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, 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 – spanning the initial reports in February 2016 and the Summer Olympics – 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. In the case of Zika, 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. Thus, we illustrate insights the proposed tools provide into potentially harmful information on social media, allowing public health researchers and practitioners to respond with a targeted and timely action.

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 and rumors 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 and 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 identified by the WHO. The Zika virus has been known for decades; it was discovered in Uganda in the 1940s and until recently it has been unnoticed. Things changed dramatically in 2015 when this mosquito-borne disease started to spread quickly across Brazil and then most of the American continent, becoming a major global health crisis. This crisis became more dramatic as the link between the Zika infection and serious brain malformation (i.e. microcephally) started to emerge. Furthermore, fears of a global pandemic started to emerge, since Zika is spread by a mosquito from the *Aedes* family, which is present in many countries. Another source of concern were the Rio summer Olympics Games, which brought international travelers to the affected areas. As at the time there was no cure or vaccine for the Zika viral infection, communication with the public was one of the most important tools to control this outbreak. These communication efforts – dealing with the detection and prevention of Zika, and also the reduction of mosquito breeding – have been challenged by the appearance of rumors that, in the best of cases, were interfering with the public health campaigns [4].

In particular, we track rumors outlined by the WHO (along with Snopes.com[1]) in the stream of nearly 13 million tweets. We employ both automated LDA-based topic discovery as well as high-precision expert-led Information Retrieval approach to identifying the relevant tweets in this stream. 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.

As discussed in the coming section, this work contributes to the literature a first application of the state-of-the-art social media analytic tools to the problem of health rumor tracking.

II. Related Work

This paper is the first of its kind to relate health informatics, machine learning and social network analysis to detect health-related rumors in a social media site. Below, we outline the

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

most relevant recent work related to the detection and tracking of health-related attitudes, as well as misinformation in other non-health related fields.

Health-related attitudes on social media. As social media captures daily thoughts and actions of millions of users, recent efforts have been made in extracting health-related behaviors and attitudes from the textual content of these websites. Recently, [5] build a classifier to detect the use of marijuana on Twitter. Attitudes toward legal drugs, including Xanax and Adderall, have also been studied by [6]. Further, [7] use association mining to health communities to discover adverse effects of drug interactions. Health-related attitudes around food on Instagram have been mined by [8]. These studies often combine the health expert knowledge with big data analytics (including machine learning, in the case of [5]), to provide insight into attitudes captured in social media interactions. Similar tools can be applied to non health-related rumor detection, as we discuss next.

Rumor detection and analysis. Separating newsworthy stories from misinformation and rumors across microblogging sites has been a popular research topic in recent years. Some focus on identifying information credibility of news propagated in social media [9], [10], [11], [12]. Others focus on detecting misleading political memes, such as the "Truthy" service presented by Ratkiewicz et al. [13], designed to detect fake political grass-roots movements (dubbed as "astroturf"). Often, machine learning models based on features related to either users or the content of propagated messages are employed. For example, [12], [11] train a classifier to automatically detect the rumors on Sina Weibo microblog site[2]. They use Sina Weibo rumor-busting service[3] as ground-truth data for the classification. Further, Jing Ma et al. [14], use both Twitter and Sina Weibo as a case study in order to detect rumors. Using time series, they capture the temporal characteristics of the rumor detection features. Authors use already labeled datasets (Twitter [9], Weibo [12]) to train and evaluate the approach. Unlike these studies, we consider a more natural setting wherein no labeled data is available, and the documents associated with the rumors must first be found. In this respect, our approach is similar to Qazvinian et al. in [15] who use content-based, network-based, and Twitter-specific features in order to track urban legends. They build Bayes classifiers using engineered features and then learn a linear function of these classifiers for rumor retrieval and classification.

Health Related misinformation detection. WHO's white paper on Risk communication in the context of Zika virus urges to "build capacity to quickly transform new information into usable, culturally-appropriate and easily understood risk communication resources that can be disseminated on multiple platforms", including social media [16]. However, little work thus far concentrated on detecting and tracking health-related rumors. Recently, Kostkova et al. [17] 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. [18], where content is selected to be annotated via crowdsourcing into pre-defined 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. [19], 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 "supervised machine learning technique" (although these were not explained in the manuscript) and observed the effect of vaccine-skeptic communities over other users' vaccination opinion. While [19] look at two Zika vaccine related memes, in this work we propose a more general methodological pipeline to track health-related rumors. Taking Zika as a case study, with the help of health professionals we expand the list of rumors to six, and examine the behavior of rumor as well as clarification efforts. Below we describe in detail our contributions to the text classification task via novel health-related features, and the application of crowdsourcing to fine-grained document labeling.

III. DATA COLLECTION

The data was collected using the Artificial Intelligence for Disaster Response (AIDR)[4] platform, which taps into Twitter Streaming Application Program Interface (API). The keywords' list contained the following (searched as quoted strings): 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. Furthermore, "zika" word is used by all English, Portuguese, and Spanish, which are the major languages of populations affected. The resulting collection of 13,728,215 tweets spans January 13 - August 22, 2016, and includes the peak of interest in Zika (in early February) and the Olympic Games in Brazil (August 5-21). Figure 1 shows the volume of the data by language.

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, Spanish and Portuguese (46%, 27% and 17% respectively). The language was determined using the language tag in the meta-data of the tweet provided by the Twitter API. Table I summarizes the global statistics of these languages distributions. It illustrates the international nature of the Zika crisis, with each language identifying a population and its diasporas affected. In this work we focus on the English data, and discuss the future work involving other languages below.
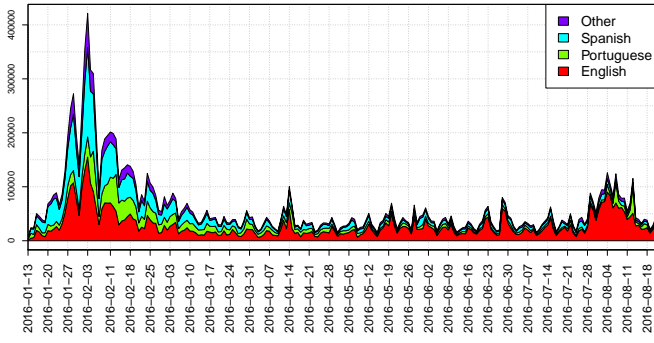
---

[2]http://weibo.com

[3]http://weibo.com/weibopiyas

[4]http://aidr.qcri.org/

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

TABLE I
DATA STATISTICS BY MAIN LANGUAGE GROUPS.

Language	Total tweets	Users	Tweets/user
English	6,267,173	1,318,293	4.75
Spanish	3,689,292	727,105	5.07
Portuguese	2,296,611	623,968	3.68
Other	1,475,139	593,221	2.49
Total	13,728,215	3,262,587 *	4.21

* or 2,546,851 unique users

To understand the geographic distribution of tweets, we use several sources to geo-locate them. We begin with the GPS attributes (*latitude*, *longitude*) of the tweet meta-data and convert them to the corresponding country name using World Borders API[5]. As 99% of tweets have missing GPS attributes, we look at location names in the *place* attribute of the tweet and convert them to exact country names. In case there is no mentioned location, we assume that the tweet location is where the user is located and get the corresponding location from the user's profile. It is worth mentioning that users' locations are messy, as they are written by the users. Thus, we use Yahoo Placemaker API[6] to map the users' place fields to GPS locations. Finally, in case no user location is mentioned, we get the tweet location by looking into other already identified location tweets tweeted by the same user (resorting to this noisy approach only if all other geo-location attempts fail). Implementing these steps, we achieve 68% coverage for English data. The top locations in decreasing number of tweets are the United States, United Kingdom, India, Canada, Nigeria, and Brazil, indicating a highly international data.

IV. RUMOR SELECTION

Our rumor selection process begins with a reliable list of information from trusted sources. We chose the WHO website as an authority for detecting and verifying rumors about Zika. As Zika was spreading further in the world, WHO provided a source listing major international rumors and misinformation about the virus. At the time of writing, WHO website [20] listed 8 statements debunking ongoing rumors. Out of these,

[5]http://thematicmapping.org/downloads/world_borders.php
[6]http://www.programmableweb.com/api/yahoo-placemaker

TABLE II
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. #StopSpreadingGMOMosquitos*
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*

4 were unsuitable, as they were not topically cohesive. For instance, one explained "Fish can help stop Zika", but did not explicitly state what the rumor which this statement would debunk would be. The process in understanding whether a topic is a rumor involved writing out an unambiguous description of the rumor, such that a tweet may be easily classified as being one. If no such description could be written, the topic was discarded. Additionally, we employed Snopes.com, which is an online authority for detecting and verifying rumors in social media, emails and other online networks [21], based on expert sourcing. From its Zika-related articles, we selected rumors which are not specific to the US. The final list of rumors, shown in Table II, 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.

V. RUMOR TRACKING

We first attempt to discover these rumors using an automated technique, as topic discovery has been used to identify rumors [14], [12] in social media. We train a Latent Dirichlet Allocation (LDA) [22] model on the english-language tweets, which then produces n "topics", grouping words which appear in similar contexts together in a topic. However, after a manual examination of n (varying from 5 to 50) topics, we do not find any topics pertaining to the above selected rumors. The vast majority of topics were informational, following by spam and jokes. Thus, we illustrate the necessity of incorporating expert knowledge in order to achieve a high-precision view of the data for our purpose.

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

	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)

A. Query Construction

We consider the task of extracting tweets relevant to our rumors as a standard Information Retrieval task. We first index the collected tweets using Indri[7], and submit a set of handcrafted interactively designed search queries (similarly to [15]). Each query is a boolean string consisting of a list of keywords that best describe the rumor. These keywords are first identified, then connected using the AND, OR and NOT operators. Each keyword is then replaced with a series of possible synonyms and replacements, all connected via the OR operator. For instance, consider the rumor saying that "vaccines cause microcephaly" (R3). Transforming this story to a query language would include several common ways of referring to vaccines, as shown in Table III. The queries are hand-crafted over at least 3 iterations of labeling the top 10 returned results.

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. We did not rely on automatic query expansion techniques such as Pseudo-relevance feedback as these automatic algorithms perform well in medical articles and not in informal unstructured text such as Twitter messages [23]. 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 III). These tweets, however, still may contain false positives, tweets that match the query but are not a rumor. For example, the following tweets are all about vaccines and microcephaly in babies (R3). The first tweet is stating that Zika vaccine causes microcephaly (**rumor**), but the second tweet clarifies that there is no evidence suggesting Zika vaccine causes microcephaly (**clarification**), and the third does not mention anything specific about the relationship between Zika vaccine and microcephaly (**other**).

(**rumor**) *Government document confirms tdap vaccine causes microcephaly.. https://t.co/4ZVLbaabbG*

(**clarification**) *Anti-vaccination extremists falsely claim that Tdap #vaccine causes microcephaly suspected to be caused by.. https://t.co/yvfHlAFKhw*

(**other**) *No cure, no vaccine for a virus that scientists believe to cause microcephaly! #microcephaly #ZikaVirus https://t.co/EuG9b1AJVw*

In the coming section, we explain the approach we take in order to distinguish between the three different types of information available in our data-set.

B. Crowdsourced Annotation

To annotate the tweets as to whether they are indeed rumors, we employ the crowdsourcing platform "Crowdflower"[8]. 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 [24], [25]. We begin by creating a task for each topic with clear instructions on the labeling of the tweets as either supporting the rumor (by outright statement or ambiguity), debunking the rumor (by clarification), or doing neither. Also for each task we create a set of no fewer than 20 "gold standard" tweets (those with known classifications) in order to test the quality of annotations throughout the jobs. If an annotator did not pass the threshold of 70% accuracy, he/she would be banned from the task and the annotations would be discarded. 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 IV shows the distribution of classes for the six rumors, with the number of tweets with propagated labels in

---

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

[8]http://www.crowdflower.com/

parentheses. Although the queries were hand-crafted to capture rumors, only 51% of final tweets were rumors (an average percentage across topics, such that no one topic dominates the statistic), and 15% clarifications, attempting to debunk these rumors. The annotator agreement (as measured in label overlap) ranged between 76% (R2) and 93% (R5) with an average of 87.7%, indicating the task differs in difficulty, but is overall clear to the annotators.

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 and we explain Figure 2 in details as follows:
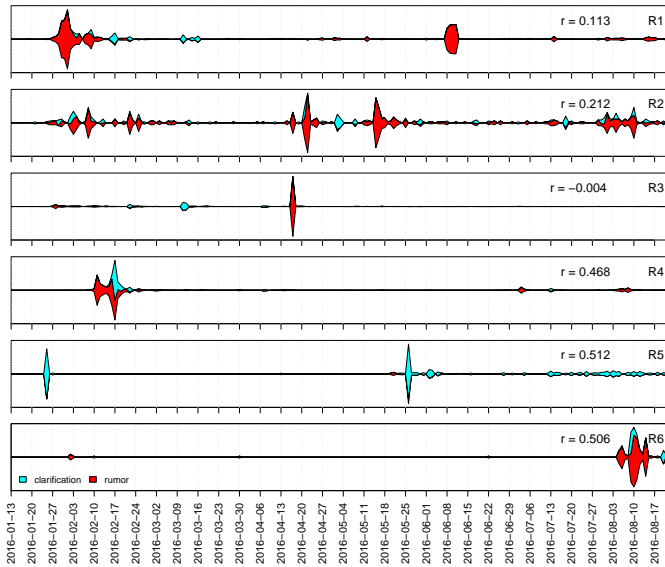

Fig. 2. Volume of the six rumors and their clarifications, along with the Pearson product-moment correlation r between the rumor and clarification volumes.

The case of mutant mosquitos: the case of R1 is interesting, since it carries over a concern about the dangers of genetically modified organisms (GMO) which has been popular for several years. For instance, the spike in July was due to an article published on The Real Strategy website claiming a link between "chemical exposure" and microcephaly[9], which gained thousands of retweets within days. However, without any interference that we detected from authoritative sources, the rumor quickly died out.

[9]For more on this rumor see http://www.huffingtonpost.com/entry/zika-monsanto-pyriproxyfen-microcephaly_us_56c2712de4b0b40245c79f7c

Have you got Zika? Flu and cold are very common diseases, therefore confusion between flu or cold and Zika might pose a serious problem for health authorities. This case is addressed in R2. Often the tweets appear to be jokes of users who feel flu-like symptoms such as:

> RT @arzel: my friend had a small cold and I caught him googling "zika virus symptoms"

Thus, although there are regular tweets on the true symptoms of Zika, there is a large proportion among these tweets that are jokes or lighthearted statements.

The killer vaccines: Similar to R1, R3's peak originated in April with an article on another 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"[10]. The article was readily believable to people who already view Monsanto negatively and might be spread by pharmaceutical companies to create an opportunity to sell new Zika virus vaccines as Dredze et al. [19] suggested in his paper. The post happened after a major WHO campaign in February and March saying "No evidence that vaccines cause microcephaly"[11]. Interestingly, the April spike receded just as quickly without any clarifications from authoritative sources.

Pesticides, immunities and coffee grounds: Others, however, did have 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.

[10]http://www.march-against-monsanto.com/1991-government-document-confirms-tdap-vaccine-causes-microcephaly/
[11]https://twitter.com/WHO/status/708317001366806528

TABLE V

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 [26]
	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 [27]
	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

For instance, Castillo et al. [9] and Qazvinian et al. [15] 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 assess their power in automatically distinguishing 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). Note that we cannot consider "clarification" class alone, as it is vastly

under-represented in our data (in part due to our focus on retrieving rumors in the previous steps).

The feature set is listed in Table V and consists of 48 features 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 [9] use Twitter features to define credibility in news topics, we build 18 similar features including the number of retweets, number of users followers and following, the presence of hashtags and mentions, the user's number of tweets, etc.

Sentiment features We consider five measures of emotional state in our dataset: count of positive/negative words, count of positive/negative smileys and sentiment score representing the strength of sentiment [26].

Linguistic features We also introduce measures to charac-

terize different linguistic styles in Twitter text [9]. We compute 17 different linguistic styles e.g: count adjectives, adverbs, pronouns, sentences, upper and lower case characters.

Readability features Authors in [27] defined the readability score as a measure of how easy it is to understand a piece of text. We introduce a set of tweet text readability measures with the intuition that more readable information are more credible. We implemented the predefined readability scores by [27] (Flech, automated, Flesch_kincaid, Gunning, and SMOG scores) in addition to computing the number of complex words and average number of syllables per word. Moreover, we counted the number of words not in word2vec news vocabulary which may signal slang language [28].

Medical/Domain features We define specialized features in the medical domain by focusing on the medical lexicon of tweets and the reliability of sources shared using URLs. First, we build a medical lexicon[12] which signals how many medical terms are used in the tweet. Prior studies [29] showed that Wikipedia is a reliable knowledge base for medical data extraction tasks. Additionally, as a source for lexical and contextual features, Wikipedia was used in the past to improve medical text relation extraction [30]. Guided by prior work, we build a specialized lexicon by crawling a total of 113 Wikipedia pages under the category of "Infectious disease", resulting in 22,123 words representing corpus M. Then, we download the same number (22,123) of the most frequent words on all of Wikipedia, representing a general corpus W. These can then be used to compute a probability of every word in specialized corpus M as: $mp_w = count_w / \sum_w M$, as well as the probability of every word in general corpus W as $wp_w = count_w / \sum_w W$. Next, for every word in every corpus, we compute $p_w = mp_w - wp_w$. Intuitively, the differences in probabilities p_w provide 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, as illustrated in Table VI (note the topmost words are more specific, while those further down the ranking are more general).

Additionally, Wikipedia references are considered trusted citations as Wikipedia increasingly includes references with high impact factor medical journals such as the *New England Journal of Medicine*, *The Lancet*, the *Journal of the American Medical Association*, and the *British Medical Journal* among the 10 most frequently cited science journals in Wikipedia in 2007 [31]. As Wikipedia pages are usually among the top results of search engine queries [32], [31], we expect people to use Wikipedia pages and references as a major source of online health information. From the same Wikipedia pages used to collect the medical lexicon, we collect a total of 2,979 referenced

TABLE VI
SELECTED "INFECIOUS DISEASE" WIKIPEDIA MEDICAL LEXICON WORDS

Word (w)	mp_w	wp_w•	p_w	Rank
*syphilis**	0.01	-	0.01	4
*bronchitis**	0.002	-	0.002	81
tetanus *	0.001	-	0.001	236
diarrhea	0.006	0.128	-0.121	13682
epidemiology	0.009	0.147	-0.138	15284
treatment	0.019	4.652	-4.633	33869
life	0.003	34.61	-34.608	35074

* Among the chosen top 13,300 words with highest p_w
• - : is when w is not in the W corpus

URLs from 441 different domains[13] including medical literature databases and news agencies. As most Twitter URLs are shortened, we expanded the URLs to detect the original domain. Finally, we manually classify tweet URL domains as *advocacy* group (advocating specific actions or policies, or claiming to be the best in providing the related information without official ties), *social_media* (YouTube, Facebook and social media helper websites that forward and aggregate content), *news* (news sources CNN, Reuters, etc.), *informational* (reliable resources providing medical information: medical companies, government sites, Snopes...) or *non-informative* (URLs having no specific domain type). Doing this, we have a total of four different domain type features where every feature is a count of the number of URLs belonging to a domain class in the tweet.

In order to pick the best features for the classification task, we employ two different automatic feature selection techniques. First, Information Gain (IG) which is a popular filtering approach that aims at removing irrelevant features after computing the gain value (amount of information a feature brings to the training set) [33]. Second, we use Greedy backward elimination technique (GBE) that starts with a model having all features, and removes features one at a time until reaching a certain performance threshold [33].

Table VII 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 VII). 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,

---

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

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

Feature	min, max	μ (σ)	IG*	GBE$^\bullet$
(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.
$^\bullet$ ✓: is in GBE best 10 feature subset, otherwise not.

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. We experiment with three different learning algorithms. First, Naïve-Bayes algorithm [34], a probabilistic based on Bayes' theorem with strong ("naïve") independence assumptions between the features. Second, Random Forest [35], which is a collection of classifiers where every classifier votes for one class and every instance is classified based on the majority class. Third, Random Decision Tree [36], a classifier that recursively builds a tree by splitting the training data based on a criterion until all partitions have the same class label. We find the best results using the Random Tree classifier using the top 10 GBE features. For training/validation process, we perform 10-fold cross validation, in which 10 experiments are performed on a different tenth of the data held out for testing, such that we take advantage of the whole dataset for both training and testing. A summary of the best classifier (Random Tree) with top 10 GBE features results are shown in Table VIII. As it shows, the classifier achieves a precision of 0.946 with 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. The final row presents the average values from across both classes. 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 are shown in Table

Class	Precision	Recall	F-measure
rumor	0.929	0.921	0.925
non-rumor	0.963	0.967	0.965
weighted average	0.946	0.944	0.945

IX. Every row of the table shows which topic is excluded in training the classifier, and then is used for testing. We find the performance is not uniform, with topics 1 and 5 having the worst precision, while topics 3 and 4 having recall under 0.500. Once again, this points to the importance of expert labeled data that is topically matched to the one in question.

Topic	% Rumor	Precision	Recall	F-Measure
R1. Zika linked to GMO	61%	0.296	0.869	0.440
R2. Flu symptoms similar to Zika	31%	0.746	0.504	0.602
R3. Vaccines cause microcephaly	71%	0.683	0.490	0.571
R4. Insecticide cause microcephaly	32%	0.594	0.432	0.500
R5. Americans are immune to Zika	32%	0.101	0.523	0.170
R6. Coffee as mosquito repellent	31%	0.688	0.688	0.688

VI. DISCUSSION

Key Findings.

Communication on Twitter around a major public health crisis is an essential component in the public health response. In our search of rumors in the stream of Zika-related tweets, we find automatic topic discovery tools such as LDA to be too coarse-grained to tease out the rumors WHO and Snopes have cited as most concerning. Thus, we incorporate the expert knowledge to compose high-precision queries to retrieve the relevant tweets. We also show that further steps are needed, as after a closer examination we find only roughly half of the captured tweets to be actual rumors. This insight shows the perils of using keyword or hashtag-based topic definition, as is done, for example by "Truthy" [13] where a topic is defined by a single hashtag, or even in Castillo et al. [9] who use Twitter Monitor algorithm to formulate keyword-based queries.

Further, within the small sample of topics we examined, we discovered a variety in terms of longevity. Topics relating to everyday activities, such as seasonal flu or coffee, can be a subject to hyperbole and humor which may propagate the misinformation. However, rumors originated from known advocacy websites such as http://www.march-against-monsanto.com/ may display a spike which quickly dissipates without correction. These websites adjust their stances to the new trending topics like Zika while maintaining their core message.

Interestingly, mainstream news websites were cited at roughly the same rate in rumor tweets (39.1%) as in others

(32.3%), including the clarifications. This emphasizes the importance of authoritative sources outside mainstream news media in setting the record straight. Further, Towers et al. [37] find that mainstream news media may help spread fear and misinformation, such as in the case of Ebola in 2014, "with each Ebola-related news video inspiring tens of thousands of Ebola-related tweets and Internet searches", effectively spreading unsubstantiated panic in the United States.

Public Health Relevance.

Detecting health rumors in a timely fashion can help public health officials tackle them before they spread. However, over-reacting to a rumor might in fact increase its damage by advertising the harmful misconceptions. In the case of the Ebola outbreak, some of the rumors circulated on the Internet, such as that drinking salty water was an effective protective measure, led to several deaths [38], [39]. Rumors around a vaccination trial for a new Ebola vaccine sparked fears for a regular *Measles* vaccine, which was being used to tackle a *Measles* outbreak at the same time [40]. Public health decisions in one country can spark rumors and mistrust in another, such as when the HPV (*Human Papilloma* Virus) vaccine campaign discontinuation in Japan sparked concerns and rumors about its safety worldwide [41]. Thus, it is imperative that the impact of health-awareness campaigns is monitored in real time, as well as internationally. The tools described here can help public health practitioners in tackling the large scale of social media streams.

Further, the case of Zika is highly complex, as much uncertainty surrounded important information. For instance, the pathogenesis of microcephaly took months to be established. Previous works have highlighted the difficulty of early detection of rumors (i.e. [42]) in public health cases – to assess the veracity of a rumor can take months of public health investigations. However, due to the unprecedented scale of the crisis, health authorities started to act before a clear link between the Zika outbreak and microcephaly was established. This was especially challenging, since it can happen that apparent rumors are in fact truth. As an example, reports on narcolepsy as a side effect of a flu vaccine in the Nordic countries were first depicted as rumors, but later, few cases were confirmed and that took years of research and still it is contested [43]. Although correlation was found in epidemiological data, some scholars argue that an increase in awareness due the hype of the "vaccination crisis" might have caused the increase in cases. Public health authorities are continuously working in a complex crisis communication dilemma, since they have to act on some level of uncertainty. In this study, we chose the rumors which have been identified by authoritative sources as certain. However, a different approach may be called for the detection of possible health rumors, which is an exciting future research direction.

In this context, we believe more work is needed in the integration of rumor monitoring with public health officials, and especially the work-flow of communication departments of public health authorities. A pipeline such as AIDR (which provided our collection) described by Imran et al. [18] wherein

volunteers provide labeled social media during a disaster to train automated methods, may also be useful for ongoing health emergencies.

Limitations.

One of the main challenges of this study is that we cannot be sure about the representativeness of the social media users compared with the general population. The demographics of social media users tend to be young, and female [44], which may be important, as some have called women "gatekeepers" of their families health [45], [46]. In addition, we need to consider that particular segments of the population are more at-risk (e.g. pregnant women) and it may be difficult to identify such users online (however, tracking this particular group of users would enlighten the effect Zika has on child-baring women). Further, the limited resources of this study were applied to only a handful of rumors – those especially brought up by WHO and Snopes – and a closer collaboration with health communication experts may provide further insight into the variety of misinformation both online and its interaction with mainstream media. Finally, Zika affected many countries, and our original dataset has covered several languages. The peculiarities of rumors in each language (and by proxy, in perhaps different cultures), could illuminate differences in the perception of medical information on social media.

VII. Conclusion

This paper presents tool pipeline incorporating expert knowledge, crowdsourcing, and machine learning for health-related rumor discovery in a social media stream. Each step of the analysis was rigorously tested by manual evaluation, providing qualitative and quantitative insight into a process needed to collect data relevant to the health communication professionals.

In particular, our study shows that tracking health misinformation in social media is not trivial, and requires some expert supervision. This can then augmented by "crowd" workers in order to provide additional annotation of the captured rumor-related tweets. We 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 find traditional media sources not to be prominent in clarifying rumors, but instead show the importance of authoritative informational sources. 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.

Continued work will address the multi-lingual nature of the dataset, and expand the efforts to cross-language analysis of rumors and their potential international spread. More studies on health rumors may provide richer test beds for building automatic classifiers not just for rumors, but for the detection of informational campaigns. Finally, a user-friendly interface similar to Kostkova et al. [17], which may involve expert input, like AIDR [18], would smooth the interaction between data scientists and health communication professionals.

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," *ACM Transactions on Information Systems*, vol. 34, no. 3, pp. 1–23, 2016. [Online]. Available: http://dl.acm.org/citation.cfm?id=2915200. 2893478{%}5Cnhttp://dl.acm.org/citation.cfm?doid=2915200.2893478

[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. [Online]. Available: http://www.who.int/bulletin/volumes/89/11/11-031111.pdf?ua=1

[4] WHO. (2016) Risk communication in the context of zika virus. http://www.who.int/risk-communication/zika-virus/risk-communication-presentation.pdf?ua=1. Accessed: 2017-01-03.

[5] A. A. Ginart, S. Das, J. K. Harris, R. Wong, H. Yan, M. Krauss, and P. A. Cavazos-Rehg, "Drugs or dancing? using real-time machine learning to classify streamed dabbing homograph tweets," in *Healthcare Informatics (ICHI), 2016 IEEE International Conference on*. IEEE, 2016, pp. 10–13.

[6] I. Seaman and C. Giraud-Carrier, "Prevalence and attitudes about illicit and prescription drugs on twitter," in *Healthcare Informatics (ICHI), 2016 IEEE International Conference on*. IEEE, 2016, pp. 14–17.

[7] H. Yang and C. C. Yang, "Harnessing social media for drug-drug interactions detection," in *Healthcare Informatics (ICHI), 2013 IEEE International Conference on*. IEEE, 2013, pp. 22–29.

[8] Y. Mejova, H. Haddadi, A. Noulas, and I. Weber, "#foodporn: Obesity patterns in culinary interactions," in *DH'15: International Conference on Digital Health 2015*. ACM, 2015, pp. 51–58.

[9] C. Castillo, M. Mendoza, and B. Poblete, "Information credibility on twitter," in *Proceedings of the 20th international conference on World wide web*. ACM, 2011, pp. 675–684.

[10] J. Leskovec, L. Backstrom, and J. Kleinberg, "Meme-tracking and the dynamics of the news cycle," in *Proceedings of the 15th ACM SIGKDD international conference on Knowledge discovery and data mining*. ACM, 2009, pp. 497–506.

[11] F. Yang, Y. Liu, X. Yu, and M. Yang, "Automatic detection of rumor on sina weibo," in *Proceedings of the ACM SIGKDD Workshop on Mining Data Semantics*. ACM, 2012, p. 13.

[12] K. Wu, S. Yang, and K. Q. Zhu, "False rumors detection on sina weibo by propagation structures," in *2015 IEEE 31st International Conference on Data Engineering*. IEEE, 2015, pp. 651–662.

[13] J. Ratkiewicz, M. Conover, M. Meiss, B. Gonçalves, S. Patil, A. Flammini, and F. Menczer, "Truthy: mapping the spread of astroturf in microblog streams," in *Proceedings of the 20th international conference companion on World wide web*. ACM, 2011, pp. 249–252.

[14] J. Ma, W. Gao, Z. Wei, Y. Lu, and K.-F. Wong, "Detect rumors using time series of social context information on microblogging websites," in *Proceedings of the 24th ACM International on Conference on Information and Knowledge Management*. ACM, 2015, pp. 1751–1754.

[15] V. Qazvinian, E. Rosengren, D. R. Radev, and Q. Mei, "Rumor has it: Identifying misinformation in microblogs," in *Proceedings of the Conference on Empirical Methods in Natural Language Processing*. Association for Computational Linguistics, 2011, pp. 1589–1599.

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

[17] P. Kostkova, V. Mano, H. J. Larson, and W. S. Schulz, "Vac medi+ board: Analysing vaccine rumours in news and social media," in *Proceedings of the 6th International Conference on Digital Health Conference*. ACM, 2016, pp. 163–164.

[18] M. Imran, C. Castillo, J. Lucas, P. Meier, and S. Vieweg, "Aidr: Artificial intelligence for disaster response," in *Proceedings of the 23rd International Conference on World Wide Web*. ACM, 2014, pp. 159–162.

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

[20] World Health Organization. (2016, September) Dispelling rumours around zika and complications. http://www.who.int/emergencies/zika-virus/articles/ rumours/en/. Accessed: 2016-07-20.

[21] A. Nourbakhsh, X. Liu, S. Shah, R. Fang, M. M. Ghassemi, and Q. Li, "Newsworthy rumor events: A case study of twitter," in *2015 IEEE International Conference on Data Mining Workshop (ICDMW)*. IEEE, 2015, pp. 27–32.

[22] B. David, N. Andrew, and J. Michael, "Latent dirichlet allocation," *Journal of machine Learning research*, vol. 3, pp. 993–1022, Jan 2003.

[23] I. Ruthven and M. Lalmas, "A survey on the use of relevance feedback for information access systems," *The Knowledge Engineering Review*, vol. 18, no. 2, pp. 95–145, 2003.

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

[25] 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," *Journal of medical Internet research*, vol. 15, no. 4, p. e73, 2013.

[26] 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.

[27] L. Feng, M. Jansche, M. Huenerfauth, and N. Elhadad, "A comparison of features for automatic readability assessment," in *Proceedings of the 23rd International Conference on Computational Linguistics: Posters*. Association for Computational Linguistics, 2010, pp. 276–284.

[28] T. Mikolov, I. Sutskever, K. Chen, G. S. Corrado, and J. Dean, "Distributed representations of words and phrases and their compositionality," in *Advances in neural information processing systems*, 2013, pp. 3111–3119.

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

[30] B. Rink, S. Harabagiu, and K. Roberts, "Automatic extraction of relations between medical concepts in clinical texts," *Journal of the American Medical Informatics Association*, vol. 18, no. 5, pp. 594–600, 2011.

[31] 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," *Journal of medical Internet research*, vol. 13, no. 1, p. e14, 2011.

[32] M. R. Laurent and T. J. Vickers, "Seeking health information online: does wikipedia matter?" *Journal of the American Medical Informatics Association*, vol. 16, no. 4, pp. 471–479, 2009.

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

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

[35] T. K. Ho, "Random decision forests," in *Document Analysis and Recognition, 1995., Proceedings of the Third International Conference on*, vol. 1. IEEE, 1995, pp. 278–282.

[36] W. Du and Z. Zhan, "Building decision tree classifier on private data," in *Proceedings of the IEEE international conference on Privacy, security and data mining-Volume 14*. Australian Computer Society, Inc., 2002, pp. 1–8.

[37] S. Towers, S. Afzal, G. Bernal, N. Bliss, S. Brown, B. Espinoza, J. Jackson, J. Judson-Garcia, M. Khan, M. Lin *et al.*, "Mass media and the contagion of fear: the case of ebola in america," *PloS one*, vol. 10, no. 6, p. e0129179, 2015.

[38] S. O. Oyeyemi, E. Gabarron, and R. Wynn, "Ebola, Twitter, and misinformation: a dangerous combination?" *British Medical*

Journal, vol. 349, no. October, p. g6178, 2014. [Online]. Available: http://www.bmj.com/cgi/doi/10.1136/bmj.g6178

[39] F. Jin, W. Wang, L. Zhao, E. Dougherty, Y. Cao, C. T. Lu, and N. Ramakrishnan, "Misinformation propagation in the age of Twitter," *Computer*, vol. 47, no. 12, pp. 90–94, 2014.

[40] The Vaccine Confidence Project, "Sierra Leone: Ebola fears hamper vaccination campaigns against Polio and Measles," http://www.vaccineconfidence.org/sierra-leone-ebola-fears-hamper-vaccination-campaigns-against-polio-and-measles/, jun 2015, accessed: 2016-10-16.

[41] H. J. Larson, R. Wilson, S. Hanley, A. Parys, and P. Paterson, "Tracking the global spread of vaccine sentiments: The global response to Japan's suspension of its HPV vaccine recommendation," *Human Vaccines and Immunotherapeutics*, vol. 10, no. 9, pp. 2543–2550, 2014.

[42] Z. Zhao, P. Resnick, and Q. Mei, "Enquiring minds: Early detection of rumors in social media from enquiry posts," in *Proceedings of the 24th International Conference on World Wide Web*, ser. WWW '15.   New York, NY, USA: ACM, 2015, pp. 1395–1405.

[43] M. C. Sturkenboom, "The narcolepsy-pandemic influenza story: Can the truth ever be unraveled?" *Vaccine*, vol. 33, pp. B6–B13, 2015.

[44] M. Duggan and J. Brenner, *The demographics of social media users, 2012*.   Pew Research Center's Internet & American Life Project Washington, DC, 2013, vol. 14.

[45] N. Calabretta, "Consumer-driven, patient-centered health care in the age of electronic information," *Bulletin. Medical Library Association*, vol. 90, no. 1, pp. 32–37, 2002.

[46] D. Warner and J. D. Procaccino, "Toward wellness: Women seeking health information," *Journal of the American Society for Information Science and Technology*, vol. 55, no. 8.