32.3% news, and 64.0% informational in non-rumors, making the presence of advocacy groups and informational sources the distinctive features, and, interestingly, not the news media. Wikipedia domains features is also among the top selected features in both techniques and this features is automatically computed and can be used more broadly.

Finally, we train a supervised classifier to predict which tweets contain rumor and which do not. We build a classifier separately for the top 10 features of IG and GBE techniques.

TABLE V
THE FEATURES SELECTED USING INFORMATION GAIN AND GREEDY BACKWARD ELIMINATION.

Feature	min, max	μ (σ)	IG*	GBE•
(T) AGE	61, 281	188 (71)	9	✓
(T) HAS MENT	0, 1	0.177 (0.381)	10	✓
(T) COUNT RT	1, 2457	394 (713)	6	✓
(S) SENTIMENT	-2.2, 1.6	-0.332 (0.71)	8	✓
(S) NEG COUNT	0, 13	0.639 (0.871)	-	✓
(L) HAS QUEST	0, 1	0.193 (0.395)	4	✓
(L) HAS EXCL	0, 1	0.023 (0.161)	5	-
(L) VERB CNT	0, 38	0.673 (0.716)	-	✓
(L) ADVB CNT	0, 102	0.682(0.936)	3	-
(L) MULT. '?/!'	0, 1	0.014 (0.12)	2	✓
(M) ADVOCACY CNT	0, 2	0.045 (0.21)	1	✓
(M) WIKI CNT	0, 1	0.253 (0.435)	7	✓

* Features are ranked desc according to information gain values.
• ✓: is in GBE best 10 feature subset, otherwise not.

We experiment with three different learning algorithms: Naïve-Bayes algorithm [22], Random Forest [23] and Random Decision Tree [24]. For training/validation process, we perform 10-fold cross validation. The best classifier - using Random Tree and the top 10 GBE features - achieves an average precision of 0.946 with an average recall 0.944 which is significantly better than a random predictor. The F-value (a harmonic mean of precision and recall) is high, indicating a good balance between precision and recall values. Note that these results are overfitted, given the limited amount of data available, feature selection on the test set, and also that the method relies on manually labeled tweets, with the addition that the dataset is already topically specialized.

As we find having training data within the topic to be extremely helpful in building accurate classifiers, we explore a more challenging scenario wherein the classifier is trained on 5 topics and tested on the 6th. The results show that the performance within topics is not uniform. For example, topics 1 and 5 have the worst precision (0.296 and 0.101 respectively), while topics 3 and 4 have recall under 0.500. Once again, this points to the importance of expert labeled data that is topically matched to the one in question.

VI. CONCLUSION

This paper presents tool pipeline incorporating expert knowledge, crowdsourcing, and machine learning for health-related rumor discovery in a social media stream. In particular, our study shows that tracking health misinformation in social media is not trivial, and requires some expert supervision. We also show the bursty and varied nature of the Zika rumors, some provoked by known advocacy groups, others propagated due to their affordance for humor or light banter. We hope this work will encourage a collaboration between health professionals and data researchers in order to quickly understand and mitigate health misinformation on social media.

ACKNOWLEDGMENT

We would like to thank Luis Fernandez-Luque, Qatar Computing Research Institute, William Schulz and Clarissa Simas, London School of Hygiene & Tropical Medicine, Per Egil Kummervold, Norut Northern Research Institute for the fruitful discussions and health related expertise.

REFERENCES

[1] H. Webb, P. Burnap, R. Procter, O. Rana, B. C. Stahl, M. Williams, W. Housley, A. Edwards, and M. Jirotka, "Digital Wildfires: Propagation, Verification, Regulation, and Responsible Innovation," *TOPS*, vol. 34, no. 3, pp. 1–23, 2016.

[2] L. Fernandez-Luque and T. Bau, "Health and social media: perfect storm of information," *Healthcare informatics research*, vol. 21, no. 2, pp. 67–73, 2015.

[3] World Health Organization, "Mixed uptake of social media among public health specialists," *Bull World Health Organ*, vol. 89, pp. 784–785, 2011.

[4] World Health Organization *et al.*, "Risk communication in the context of zika virus: interim guidance," pp. 3–4, March 2016.

[5] C. Chew and G. Eysenbach, "Pandemics in the age of twitter: content analysis of tweets during the 2009 h1n1 outbreak," *PloS one*, vol. 5, no. 11, p. e14118, 2010.

[6] P. Kostkova, V. Mano, H. J. Larson, and W. S. Schulz, "Vac medi+ board: Analysing vaccine rumours in news and social media," in *ACM DH6*. ACM, 2016, pp. 163–164.

[7] M. Imran, C. Castillo, J. Lucas, P. Meier, and S. Vieweg, "Aidr: Artificial intelligence for disaster response," in *ACM WWW'15*. ACM, 2014, pp. 159–162.

[8] M. Dredze, D. A. Broniatowski, and K. M. Hilyard, "Zika vaccine misconceptions: A social media analysis," *Vaccine*, vol. 34, pp. 3441–3442, 2016.

[9] World Health Organization. (2016, September) Dispelling rumours around zika and complications. Accessed: 2016-07-20.

[10] A. Nourbakhsh, X. Liu, S. Shah, R. Fang, M. M. Ghassemi, and Q. Li, "Newsworthy rumor events: A case study of twitter," in *ICDMW*. IEEE, 2015, pp. 27–32.

[11] V. Qazvinian, E. Rosengren, D. R. Radev, and Q. Mei, "Rumor has it: Identifying misinformation in microblogs," in *EMNLP*. ACl, 2011, pp. 1589–1599.

[12] B. Yu, M. Willis, P. Sun, and J. Wang, "Crowdsourcing participatory evaluation of medical pictograms using amazon mechanical turk," *JMIR*, vol. 15, no. 6, p. e108, 2013.

[13] H. Zhai, T. Lingren, L. Deleger, Q. Li, M. Kaiser, L. Stoutenborough, and I. Solti, "Web 2.0-based crowdsourcing for high-quality gold standard development in clinical natural language processing," *JMIR*, vol. 15, no. 4, p. e73, 2013.

[14] W. Lowe, K. Benoit, S. Mikhaylov, and M. Laver, "Scaling policy preferences from coded political texts," *Legislative studies quarterly*, vol. 36, no. 1, pp. 123–155, 2011.

[15] L. Feng, M. Jansche, M. Huenerfauth, and N. Elhadad, "A comparison of features for automatic readability assessment," in *COLING: Posters*. ACL, 2010, pp. 276–284.

[16] C. Castillo, M. Mendoza, and B. Poblete, "Information credibility on twitter," in *ACM WWW*. ACM, 2011, pp. 675–684.

[17] T. Mikolov, I. Sutskever, K. Chen, G. S. Corrado, and J. Dean, "Distributed representations of words and phrases and their compositionality," in *NIPS*, 2013, pp. 3111–3119.

[18] J. Friedlin and C. J. McDonald, "An evaluation of medical knowledge contained in wikipedia and its use in the loinc database," *JAMIA*, vol. 17, no. 3, pp. 283–287, 2010.

[19] B. Rink, S. Harabagiu, and K. Roberts, "Automatic extraction of relations between medical concepts in clinical texts," *JAMIA*, vol. 18, no. 5, pp. 594–600, 2011.

[20] J. M. Heilman, E. Kemmann, M. Bonert, A. Chatterjee, B. Ragar, G. M. Beards, D. J. Iberri, M. Harvey, B. Thomas, W. Stomp *et al.*, "Wikipedia: a key tool for global public health promotion," *JMIR*, vol. 13, no. 1, p. e14, 2011.

[21] M. Cord and P. Cunningham, *Machine learning techniques for multimedia: case studies on organization and retrieval*. Springer, 2008.

[22] H. Zhang and J. Su, "Naïve bayesian classifiers for ranking," in *Machine Learning: ECML 2004*. Springer, 2004, pp. 501–512.

[23] T. K. Ho, "Random decision forests," in *ICDAR*, vol. 1. IEEE, 1995, pp. 278–282.

[24] W. Du and Z. Zhan, "Building decision tree classifier on private data," in *CRPIT '14*. Australian Computer Society, Inc., 2002, pp. 1–8.