Fall 12-13-2019

# VaxInsight: an artificial intelligence system to access large-scale public perceptions of vaccination from social media

Jingcheng Du

# VaxInsight: an artificial intelligence system to access large-scale public perception of vaccination from social media

A
Dissertation

Presented to the Faculty of
The University of Texas
Health Science Center at Houston
School of Biomedical Informatics
in Partial Fulfilment of the Requirements for the Degree of

Doctor of Philosophy

By

Jingcheng Du, B.S.

University of Texas Health Science Center at Houston

2019

Dissertation Committee:

Cui Tao, PhD[1], Chair
Hua Xu, PhD[1]
Yong Chen, PhD[2]
Sahiti Myneni, PhD[1]
Trevor Cohen, PhD[3]

[1]School of Biomedical Informatics, UTHealth
[2]Perelman School of Medicine, University of Pennsylvania
[2]School of Medicine, University of Washington

# Dedication

I dedicate this dissertation to my parents, Yunwei Du and Mei Tao, for their ever-present love, support, and encouragement. This dissertation is also dedicated to my loving, supportive, and beautiful wife, Linqing Su.

# Acknowledgements

I am deeply indebted to my mentors, collaborators, and friends, all of whom played a tremendous role in the completion of my dissertation.

I am truly grateful to my primary advisor, Dr. Cui Tao, for her brilliant and dedicated mentorship over the last few years. I came with an engineering background and zero experience in medical informatics. Dr. Tao guided me to where I am today. I would also like to thank my committee members: Dr. Yong Chen and Dr. Hua Xu, who have been very supportive of my research and career since the beginning. I also recognize Dr. Sahiti Myneni and Dr. Trevor Cohen for providing critical support in addressing the challenges of this dissertation research.

I am very appreciative of the mentors, fellows, and staff involved in the CPRIT-UTHealth – Fellowship in Innovation for Cancer Prevention Research. The CPRIT fellowship provided generous support and rigorous training in preparation for an academic career. In particular, I would like to thank Dr. Patricia Mullen, Dr. Roberta Ness, and Dr. David Loose for their insightful mentoring over the last few years.

I would like to express gratitude to the faculty of UTHealth and beyond for their critical guidance and support for my doctoral studies: Dr. Zhiyong Lu, Dr. Ross Shegog, Dr. Degui Zhi, Dr. Kirk Roberts, Dr. Yang Gong, Dr. Yaoyun Zhang, Dr. Jun Xu, Dr. Jiang Bian and Dr. Chongliang Luo. I would also like to acknowledge my friends and colleagues at UTHealth SBMI and the National Institutes of Health for all of their assistance.

In addition, I would like to thank CPRIT (RP160015, RP140103), NIH (R01LM011829, R01AI130460, 1R01LM012607, R01AI116794 and R01LM009012) for their support of this study. I thank Dr. Deborah Baskin for her help on English editing.

# Abstract

Vaccination is considered one of the greatest public health achievements of the 20th century. A high vaccination rate is required to reduce the prevalence and incidence of vaccine-preventable diseases. However, in the last two decades, there has been a significant and increasing number of people who refuse or delay getting vaccinated and who prohibit their children from receiving vaccinations. Importantly, under-vaccination is associated with infectious disease outbreaks. A good understanding of public perceptions regarding vaccinations is important if we are to develop effective vaccination promotion strategies. Traditional methods of research, such as surveys, suffer limitations that impede our understanding of public perceptions, including resources cost, delays in data collection and analysis, especially in large samples. The popularity of social media (e.g. Twitter), combined with advances in artificial intelligence algorithms (e.g. natural language processing, deep learning), open up new avenues for accessing large scale data on public perceptions related to vaccinations.

This dissertation reports on an original and systematic effort to develop artificial intelligence algorithms that will increase our ability to use Twitter discussions to understand vaccine-related perceptions and intentions. The research is framed within the perspectives offered by grounded behavior change theories. Tweets concerning the human papillomavirus (HPV) vaccine were used to accomplish three major aims: 1) Develop a deep learning-based system to better understand public perceptions of the HPV vaccine, using Twitter data and behavior change theories; 2) Develop a deep learning-based system to infer Twitter users' demographic characteristics (e.g. gender and home location) and investigate demographic differences in public perceptions of the HPV vaccine; 3) Develop a web-based interactive visualization system to monitor real-time Twitter discussions of the HPV vaccine.

For Aim 1, the bi-directional long short-term memory (LSTM) network with attention mechanism outperformed traditional machine learning and competitive deep learning algorithms in mapping Twitter discussions to the theoretical constructs of behavior change theories. Domain-specific embedding trained on HPV vaccine-related Twitter corpus by *fastText* algorithms further improved performance on some tasks. Time series analyses revealed evolving trends of public perceptions regarding the HPV vaccine. For Aim 2, the character-based convolutional neural network model achieved favorable state-of-the-art performance in Twitter gender inference on a Public Author Profiling challenge. The trained models then were applied to the Twitter corpus and they identified gender differences in public perceptions of the HPV vaccine. The findings on gender differences were largely consistent with previous survey-based studies. For the Twitter users' home location inference, geo-tagging was framed as text classification tasks that resulted in a character-based recurrent neural network model. The model outperformed machine learning and deep learning baselines on home location tagging. Interstate variations in public perceptions of the HPV vaccine also were identified. For Aim 3, a prototype web-based interactive dashboard, *VaxInsight,* was built to synthesize HPV vaccine-related Twitter discussions in a comprehendible format. The usability test of *VaxInsight* showed high usability of the system.

Notably, this maybe the first study to use deep learning algorithms to understand Twitter discussions of the HPV vaccine within the perspective of grounded behavior change theories. *VaxInsight* is also the first system that allows users to explore public health beliefs of vaccine-related topics from Twitter. Thus, the present research makes original and systematical contributions to medical informatics by combining cutting-edge artificial intelligence algorithms and grounded behavior change theories. This work also builds a foundation for the next generation of real-time public health surveillance and research.

# Vita

2014 to present...................................................Ph.D. candidate/student, Biomedical Informatics, University of Texas Health Science Center at Houston, USA

2013 to 2014 ...............................................................Graduate student, Electronics Engineering, Shanghai Jiao Tong University, China

2009 to 2013 .........................................................................B.S, Mechanical Engineering, Xidian University, China

## Selected First-authored Publications

- Du J, et al. Leveraging Deep Learning to Understand Health Beliefs about the Human Papillomavirus Vaccines from Social Media. Nature npj Digital Medicine 2019 Apr 15;2(1):27. (DOI: 10.1038/s41746-019-0102-4)
- Du J, et al. ML-Net: multi-label classification of biomedical texts with deep neural networks. Journal of the American Medical Informatics Association. 2019 Jun 24;26(11):1279-85. (DOI: 10.1093/jamia/ocz085, PMID: 31233120)
- Du J, et al. Gene2Vec: Distributed Representation of Genes Based on Co-Expression, BMC genomics. 2019 Feb;20(1):82. (DOI: 10.1186/s12864-018-5370-x, PMID: 30712510)
- Du J, et al. Extract Psychiatric Stressors from Twitter Using Deep Learning, BMC Medical Informatics and Decision Making 2018 Jul;18(2):43. (DOI: 10.1186/s12911-018-0632-8, PMID: 30066665, PMCID: PMC6069295)
- Huang J*, Du J*, et al. Characterization of the differential adverse event rates by race/ethnicity groups for HPV vaccine by integrating data from different sources, Frontiers in Pharmacology. 2018;9:539. (DOI: 10.3389/fphar.2018.00539, PMID: 29896103, PMCID: PMC5987677).
- Du J, et al. Public Perception Analysis of Tweets during the 2015 Measles Outbreak Using a Neural Network Model. Journal of Medical Internet Research, 20(7), p.e236. (DOI: 10.2196/jmir.9413, PMID: 29986843, PMCID: PMC6056740)

- Du J, et al. Analysis of Individual Differences in Vaccine Pharmacovigilance Using VAERS Data and MedDRA System Organ Classes: A Use Case Study with Trivalent Influenza Vaccine. Biomed Inform Insights 2017;9:1178222617700627. (DOI:10.1177/1178222617700627, PMID: 28469434, PMCID: PMC5391193)
- Cai Y*, Du J*, et al. A Signal Detection method for Temporal Variation of Adverse Effect with Vaccine Adverse Event Reporting System data. BMC Medical Informatics and Decision Making, 2017 Jul;17(2):76. (DOI: 10.1186/s12911-017-0472-y, PMID: 28699543, PMCID: PMC5506615)
- Du J, et al. Leveraging Machine Learning Based Approaches to Assess Human Papillomavirus Vaccination Sentiment Trends with Twitter Data. BMC Medical Informatics and Decision Making 2017 Jul;17(2):69. (DOI: 10.1186/s12911-017-0469-6, PMID: 28699569, PMCID: PMC5506590)
- Du J, et al. Optimization on machine learning based approaches for sentiment analysis on HPV vaccines related tweets. Journal of Biomedical Semantics. 2017 Mar 3;8(1):9. (DOI: 10.1186/s13326-017-0120-6, PMID: 28253919, PMCID: PMC5335787)
- Du J, et al. Trivalent influenza vaccine adverse symptoms analysis based on MedDRA terminology using VAERS data in 2011. Journal of Biomedical Semantics. 2016 Jun 14;7(1):1. (DOI: 10.1186/s13326-016-0056-2, PMCID: PMC5192591)

## Field of Study

Biomedical Informatics

Table of Contents

# List of Tables

# List of Figures

# Chapter 1: Introduction and Literature Review

## 1.1. Vaccine Refusal and Hesitancy

A vaccine is a biological preparation that is made from very small amounts of weak or dead germs to provide active acquired immunity for a particular disease. Vaccines are among the most effective tools available for preventing infectious diseases. As vaccines are given to healthy people, they are held to very high safety standards.[1,2] Overwhelming scientific evidence also verifies the safety of vaccines.[3] Every licensed and recommended vaccine goes through years of rigorous safety testing and is monitored continuously for safety signals after being released to the public.[4] Vaccination is the administration of vaccines. Due to the widespread deployment of vaccination programs, dramatic declines in mortality and morbidity are found for various vaccine-preventable diseases (VPD), such as polio and smallpox.[5] Vaccination is considered one of the greatest public health achievements of the 20th century.[6]

High vaccination rates are required in order to significantly reduce the prevalence and incidence of VPD.[7] However, in the last two decades, there is a significant and increasing number of people who refuse or delay vaccinations for themselves and for their children.[8–10] A recent report from the American Academy of Pediatrics (AAP) shows that "within a 12-month period, 74% of the pediatricians report encountering a parent who refused or delayed one or more vaccines.".[11] According to a nationwide survey, one in ten parents don't follow the recommended vaccination schedules for their children as developed by the U.S. Centers for Disease Control and Prevention (CDC). Even for parents who adhere to the schedule, one in four say that they feel it may not be the best or safest way to immunize youngsters.[12,13]

State-level rates of nonmedical exemptions, considered as the primary measure of vaccination refusal in the United States, also are continuing to increase.[9,14] Unvaccinated individuals not only put themselves at risk for infectious diseases but also pose a public health threat to communities.[15] Under-vaccination is associated with infectious disease outbreaks.[16] For example, the United States declared the elimination of measles in 2000. However, in recent years, there has been a resurgence in measles outbreaks in both the U.S. and elsewhere in the world.[17–19] Besides the outbreak of VPD, vaccination refusal is also associated with higher inpatient admission and emergency department utilization rates, increased morbidity, and death.[20] As a result, there is growing concern over the success of immunization promotion.[7,21]

There is great diversity in motivations for vaccination refusal or delay, including fear of adverse reactions, concerns over efficacy, distrust of government and pharmaceutical companies, and doubts about the reliability of sources of information for decision making.[15,22–24] Dubé et al summarize the major determinants of vaccination acceptance or rejection as: contextual, organizational, and individual (see *Table 1* [7]). These determinants can vary based on type of vaccine, and their impact on vaccination behaviors change over time.

An increase in anti-vaccination campaigns, which rely mostly on rhetorical arguments, incites more fear and distrust within the population. [7,25] For instance, with the emergence of Web 2.0 and the popularity of social media, anti-vaccination campaigns disseminate misinformation both rapidly and over an unprecedentedly large population.[7,26] Such dissemination may be associated with a decline in the willingness to get vaccinated .[27] This was the case following the now discredited study by Andrew Wakefield that claimed a link between autism and the MMR vaccine.[28–30]

Table 1. Main determinants of parents' vaccination decisions

| Determinants type | Determinants example |
|---|---|
| Contextual | Historical, political and sociocultural influences<br>Communication and media environment |
| Individual | Sociodemographic characteristics<br>Knowledge and attitudes<br>Past experiences with health and vaccination services<br>Trust in health system and healthcare providers |
| Organizational | Availability and quality of vaccination services<br>Health staff motivation and attitudes<br>Vaccine-specific issues |

### 1.2. Use of Social Media for Understanding Public Perceptions

A good understanding of both the causes and contexts leading to vaccination hesitancy and refusal is a first and important step in developing effective vaccination promotion strategies.[31–34] Traditional survey methods [35–37] are effective approaches with which to gather data on vaccination hesitancy and refusal. However, surveys also suffer from significant limitations: 1) most survey-based methods are expensive and labor consuming, which makes them difficult to administer to large populations;[38,39] 2) they are often unable to reach vulnerable populations such as minorities, the poor, and young adults;[24] 3) surveys require substantial amounts of time in order to collect and analyze data, creating delays in uncovering current opinions;[40] 4) survey-based methods, which often measure public opinions at a specific point in time, have difficulties in tracking changes in the reasons behind vaccination refusal and hesitancy;[15,41] 5) survey-based methods also suffer from social desirability bias, which refers to the tendency to give responses that generally are considered appropriate instead of choosing answers that accurately reflect feelings.[24,38,42,43] Therefore, in order to prevent crises from under-vaccination and in order to stop potential VPD outbreaks, new, innovative, and well-designed methods need to be developed. Such methods need to facilitate an understanding of public perceptions of vaccines across different communities in real time, in order to provide instant feedback and alerts to health professionals.

One method that can be useful is one that relies on social media. Social media is defined as “a group of Internet based applications that build on the ideological and technological foundations of Web 2.0, and that allow the creation and exchange of user generated contents”.[44] Social media is a convenient way for users to generate, share, receive, and comment on social content.[45] The popularity of mobile devices (e.g. smartphones and tablets) accelerates the penetration of social media into all aspects of life. According to *Global Digital Report* 2019,[46] 3.48 billion out of 7.68 billion people, worldwide, are active social media users. In the U.S., 72% of adults used at least one social media platform in 2019, compared with only 8% in 2008. Social media is most popular among young adults, with more than 90% of individuals between the ages of 18 and 29 using one or more platforms. However, the use of social media among other age groups also is growing. Therefore, the population of social media users is becoming more representative of the broader population.[47]

The potential of social media to improve public health is great, especially as more and more people, including the general public and health professionals, use it to share and discuss health-related information.[45] As can be seen in *Figure 1*, social media-related publications in

PubMed grew dramatically over the last two decades. This increase demonstrates great value for many health-related issues, including disease management and patient communication,[48–50] the monitoring of adverse drug reactions (ADRs), [51–54] the facilitation of medical education,[55] the promotion of behavior change,[56,57] public health surveillance,[58–60] and the detection of mental illness and suicide.[61–63]

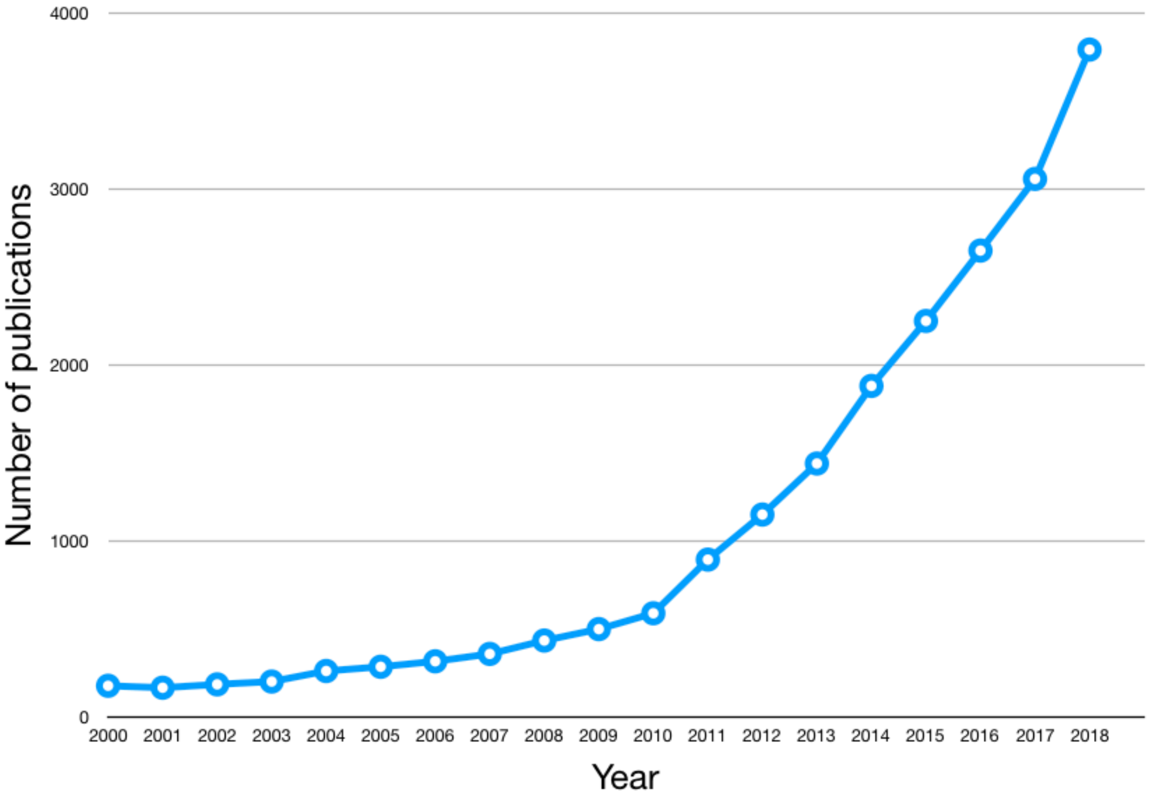


Figure 1. Social media related PubMed publications

Social media are not only sources of information about general health-related issues. They also serve as important sources for assessing public perceptions of vaccines. And, prior research documents that exposure to information on social media can impact attitudes and behavior.[64] This is observed in the use of social media by anti-vaccination movements that employ these platforms as primary communication tools.[24] Furthermore, several studies show that individuals' vaccination refusal and hesitancy is associated with vaccine-related information on the internet and social media.[7,65,66]. Thus, it is increasingly necessary to understand and monitor the information that is shared on social media platforms if health professionals aim to promote vaccination and reduce vaccine preventable diseases. This is beginning to occur around vaccines for HPV,[67–70], influenza,[58,71–75] Zika,[76–78] hepatitis,[79,80] polio,[81] measles,[82] as well as in terms of general vaccine topics.[41,83–86]

### 1.3. Existing Challenges

Although much effort has been devoted to analyzing health-related social media discussions, major challenges still exist for fully leveraging data to understand public perceptions of vaccination. Many previous studies focus on developing semi-automatic methods to understand social media discussions of vaccines, including manual coding and hashtag or keywords analysis.[69,73,88–91,74,76,79,80,82,84,85,87] However, these semi-automatic methods are limited by the lack of scalability and accuracy. In recent years, scalable automatic approaches based on machine learning (ML) have been developed to understand the contents of social media posts, including unsupervised ML methods [92,93], supervised ML methods,[43,67,75,94–96] and mixed approaches.[41,58,83] However, most of these efforts have yet to address the characterization of content at the level of granularity conducive to understanding public perceptions that provide actionable insights. These approaches also miss important data, including demographic information. This is unfortunate as such information is associated with the willingness to get vaccinated.[7,97]

Previous studies also find marked differences in vaccination coverage and health beliefs among different subgroups within the population.[97–102] For example, one study shows that women are more likely than men to believe a parent who claims that their child was injured by a vaccine.[102] Location is also associated with vaccination endorsement. Under-immunization and vaccination refusal are found to cluster geographically.[8] However, these demographic attributes, which are often collected by survey methods, are commonly missing in some social media platforms (e.g. Twitter). The lack of such information makes it challenging to investigate demographic differences in public perceptions across different subpopulations, and prevents comparisons with findings from traditional surveys.

Tracking public perceptions of vaccines in real time can assist public health professionals in examining health policies and in delivering timely responses. However, available systems use data collected by traditional survey methods and don't provide real-time information regarding vaccination attitudes and health beliefs.[103] Some pioneering work like *Vaccine Sentimeter*[104] and *VaccineWatch*[104] provide real-time surveillance regarding vaccines from internet data. However, their analyses are limited by the reliance on count data or sentiment scores. To generate actionable insights, more granular information needs to be provided, especially about health beliefs.

## 1.4. Summary

This chapter provides a discussion of vaccination refusal and hesitancy as well as their determinants and consequences. The use of social media to understand public perceptions was reviewed and both significance and challenges were discussed. To date, little is known about social media contents on a granular level (e.g. user health belief, user demographics). As a result, current research is unable to trigger actionable insights for health professionals. Additionally, it is still not clear how well these cutting-edge methods can facilitate the needed analyses, despite recent advances in machine learning and deep learning, especially in natural language processing (NLP). The present study addresses these challenges through the following specific aims:

**Aim 1: Develop a deep learning-based system to automatically understand public perceptions of the HPV vaccine from the perspective of grounded behavior changes theories and by using Twitter data.**

The present study will use Twitter data and focus on the human papillomavirus (HPV) vaccine as a use case. A deep learning-based framework will be developed to automatically align Twitter discussion feeds to primary constructs from the health belief model and the theory of planned behavior. Trends and changes in HPV vaccine perceptions using time series analyses will be identified. Relevant work on this Aim is reported in Chapter 2.

**Aim 2: Develop deep learning-based systems to infer users' demographic attributes and investigate demographic differences in public perceptions of the HPV vaccine using Twitter data.**

Deep learning-based frameworks will be developed to automatically infer demographic information from Twitter users. Specifically, the present study will focus on gender and location (i.e. map Twitter users to U.S. state levels) inference. Demographic differences in public perceptions (predicted from Aim 1) of HPV vaccines among sub-populations will be investigated using Chi-square. Relevant work on this Aim is reported in Chapter 3 (on gender inference) and in Chapter 4 (on location inference).

**Aim 3: Develop a web-based interactive visualization system to monitor real-time Twitter discussions of HPV vaccines.**

The deep learning-based frameworks developed in Aims 1and 2 will be combined in order to build a web-based online surveillance system to real-time track public perceptions of the HPV vaccine by using Twitter data. This system will allow users to track and identify changes in public perceptions of HPV vaccines within a particular population. Users can also use the system to visually compare differences in public perceptions among different subpopulations. In addition, users can verify machine generated predictions and provide feedback on the predictions in order to further improve the accuracy of deep learning algorithms. The usability of the system will be evaluated by graduate students and faculty in public health and medical informatics through the System Usability Scale. Relevant work on this Aim is reported in Chapter 5.

# Chapter 2: Deep Learning and Behavioral Theory: An Improved Analytic Method to Understand HPV Vaccination Intentions from Twitter Discussions

## 2.1. Introduction

The human papillomavirus (HPV) is the most common sexually transmitted infection in the U.S., with 14 million new HPV infections each year.[105] HPV infections cause about 33,700 cases of cancer every year in the U.S., including cervical, vaginal, vulvar, penile, and anal cancer.[106,107] The HPV vaccine has been available since 2006 to protect against HPV-associated cancers and is recommended for all girls and boys who are 11 to 12 years old. Unfortunately, compared to other recommended vaccines, the HPV vaccination rate remains suboptimal, with about half (51 percent) of adolescents failing to remain up to date.[108] This may be due, in part, to the negative attitudes towards HPV vaccination that are held commonly by parents of adolescents.[109] Nonetheless, unvaccinated individuals are more vulnerable to the virus and pose a public health threat to communities.

Historically, vaccination-promotion strategies have been informed by knowledge-deficit models that attribute public skepticism or ambivalence to a lack of information and understanding.[110,111] There is still a significant proportion of the public, however, that opposes the vaccine, regardless of the evidence.[111] The factors associated with HPV vaccination hesitancy are multi-dimensional and vary by individual.[109,112–115] Incorporating behavior change theories into research can facilitate an understanding of these factors and aid in the development of strategies to influence specific health-related behaviors.[118] For instance, associations are found between theoretical constructs within the health belief Model (HBM),[116–119] and the theory of planned behavior (TPB),[120–122] as they relate to HPV vaccination intention and uptake. HBM and TPB have demonstrated efficacy in providing a description of the antecedents to vaccination behavior, including predicting the likelihood of behavioral initiation.[118,120,123]

Social media offer opportunities to reach large populations, while mitigating the limitations of traditional surveys, including resource costs, inability to track changes in real time, and delays between data collection and availability.[15,40,41] Millions of users generate, share, receive, and comment on various topics, including those in health-related domains.[124] Social media offer an unprecedented level of reach to improve public health.[45] Initial semi-automatic methods to understand social media discussions on vaccines have included manual coding and hashtag or keywords analysis,[80,82,86,90] but these are limited, respectively, by their lack of scalability and inaccuracies. Machine-learning methods have emerged to address these limitations and to improve the precision with which public perception of vaccines can be understood.[94,96] This is particularly true of the HPV vaccine.[67,92,93,95,125,126] Recent studies apply deep-learning methods to analyze social media (e.g., Twitter) messaging data [127–129] and demonstrate the superiority of deep learning in comparison with traditional machine-learning efforts.[130,131] Deep learning is a set of advanced machine-learning algorithms that achieve state-of-the-art performance in many natural language processing (NLP) tasks.[132–134] It does so while dramatically reducing the overhead of feature engineering that is required for most traditional machine-learning-based approaches. Prior work also demonstrates the superiority of deep-learning-based methods.[131]

The purpose of the present study is to leverage deep-learning and machine-learning algorithms to automatically align Twitter discussion feeds to HBM and TPB constructs. This study is significant in its (1) description of a deep learning-based framework to map Twitter discussions to the constructs of multiple behavior change theories; (2) comprehensive evaluation of deep-learning and machine-learning algorithms for this purpose; and (3) identification of trends and changes in HPV vaccine perceptions using time series analyses. The study's innovative way in which to categorize messages is informed by theory-based constructs. As a result, attitudes can be differentiated and future resource messaging can be fine-tuned.

## 2.2. Materials and Methods

### 2.2.1. Study overview

Twitter is one of the most popular social media platforms in the world, with 326 million active users, monthly.[135] Twitter is recognized as one of the major and most credible sources for accessing public opinions on various topics, from politics [136] to public health.[137] However, an accurate understanding of Twitter discussions is considered challenging in light of the unique characteristics of Twitter text (e.g., short text, cyber slang, emoticons).[138] The current study provides (1) a comprehensive evaluation of deep-learning and machine-learning methods on Twitter text classification for theoretical constructs informed by HBM and TPB; (2) a prediction of the constructs from the un-label Twitter dataset, using pre-trained deep-learning models; and (3) a description of trends in these theoretical constructs over time, using time series analyses. An overview of the study design can be seen in *Figure 2*.

This study received expedited review and IRB approval from the Committee for the Protection of Human Subjects at The University of Texas Health Science Center at Houston. Waiver of informed consent was granted by the IRB due to the retrospective design of the study. The approved IRB protocol number is HSC-SBMI-16–0291.

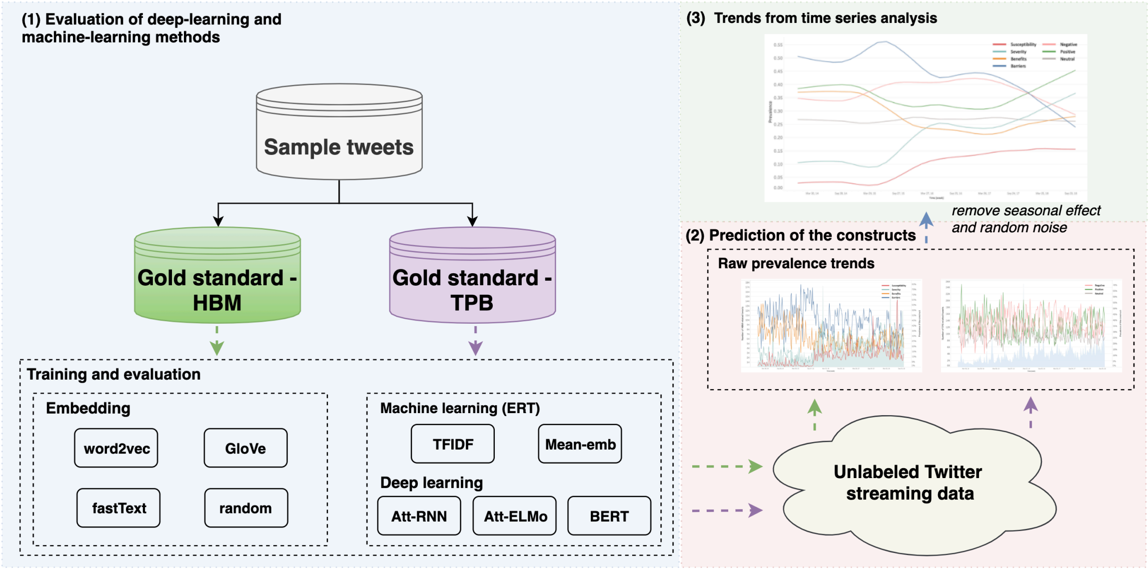


Figure 2. Overview of the study design

### 2.2.2. Operationalization of theoretical constructs

The principal HBM constructs used in this study are *perceived susceptibility* (e.g., likelihood of contracting HPV), *perceived severity* (e.g., degree of negative health effects of HPV), *perceived benefits* (e.g., positive outcomes of getting the HPV vaccine), and *perceived barriers* (e.g., negative aspects of HPV vaccination). HBM constructs, definitions, and examples are provided in (*Appendix A*). TPB constructs are represented by the attitudes construct that is comprised of fact-based statements, inclusive of reference to norms and/or behavioral controls but defined by *positive*, *negative*, and *neutral* valence. There are several other constructs in TPB that also could influence HPV vaccination behavior, such as *subject norm* and *perceived behavior control.* However, the low prevalence of these constructs in Twitter discussions, precludes their inclusion in this present study. Therefore, this study covers only the major constructs noted above (*Figure 3*).

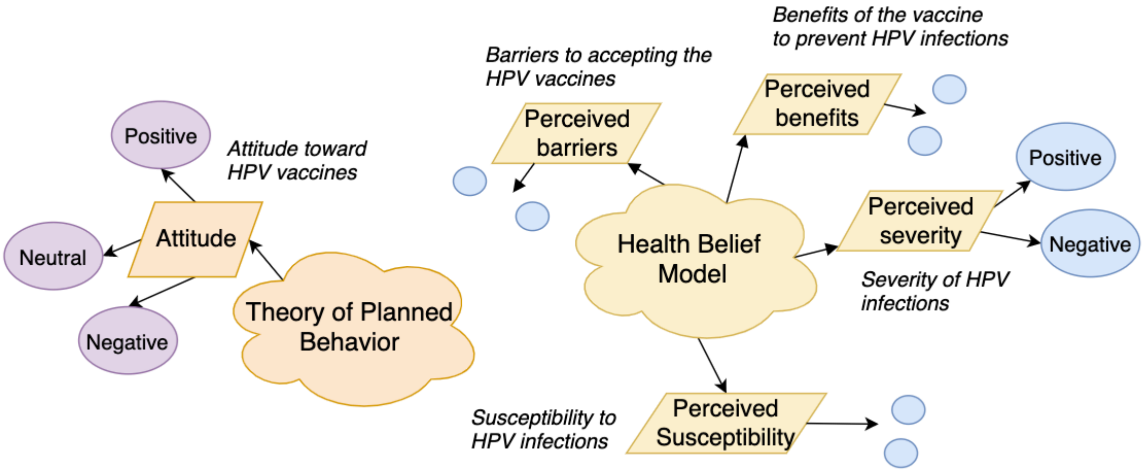


Figure 3. Key constructs of HBM and TPB in this study

### 2.2.3. Twitter corpus collection and annotation

A set of keywords (i.e., HPV, human papillomavirus, Gardasil, and Cervarix) was used to collect English-language tweets by using Twitter streaming API (~1% of the entire stream volume) from January 1, 2014, to October 26, 2018. A total of 1,431,463 English-language tweets were collected. The gold-standard corpus (tweets with the annotation) was acquired from a previous study.[131] Specifically, for HBM, the focus was on the four primary constructs: *perceived susceptibility*, *perceived severity*, *perceived benefits*, and *perceived barriers* (*Appendix A*). Three reviewers (two graduate-level research assistants and one postdoctoral fellow in health informatics) were trained. They then categorized a subset of 6,000 tweets based on the relevance of these tweets to the HBM constructs. Each tweet was assigned to none (not related to HBM), one, or multiple HBM constructs. The inter-annotator agreement for each construct ranged from 0.727 to 0.834. For TPB, the focus was on an amalgamated construct of *attitude*. Gold standard data were acquired from a previously published study.[139] Three reviewers (two graduate-level research assistants in public health and one graduate-level research assistant in health informatics) categorized the same 6,000 tweets based on attitudes toward the HPV vaccine as expressed within the tweets. The reviewer first decided whether the tweet was related to the construct of attitude toward the HPV vaccine. If it was related, the reviewer further decided whether it was *positive*, *negative*, or *neutral*. The overall inter-annotator agreement was 0.851.

### 2.2.4. Deep learning-based framework

The understanding of Twitter content within the perspective of behavior change theories was framed to text-classification tasks. Deep-learning classifiers for HBM and TPB were built. For the four primary HBM constructs, the tweet was categorized first based on whether it was relevant to any of the HBM constructs and then relevant tweets were categorized into the four primary HBM constructs, using binary classification (one classifier for one construct). A similar process was followed for TPB constructs. A tweet was categorized first in terms of its relevance to the construct of attitude toward the HPV vaccine; and then the relevant tweet was categorized according to one of three attitudes: *positive*, *negative*, and *neutral*.

*Twitter word embedding*

Many machine-learning and almost all deep-learning algorithms are incapable of processing strings and text in their raw form. Pre-trained word embedding, which provides distributed representations of words in a vector space, can help learning algorithms achieve better performance in natural language processing tasks.[140] Various word embedding models have been proposed in recent years, including *word2vec*,[140] *GloVe*,[141] and *fastText*.[142]

*Word2vec* is one of the most popular techniques to learn word embedding There are two main training algorithms for *word2vec*: continuous bag of words (CBOW) and skip-gram. CBOW uses the context of the word to predict a target word while skip-gram uses a word to predict a target context. For the present study, skip-gram was chosen as the algorithm as it works better for infrequent words.[143]

*GloVe* stands for global vectors for word representation. *GloVe* was proposed by Pennington et al as a count-based method to learn word vectors. Different from *word2vec,* which leverages the predictive model (i.e. neural network) to learn word vectors, *GloVe* learns word vectors from aggregated global word-word co-occurrence.

*fastText* is a more recent method of word embedding, which is based on a skip-gram model. However, contrary to *word2vec*, where the morphology of words is ignored, each word is represented as a bag of character n-grams in *fastText*. A word vector representation is associated with each character n-grams.

Twitter word embedding was trained by applying the above three models to the un-labeled HPV-related Twitter corpus (which were termed W2V HPV, GloVe HPV, and FT HPV). For all three models, window size was set at 5, maximum iteration at 20, and dimension size at 200. The use of these Twitter word embeddings was evaluated on a recurrent neural network (RNN) with attention mechanism.[131] For comparison purposes, the use of pre-trained 200-dimension GloVe Twitter embedding (trained 2 billion tweets from the general domain, which we term GloVe General) and the use of random embedding were also evaluated.

*Deep learning-based classifiers*

Twitter text classification tasks often are considered more challenging due to the many unique characteristics of Twitter text, such as very short text, frequent occurrence of incorrect spellings, cyber slang, and emoticons.[138,144,145] Deep learning based approaches significantly improve performance in text classification tasks.[146–150] Thus, three competitive deep learning-based algorithms were evaluated in the current study: *Att-RNN*, *Att-ELMo* and *BERT*.

*Att-RNN* is an RNN with attention mechanism.[131] *Att-RNN* consists of several layers, including: 1) a word embedding layer to map Twitter text token into high dimensional vectors; 2) a bidirectional long short-term memory (Bi-LSTM) layer [151] to capture both forward and backward information of the corresponding Twitter text; 3) an attention layer to add on top of the Bi-LSTM layer to further augment the sequence model by capturing the salient portions and context; 4) a Softmax layer as the output layer for classification. The architecture of *Att-RNN* can be seen in *Figure 4*.

*Att-ELMo* is an attentive sequence model based on the Embeddings from Language Models (ELMo).[152] Traditional word embedding methods assign a static high dimensional vector to a word, regardless of its context. However, a word could have multiple context-dependent meanings. ELMo is a deep contextualized word embedding method that can look at the entire context before assigning each word its embedding vector. *Att-ELMo* first adopts the pre-trained ELMo (which was loaded from https://tfhub.dev/google/elmo/2) to map each word in the tweet to high-dimensional vectors. Then, similar to *Att-RNN*, word vectors are then fed to a bidirectional RNN, followed by the attention mechanism. A Softmax layer serves as the output layer for classification.

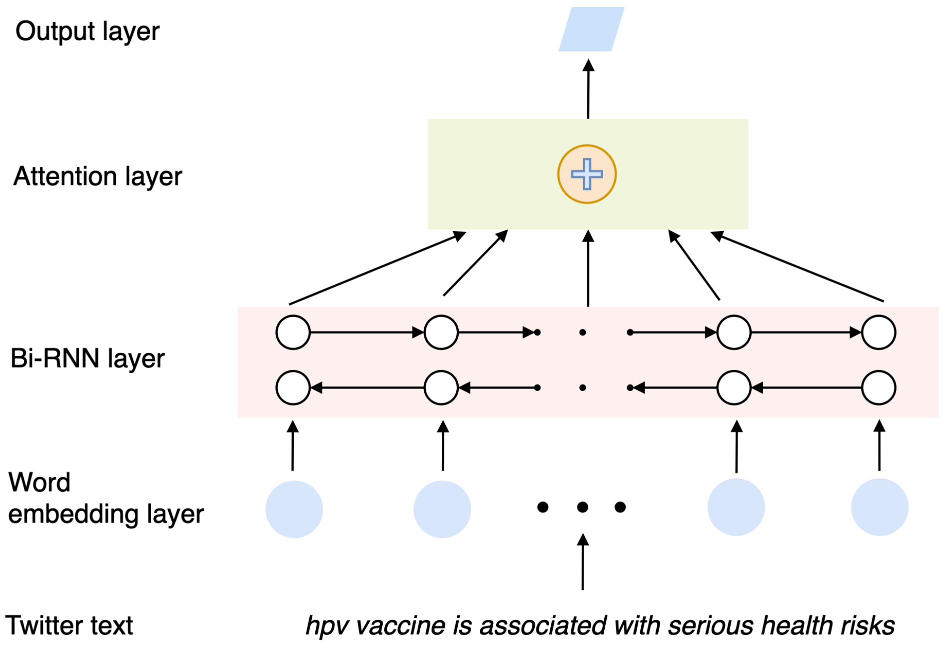


Figure 4. The architecture of the attentive recurrent neural network (RNN) for Twitter text classification

*BERT* stands for Bidirectional Encoder Representations from Transformers. *BERT* is a new language representation model based on Transformer architecture.[153] Transformer relies entirely on self-attention to compute representations of its input and output without using sequence-aligned RNNs (e.g. LSTM).[154] Contrary to recurrent models, Transformer allows for significantly more parallelization. *BERT* achieved state-of-the-art performance in 11 natural language processing tasks.[153] A pre-trained *BERT* model can be fined tuned with just one additional layer to other tasks. The pre-trained *BERT* model (BERT-Large, Uncased) was loaded and fine-tuned in the present study's Twitter text classification dataset.

### 2.2.5. Evaluation

*Machine learning-based classifiers*

Several classic machine-learning algorithms (e.g., support vector machines, logistic regression, random forest) were tested and extremely randomized trees were chosen (ERT)[155] as the baseline algorithm due to its better performance on most of the tasks. Two types of features were evaluated: (1) mean-embedding - all of the tokens were mapped to high-dimensional vectors using pre-trained word embedding and took the averaged word vectors for all words in each tweet as the feature (which was termed *mean-emb*) and (2) term frequency-inverse document frequency (TF-IDF) - TF-IDF is a numerical statistic that is intended to reflect how important a word is to a document in a corpus.[156]

### *Experiment setting*

Machine-learning and deep-learning algorithms on HBM and TPB were evaluated. These are both two-step classifications: (1) the content of a tweet was first classified as to whether it was related to HBM or TPB. In this step, all gold-standard tweets (6,000 in total) were divided into training, validation, and testing sets with a proportion of 7:1:2; (2) then, for HBM-related (3,264 in total) and TPB-related (3,984 in total) tweets, the tweets were classified into each specific construct. The HBM and TPB-related tweets were divided into training, validation, and testing sets with a proportion of 7:1:2, respectively. For each task, the models were trained on the training set; hyper-parameter tuning was performed on the validation set; performance was evaluated on the testing set. Random sampling of the tweets was repeated 10 times (with replacement) with the same proportion and the average metrics for each model were calculated.

### *Evaluation metrics*

For all of the binary classifiers (e.g., to classify the tweet as HBM-related or TPB related or to each HBM construct), sensitivity, specificity, accuracy, precision, recall, and F-1 score were calculated. For the multi-class classifier (i.e., to classify the tweet into one of three attitudes), overall accuracy, as well as precision, recall, and F-1 score for each attitude (i.e., positive, negative, or neutral) were computed.

### 2.2.6. Time series analyses of trends for theoretical constructs

*Prediction*

The best-performing model (i.e., *Att-RNN* with FT HPV embedding) was selected for predicting un-labeled data in the tweets collection. Random sampling of the tweets was repeated and trained on the *Att-RNN* model 10 times (the same strategy described in the Experiment setting) in order to reduce variance in the deep-learning models.[157] The final prediction of the all the un-labeled tweets was based on majority voting of the predictions from 10 models.

*Time series analysis*

The prevalence of each construct was defined by calculating the ratio of the count of tweets that were classified to that construct to the total count of tweets that were classified to the corresponding theory. The prevalence of each construct was calculated for each week. Time series analyses were conducted on the weekly prevalence data in order to extract the trend of the constructs. Specifically, prevalence was decomposed into seasonal, trend, and random noise components using locally estimated scatterplot smoothing (LOESS).[158] The decomposition was done by the R function “stl”.[159]

## 2.3. Results

### 2.3.1. The impact of word embedding

The performance of the different word-embedding techniques can be seen in *Appendix B*, *Appendix C*, and *Appendix D*. The use of pre-trained Twitter word embedding boosted overall performance compared with the use of randomly initialized embedding. The present study trained domain-specific embedding on a relatively small corpus (i.e., HPV vaccine-related Twitter corpus). It provided comparable performance with the general Twitter embedding (GloVe Twitter), trained on billions of tweets on the tasks. In particular, the use of FT HPV embedding led to the best performance in the majority of the tasks.

### 2.3.2. The comparison of classification algorithms

The comparison of different classification algorithms can be seen in *Appendix E*, *Appendix F*, and *Appendix G*. Machine learning with the TFIDF feature achieved higher accuracy and F-1 score than did the same learning algorithm with the mean-emb feature. In general, deep-learning algorithms (e.g., *Att-RNN*, *Att-ELMo*, *BERT*) demonstrated superiority over machine-learning algorithms on most of the tasks. The *Att-RNN* model achieved the best accuracy on almost all tasks (except for the TPB-related task, for which it ranked second) and the best micro-average F-score on TPB-attitude classification. The *BERT* model achieved the highest accuracy on the TPB-related task, as well as the highest F-1 score on HBM-barriers, HBM-benefits and TPB-related tasks.

### 2.3.3. Trends for theoretical constructs

There were dramatic fluctuations in the prevalence of each construct (*Figure 5* and *Figure 6*). In addition, there were increasing trends in the total count of theory-related tweets (i.e., *HBM-related*, and *TPB-related*) over the years of the study. These fluctuations demonstrated increasing interest in discussing the HPV vaccine on Twitter. Time-series analyses further extracted smooth trends for each construct (*Figure 7*). As can be observed among the HBM-related constructs, there was a decreasing trend in the prevalence of *barriers* and an increasing trend in the prevalence of *severity*; the prevalence of *benefits* decreased from early in 2015 to the middle of 2016 and remained stable thereafter; *susceptibility* showed an opposite trend, as prevalence started increasing from early in 2015 and remained relatively stable after the middle of 2016. Among attitudes toward the HPV vaccine, *neutral* attitude stayed stable over the years; from the middle of 2017, *positive* attitude toward the HPV vaccine showed an increasing trend, while *negative* attitude showed a decreasing trend.

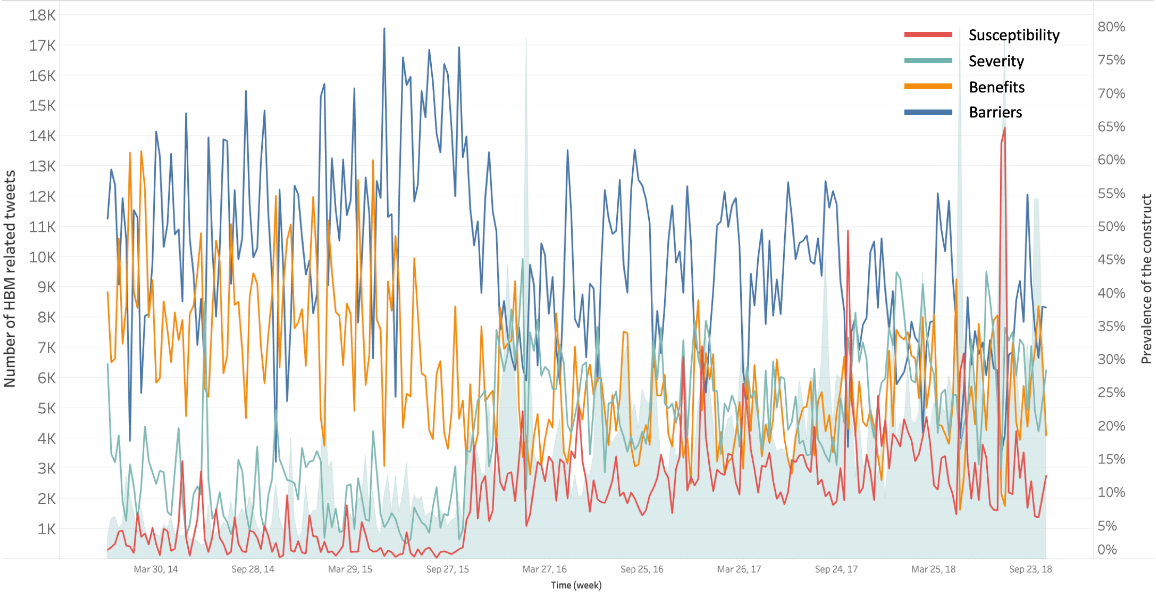


Figure 5. Prevalence of constructs from health belief model (HBM). The green shadowed area represents the total count of HBM-related tweets for each week, and the colored lines represent the prevalence of each construct

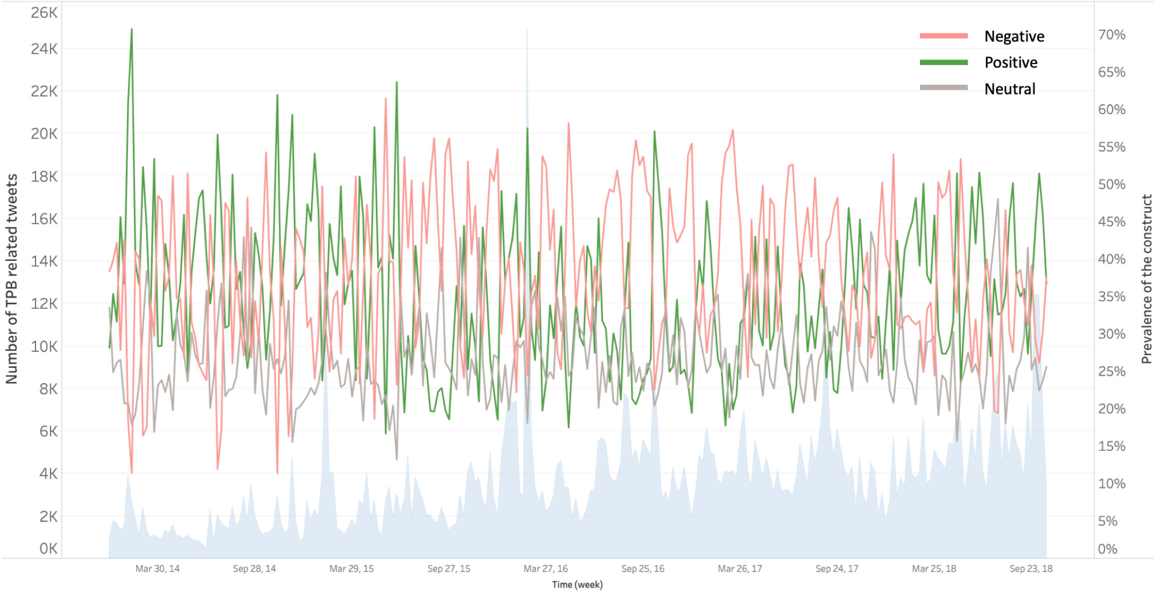


Figure 6. Prevalence of attitude from theory of planned behavior (TPB). The blue shadowed area represents the total count of TPB-related tweets for each week, and the colored lines represent the prevalence of each attitude

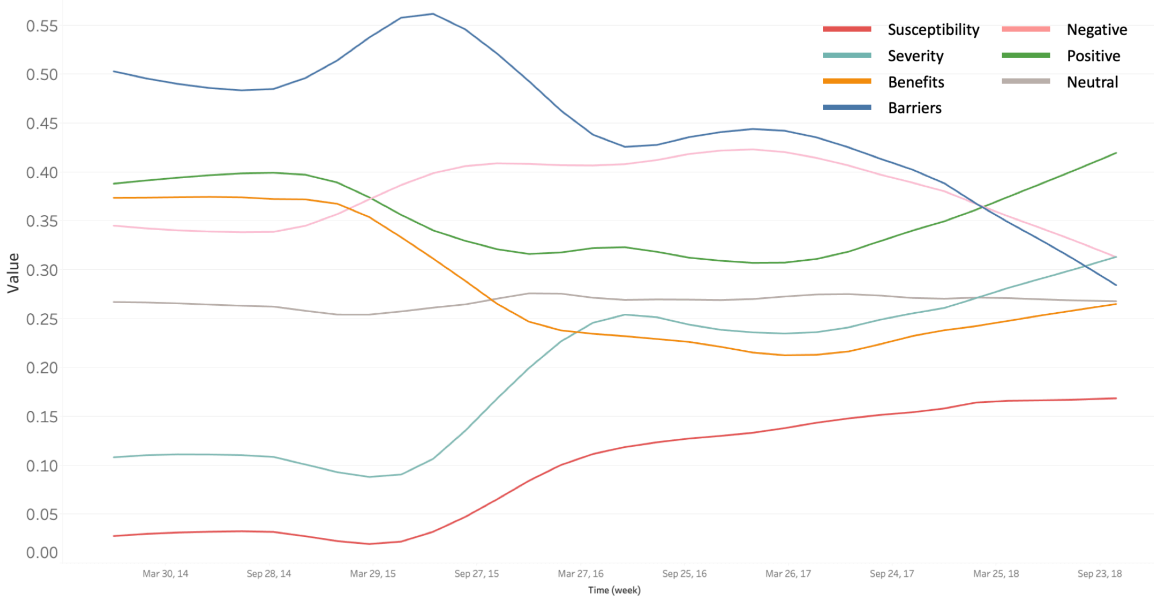


Figure 7. Trends of theoretical constructs after removing seasonal effect and random noise

## 2.4. Discussion

One major goal of the present study is to evaluate computational algorithms for the classification of HPV vaccine Twitter discussions, as they relate to the constructs of behavior change theories. One finding is that pre-trained word embedding can improve the performance of deep-learning models. Word embedding is particularly important for deep learning-based NLP models, for which it can provide a dense representation of the semantic information for the words. Although GloVe Twitter embedding has been widely adopted in Twitter-related NLP systems,[130,131,160,161] it is valuable to see that the domain-specific embedding trained on only ~1.4 million tweets can provide comparable or even better performance on the majority of tasks. Although domain-specific embedding might not be able to capture all of the necessary semantics, it can be more representative of specific domain-related tasks.[162]

Another finding is that machine-learning algorithms with the TFIDF feature provide relatively strong baseline performance on most of the tasks. Although deep-learning algorithms achieve better performance, the improvement over machine-learning algorithms on most of the tasks is not dramatic. Machine-learning algorithms can still be a good option for some tasks especially in light of the higher computation cost of deep-learning algorithms. BERT, a recent breakthrough in NLP, has advanced state-of-the-art performance in multiple general domain NLP tasks.[153] A few studies in the Twitter domain also show the superiority of BERT over other machine-learning and deep-learning algorithms.[163,164] The present study also shows that the default BERT can achieve performance comparable to the best algorithm (i.e., *Att-RNN*). Recent studies indicate that the transfer learning of BERT in the biomedicine domain can advance existing state-of-the-art performance.[165] It can be expected that the transfer learning of BERT to the Twitter domain can further advance performance on Twitter-related tasks.

The retrospective analysis of theoretical constructs, including health beliefs and attitudes toward the HPV vaccine, provide a better understanding of public perceptions and their evolving trends in terms of multiple dimensions. Trends in public perceptions could be one of the measures of health policy efficacy and a signal for public vaccination acceptance. The present study demonstrates that, in recent years, *positive* attitude toward the HPV vaccine is increasing in Twitter space, which could benefit from the significant efforts of public health professionals on HPV vaccination promotion. The increase in *perceived severity* could have resulted from the shift in promotion strategy. This shift placed more emphasis on the importance of the HPV vaccine for cancer prevention than for genital warts. The Centers for Disease Control and Prevention's (CDC) initiative to use a cancer-oriented message may have pushed providers to be proactive and presumptive in recommendations. The CDC also bundled the HPV vaccine with the larger group of mandated vaccines, which, in turn, may be having an effect on trends in public perceptions. Importantly, the approach used in the current study permits an automatic understanding of an individual's health beliefs and attitudes toward the HPV vaccine, which could facilitate further innovative and customized vaccination promotion strategies.

A limitation of the present study is that the treatment of predicted labels as true labels for the time-series analyses could have led to information bias due to misclassification rates.[166,167] A further limitation was that the gold-standard corpus was limited to 6,000 tweets. This may not be representative of the un-labeled tweets collection (~1.5 million tweets), and the shift in data distribution between labeled and un-labeled data could have brought additional bias to the prediction. To mitigate this, it is recommended that future studies add more representative tweets to the gold-standard corpus. In addition, other novel computational frameworks (e.g., domain-adversarial training[168]) may overcome the shift in data distribution. Future evaluation of these frameworks is indicated.

## 2.5. Conclusion

The present study evaluated various machine-learning and deep-learning algorithms in order to map HPV vaccine-related Twitter discussions to the constructs of grounded behavior change theories. Deep-learning algorithms outperformed machine-learning algorithms on the tasks. Pre-trained word embeddings were effective in improving the performance of deep-learning algorithms. Domain-specific embedding led to comparable or higher performance compared with pre-trained embedding from the general domain. Time-series analyses on the predicted constructs revealed evolving trends in public perceptions of the HPV vaccine. This study could benefit vaccination promotion programs by providing an automatic understanding of the population- and individual-level health beliefs and attitudes toward the HPV vaccine. A similar strategy can be applied to an understanding of public perceptions on other health-related topics, such as measles outbreaks, influenza vaccines, and so forth.

# Chapter 3: Gender Differences in Public Perceptions of the HPV Vaccination

## 3.1. Introduction and Related Work

Gender plays an important role in shaping awareness, perceptions, and intentions as they relate to vaccinations. In one study, women students in Pakistan show significantly greater awareness of vaccines for hepatitis than do men.[175] Perceptions of the risks associated with vaccinations also differ by gender, with women expressing more concern over efficacy and safety.[97, 174] And, men and women are dissimilar in their intentions to get vaccinated, despite the fact that neither the disease nor the vaccine, itself, is specific to any one gender. Research in France, Greece, and the Netherlands suggests that men are more willing to get vaccinated.[97]

In addition, although HPV vaccination rates are rising, there are marked gender differences, with 65% of girls and only 56% of boys receiving a first dose.[169] Survey research also demonstrates gender disparities in health beliefs and knowledge about HPV and HPV vaccines.[100,101,170–172] These patterns may be due, in part, to the feminization of HPV in which concerns over sexuality and the transmission of disease tend to be greater for girls than they are for boys [176]. Such gendered perspectives may have led, as well, to the fact that the vaccine was first approved for girls, only.

Research on public perceptions of the HPV vaccine is critical. However, reliance on social media platforms, such as Twitter, in order to obtain data, is problematic. While survey-based methods tend to gather gender and other demographic data, such data are often missing or hard to discern on social media. The lack of demographic information makes it challenging to investigate differences in public perceptions across different subpopulations and to validate and compare findings from studies using traditional survey-based methods. Nonetheless, there are some recent efforts aimed at understanding public perceptions of vaccinations that apply machine or deep learning approaches to Twitter posts [94,131]. But, they are unable to address gender differences due to the lack of reliable gender information provided in tweets. In an attempt to overcome this limitation, Huang et al leveraged the Demographer [173] to infer gender attributes from Twitter users [75] and studied how vaccine tweet counts varied by gender. However, Demographer and other name-based inferring tools [174,175] are insufficient as the name of specific Twitter users is sometimes not available, is inaccurate, or is gender neutral.

Other approaches using traditional machine learning-based approaches have been tried. These studies framed gender prediction as binary classification tasks and proposed machine learning-based approaches (e.g. support vector machines) with extensive feature engineering for gender prediction. Typically, word and character-level n-grams based-features were used [176,177]. Additional elements, including emojis, part-of-speech (POS) tags, latent semantic analysis(LSA), and lexicon features also were adopted in some machine learning-based systems [178,179]. Demographic attributes, such as gender, were found to be associated with linguistic features of user-generated posts [180,181]. Overall, traditional machine learning-based approaches with extensive feature engineering achieved high performance in Twitter gender inference challenges.

The use of deep learning-based methods has advantages over traditional approaches in feature engineering and can achieve state-of-the-art performance in many natural language processing (NLP) tasks.[134] This was attempted with Twitter, specifically with regard to gender inference. However, the performance of deep learning-based methods using Twitter textual data were found to be suboptimal compared with traditional machine learning-based approaches that had thorough feature engineering. For example, in a 6th Author Profiling Task at PAN 2018,[182] the top three systems in gender prediction (using English Twitter textual data only) all adopted non-deep learning-based approaches. Sierra et al leveraged a feed-forward neural network with *fastText* embedding and ranked 4th in the English Twitter text category;[183] Takahashi et al designed a recurrent neural network for text and ranked 7th in that category.[184] In a 5th Author Profiling Task at PAN 2017,[185] Miura et al proposed a neural attention network to integrate both word and character information;[186] however, this study only ranked 6th in English gender inference.

In the current work, a novel deep learning-based approach was employed to infer the gender of Twitter users by utilizing English Twitter textual data as the input. As a use case, the model also was leveraged to investigate gender differences in public perceptions of the HPV vaccine on Twitter space. . The present study contributes to the field in three main ways:

1) The character-based convolutional neural network model with embedding fusion for gender inference using English Twitter textual data as the input compared favorably with the state-of-the-art performance in a recent Twitter Author Profiling (i.e. gender inference) Challenge Task.
2) Multiple competitive machine learning-based and deep learning-based algorithms related to gender inference were compared.
3) An evaluation of the HPV vaccine-related Twitter corpus (described in Chapter 2) identified gender differences in public perceptions of the HPV vaccine. The findings were largely consistent with previous survey-based studies.

## 3.2. Method

There were two major steps in the present study: 1) to evaluate a convolutional neural network-based deep learning model for English Twitter gender inference and then to use the model in a recent open challenge task: 6th Author Profiling Task at PAN 2018;[182] 2) to leverage the trained model for gender inference on Twitter users who discussed the HPV vaccine and then to uncover any gender differences in public perceptions regarding the vaccine.

### 3.2.1. Datasets

*Author Profiling dataset*

Author Profiling Tasks at PAN are a series of international challenges which aim to classify texts into classes based on the stylistic choices of their authors. In the current study, models were evaluated on the Author Profiling Tasks at PAN 2018, which is focused on gender identification in Twitter. In this challenge and for each Twitter user, a total of 100 tweets and 10 images were provided. Three datasets in different languages, including English, Arabic and Spanish were provided. The current study then focused only on English tweets and only on text data (image data were excluded). This process provided a balanced corpus with regard to gender. It resulted in 3,000 Twitter users for training and 1,900 Twitter users for testing.[182] The training dataset was used to develop a gender identification model that evaluated the model's performance on the testing dataset.

*Twitter HPV vaccine-related dataset*

A set of HPV vaccine-related keywords were employed to collect 1,431,463 English tweets by using Twitter streaming API between the dates of Jan. 1, 2014, to Oct. 26, 2018. A subset of 6,000 tweets was annotated based on its relevance to four health belief model (HBM) constructs (i.e. *perceived susceptibility*, *perceived severity*, *perceived benefits*, and *perceived barriers*) and one theory of planned behavior (TPB) construct (i.e. *positive*, *negative*, and *neutral attitudes*). Gold standard data were used to train and evaluate an attentive recurrent neural network. Random sampling of the tweets was repeated in the gold standard corpus and training 10 times for each construct. The final prediction of all the un-labeled tweets was based on a community ensemble (i.e. majority voting) of 10 models. A detailed description of the Twitter data and of the models is provided in Chapter 2's Method section.

Up to 100 tweets were then collected from each unique Twitter user in our corpus from December 2018 to January 2019 by using Tweepy.[187] There were 486,116 unique Twitter user IDs derived from 1,431,463 tweets. From these tweets, 275,753 of these IDs were still active during the collection period. Tweets from those Twitter accounts that were not valid during the entire collection period were removed from the Twitter corpus. After excluding inactive users' tweets, 1,052,770 tweets (73.54%) were included in the final analysis. Of those tweets that were included in the analysis, 740,910 were classified as related to HBM. Additionally, 75,376, 164,793, 202,566 and 317,863 tweets were classified, respectively, as *perceived susceptibility*, *perceived severity*, *perceived benefits*, and *perceived barriers*. Furthermore, 761,449 tweets were classified as related to the TPB attitude, among which 275,203, 280,532 and 205,714 tweets were classified as *positive*, *negative*, and *neutral,* respectively. Among active Twitter users, 266,316 (96.58%) had at least 100 tweets during the December 2018 to January 2019 time frame. The average number of collected tweets for each user was 98.05. Other basic statistics from these datasets for training, testing, and prediction are summarized in *Table 2*.

Table 2. Descriptors and basic statistics for three Twitter gender datasets

| Dataset | Number of Twitter users | Number of tweets per user | | | | Number of tokens per user | | | |
|---|---|---|---|---|---|---|---|---|---|
| | | Mean | Min | Max | SD | Mean | Min | Max | SD |
| Training | 3,000 | 100 | 100 | 100 | 0 | 1,681 | 412 | 3,338 | 306.94 |
| Testing | 1,900 | 100 | 100 | 100 | 0 | 1,679 | 754 | 2,765 | 297.30 |
| Prediction | 275,753 | 98.05 | 1 | 100 | 11.66 | 1,916 | 2 | 7,995 | 551.78 |

### 3.2.2. Convolutional neural network with embedding fusion

The convolutional neural networks (CNN) model is commonly used in various computer vision tasks [188] and demonstrates excellent performance in the NLP field, including text classification tasks.[127,146,189,190] CNN utilizes layers with convolving filters to extract local features. It can be trained in parallel in order to reduce training time as compared to the recurrent neural networks (RNN) model, which needs to be processed sequentially. Kim proposed a simple CNN with one layer of convolution on top of word embeddings followed by a max-pooling layer and achieved excellent results on multiple sentence classification benchmarks.[146]

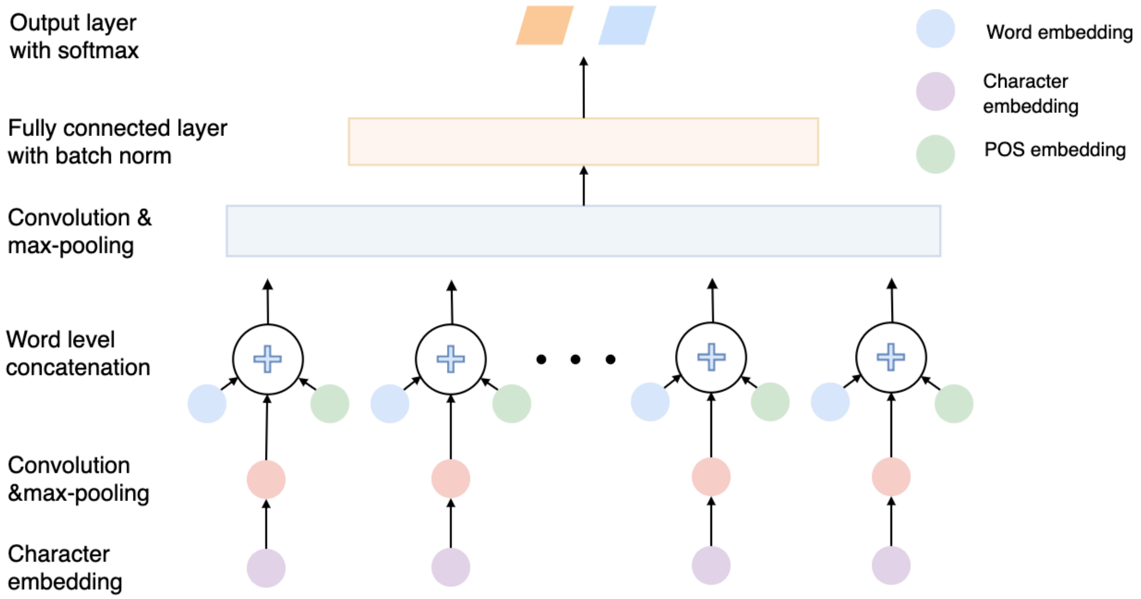


Figure 8. The architecture of convolutional neural network for Twitter gender prediction

The present study extended Kim's CNN model by adding embedding fusion to model word morphology information and sentence syntactic information from the Twitter posts. The overall architecture of the framework is illustrated in *Figure 8*. Specifically, a character layer was designed which took the character embedding of each character within a tweet token as the input and output the summary of characters for each token using a convolutional layer followed by a max-pooling layer. The character layer ensured mapping both the in-vocabulary words and the out-of-vocabulary words (e.g. incorrect spellings) to high dimensional vectors. The word layer concatenated the output of the character layer, word embedding and part of speech (POS), and concatenate them together in order to more comprehensively represent the linguistic features of each word. The output of the word layer then was fed to another convolutional layer and max-pooling layer to represent the information of the Twitter user. A dense layer was added with batch normalization on top of the pooling layer. The output layer was a fully connected layer with Softmax outputs. L2 regularization and dropout was added to avoid overfitting. The major parameter settings for the proposed model can be seen in *Table 3.* This deep neural network was termed CNN_char_pos.

Table 3. Major hyper-parameter settings for the proposed character-based convolutional neural network with embedding fusion

| Hyper-parameter | Setting |
|---|---|
| Learning rate | 0.001 |
| Batch size | 64 |
| POS embedding dimension | 10 |
| Character embedding dimension | 50 |
| Pre-trained word embedding | GloVe Twitter embedding (d=200) |
| L2 regularization | 0.00001 |
| Dropout rate | 0.2 |
| No. of filters: word/character level | 2048/50 |
| Maximum length of token | 2,000 |
| Maximum length of character | 30 |
| Filter size: word/character level | 1,2,3/3 |

### 3.2.3. Experiment setting

*Data preprocessing*

For each tweet, we adopted the script [191] to perform preprocessing (e.g. user name normalization, URL normalization, lowercase), and then NLTK TweetTokenizer was used for tokenization, and Taggers was utilized for POS Tagging.[192] For each user, all of the available tweets (up to 100) were combined into a single document, which served as input for the machine learning and deep learning models.

*Baseline models*

Several competitive traditional machine and deep learning models were chosen for comparison. For traditional machine learning algorithms, term frequency-inverse document frequency (TF-IDF) was used as features which were then employed to evaluate several algorithms, including support vector machine (SVM), logistic regression (LR), and extra tress (ET). For deep learning algorithms, the basic CNN model with word embedding only (termed CNN), the CNN model with word embedding and character embedding (termed CNN_char), and an attentive bi-directional recurrent neural network model (termed RNN) with word embedding, only, were evaluated.

*Cross fold validation & community ensemble*

The training dataset was split into 5 folds. For each fold, 4 folds data were used as the training set while the remaining fold was used as the validation set. The model which achieved the highest accuracy on the validation corpus was selected and evaluated on the official testing data. The mean accuracy of models from 5 folds on the testing set and the accuracy of ensemble models from 5 folds after a majority voting were calculated.

### 3.2.4. Chi-square test

Gender differences in perceptions regarding the HPV vaccine were investigated using Chi-square. Differences were measured by the odds ratio of gender versus perceptions with respect to the number of tweets mapped to the constructs of HBM and TPB. This was done for each year, from 2014-2018, using Twitter data . It should be noted that the year 2018 had Twitter data only up to October. Chi-square was used to determine whether there was a significant difference between the frequencies based on gender. To adjust for multiple comparisons, a Bonferroni correction was used with a nominal significance level of 0.05 and with 35 comparisons (7 tests each year).

## 3.3. Results and Discussion

### 3.3.1. Comparison of algorithms on gender inference

The comparison of different algorithms on gender inference can be seen *in Table 4*. The mean, standard deviation (SD) and voting accuracy of 5 folds for each algorithm on the testing dataset were calculated. CNN models outperformed SVM and RNN models on Twitter gender inference tasks. The community ensemble model further improved the accuracy for all of the algorithms. The ensemble of models from 5 folds led to better performance compared to the mean accuracy of 5 folds. The ensemble of CNN_char_pos achieved the highest accuracy among all of the models and slightly higher accuracy compared to the best results (0.8221 from [179]) reported in Task at PAN 2018.

Table 4. Comparison of algorithms on Twitter gender inference

| | SVM | RNN | CNN | CNN _char | CNN_char _pos |
|---|---|---|---|---|---|
| Mean | 0.7902 | 0.7874 | 0.8019 | 0.8127 | 0.8128 |
| SD | 0.0035 | 0.0106 | 0.0066 | 0.0018 | 0.0060 |
| Voting | 0.7968 | 0.8047 | 0.8153 | 0.8189 | 0.8237 |

### 3.3.2. Gender differences in public perceptions

The trained ensemble model of CNN_char_pos was then used to infer Twitter user gender on the HPV Twitter corpus: the majority voting of predictions from models trained on each fold was taken as the final prediction. Among 275,753 active Twitter users, 137,506 (49.87%) were inferred to be women. Out of 1,052,770 tweets in the HPV vaccine Twitter corpus, 550,224 tweets (52.26%) were inferred to be sent by women Twitter users. Coverage of the ensemble model was evaluated further by calculating the mean of prediction probability (i.e. the value of Softmax output) of 5 folds on the prediction of the HPV Twitter corpus. At the Twitter user level, 183,394 (66.51%) had a higher average probability than 0.9; 215,722 (78.23%) had a higher average probability than 0.8. At the tweet level, 653,860 (62.11%) had a higher average probability than 0.9; 790,275 (75.01%) had a higher average probability than 0.8 on their predicted gender. The gender inference model showed high coverage on the corpus.

The odds ratios of gender versus HBM and TPB measures for years 2014-2018 were plotted in *Figure 9*. Odds ratios greater than 1 meant that men had higher positive rates than women. The Chi-square for all the odds ratios were significant, with p-values less than 0.001. It was shown that for all of the constructs, men had lower prevalence rates than women, except for HBM *barriers* and TPB *positive.* Here*,* the odds ratios were above 1.7 and 1.5, respectively, for all five years. The results were largely consistent with findings from previous survey-based studies. For example, in a Korean sample,[101] men perceived more barriers to HPV prevention than did women, while women perceived more benefits and knowledge of the HPV vaccine. Additionally, in a study of African-American college students, [100] found that men scored higher on the perceived barriers to the HPV vaccine and lower on perceived severity and perceived benefits than did women.

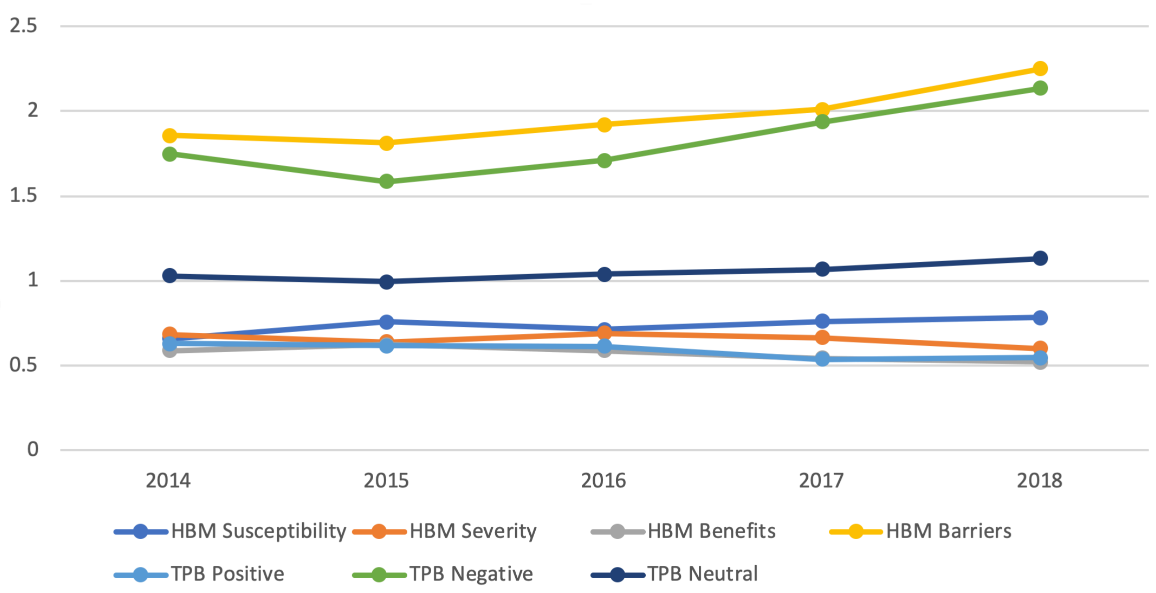


Figure 9. Difference in frequencies of tweets aligned to HBM and TPB constructs regarding HPV vaccine between gender groups (male vs female).

## 3.4. Discussion and conclusion

In the present study, a character-based CNN model with embedding fusion for Twitter gender inference was evaluated with users' historical English Twitter posts as the input. The performance of the proposed model compared favorably with the state-of-the-art performance in a recent Author Profiling Task. The comparison of embedding fusion shows the efficacy of using character embedding and POS embedding in Twitter gender inference. The trained models on an HPV vaccine-related Twitter corpus were leveraged and they identified public perception differences regarding the HPV vaccine between gender groups. The findings were largely consistent with previous survey-based studies and showed the potential of using social media and deep learning models to understand differences in public perceptions of public health related topics for different demographic groups.

# Chapter 4: Exploring Interstate Variations in Public Perceptions of the HPV Vaccination from Twitter Using Deep Learning

## 4.1. Introduction and Related Work

Interstate variations in law, policy, religion, history, socioeconomic status, and other important institutions and indicators are well-documented. Therefore, it is conceivable that such regional differences may also exist with regard to vaccine health belief and coverage. Although there is no federal vaccination law, each state has its own set of laws that permit certain exemptions for medical (all 50 states), religious (45 states), and philosophical (15 states) reasons.[193]

According to a national immunization survey, HPV vaccine coverage showed higher interstate variations compared with other vaccines. For example, in 2017, Missouri had the lowest coverage (85.8 %) for the MMR vaccine for children between the ages of 19 to 35 months, while Massachusetts had the highest (98.3%);[194] Louisiana had the lowest coverage (35.3%) for influenza vaccinations for persons six months and older, while Massachusetts had the highest (49.7%);[97] However, for the HPV vaccine, Wyoming had the lowest coverage (>=one dose, 46.9%) for persons between the ages of 13 and 17, and DC had the highest (>=one dose, 91.9%).[195] Such high interstate variations in HPV vaccine coverage suggest that it may be interesting and worthwhile to study variation as it applies to vaccination health beliefs and attitudes.

Previous studies leveraged national surveys and interviews to access health beliefs and attitudes regarding vaccination in the U.S..[65,196] However, few studies were able to access interstate variations in public perceptions. Additionally, traditional survey-based methods suffer limitations in accessing interstate variations possibly due to the cost. However, the use of social media, especially Twitter, for accessing interstate variations can be a quick, inexpensive, and doable approach. Shapiro et al leveraged Twitter data to examine intercountry variation in HPV vaccine health beliefs.[67] In their study, individuals in the UK reported a greater concern with the vaccine than did those in Canada and Australia. Dunn et al used information exposure derived from Twitter to explain differences in state-level HPV vaccine coverage. They applied topic model to classify tweets and used proportional exposure to each topic to construct multivariable models for predicting vaccine coverage. They found that measures of information exposure derived from Twitter explained differences in coverage.[92]

Although the geo-location of Twitter users is important for public health-related studies, the location information of Twitter users is not typically or directly available on the platform. Therefore, in the last few years, Twitter location inference problems have received much attention.[181,197–199] There are three types of Twitter-related locations, *home location*, *tweet location*, and *mentioned location*. *Home location* refers to the long-term residential address of Twitter users, whereas *tweet location* refers to the place where a tweet is posted, and *mentioned location* refers to the location names stated in the Twitter text. [181] *Home location* is more important for applications such as public health monitoring and public opinion polling estimation.[181] *Home location* is typically available on Twitter users profiles. However, as it is optional for Twitter users to complete their profiles, *home location* information is very sparse. One study found that only 21% of Twitter users reported their residential cities in their profiles.[200]

The focus of the present study was on the inference of Twitter users' *home location*. *Home location* can be represented in different levels of granularity, including administrative regions, geographical grids, among others. Most existing studies predicted city-level *home location* for Twitter users; few studies predicted state-level location information. Additionally, traditional efforts were based mostly on probabilistic models. For example, Cheng et al proposed methods that first identified words in tweets with a strong local geo-scope, and then modeled the spatial word usage to infer the location.[199] Some studies also explored the use of supervised machine learning algorithms with word features, including Naive Bayes, logistic regression, among others and , to infer locations of the users.[201,202] Recent efforts evaluated the use of deep learning for location estimation. Rahimi leveraged a multilayer perceptron with one hidden layer to build the Twitter location classifier.[203] Miura et al combined users' posts chronologically and proposed a recurrent neural network to predict *home location*.[204] The current study used administrative regions at the U.S. state level, where decisions actually are made about vaccine policies.

The present study applied deep learning-based approaches for Twitter users' home location inference at the U.S. state level, utilizing chronological English Twitter textual data as the input. Interstate variations in public perceptions of the HPV vaccine were investigated. Specifically,

1) The inference of Twitter users' *home location* as text classification tasks was framed and a set of deep learning and machine learning algorithms on a large open Twitter location dataset were evaluated.

2) The use of transfer learning for *home location* inference on the HPV vaccine-related Twitter dataset (described in Chapter 3) was evaluated.
3) The HPV vaccine related Twitter dataset was evaluated using Chi-square and it identified interstate variations in public perceptions of the HPV vaccine.

## 4.2. Method

### 4.2.1. Datasets

*Open Twitter location dataset (Lee dataset)*

Machine learning and deep learning models were evaluated using the dataset provided by Lee et al.[199] The original dataset contained two separate datasets for training and testing. The training set contained 101,358 Twitter users and 3,744,925 tweets from these users. All of the locations of users were self-labeled (in their profiles) at the U.S. city-level granularity. The test set contained 5,119 Twitter users and 5,126,035 tweets from the users. All of the locations of users in the test set were uploaded from their smart phones with the form of "UT: latitude, longitude." As we only considered the self-reported profiles as ground truth, we evaluated our methods on the training data only. An open-source script to map self-reported cities in the U.S was applied as the present study focused on U.S. state-level geo-tagging.[205] The script leveraged lexicon-based rules to map the U.S. city name to a U.S. state. For example, "Miami, FL" to "FL", "Texas" to "TX", "Texas, USA" to "TX", "Syracuse, NY" to "NY". Out of 101,358 users in the training set, 98,429 (97.11%) users' city names were able to be mapped to a valid U.S. state. The present study assumed the U.S. state inferred by the lexicon-based script as a ground truth label. Twitter users in the Lee training set were randomly split into train, validation, and test sets with a proportion of 7: 1: 2.

*The HPV vaccine-related Twitter location dataset (HPV dataset)*

As described in section 3.2.1., Twitter posts (up to 100) were collected for 275,753 Twitter users in our HPV vaccine-related Twitter corpus. These users all had accounts that were active during the collection period. Out of these Twitter users, 70,858 Twitter users' location profiles could be mapped to one of the U.S. states using the lexicon-based script. It was assumed that the U.S. states' labels inferred by the script from users' self-reported profiles were ground truth labels. The 70,858 users were split randomly into train, validation, and test sets with a proportion of 7: 1: 2. These sets were used to evaluate transfer learning.

### 4.2.2. The recurrent neural network for Twitter user *home location* inference

Twitter *home location* inference was framed as text classification and we proposed a bi-directional recurrent neural network with embedding fusion for such tasks. The overall architecture of the proposed framework is illustrated in *Figure 10*. Similar to the character-based convolutional neural network that was proposed in section 3.3.2., there was a character layer at the bottom, which took the character embedding of each character within a tweet token as the input and output the summary of characters for each token using convolution and max pooling. The character layer ensured mapping both the in-vocabulary words and the out-of-vocabulary words to a high dimensional vector. The word layer concatenated the output of the character layer, word embedding, and part of speech (POS) embedding to more comprehensively represent the linguistic features of each word. The output of the word layer was fed to a bi-directional recurrent neural network (RNN) layer to represent the backward and forward information of the Twitter user. Long short-term memory (LSTM) was chosen as the RNN unit. The attention layer was added on top of the RNN output to capture the salient portions and context.[206,207]

Several fully-connected layers were added on top of the attention layer. The output layer was a fully connected layer with Softmax outputs.

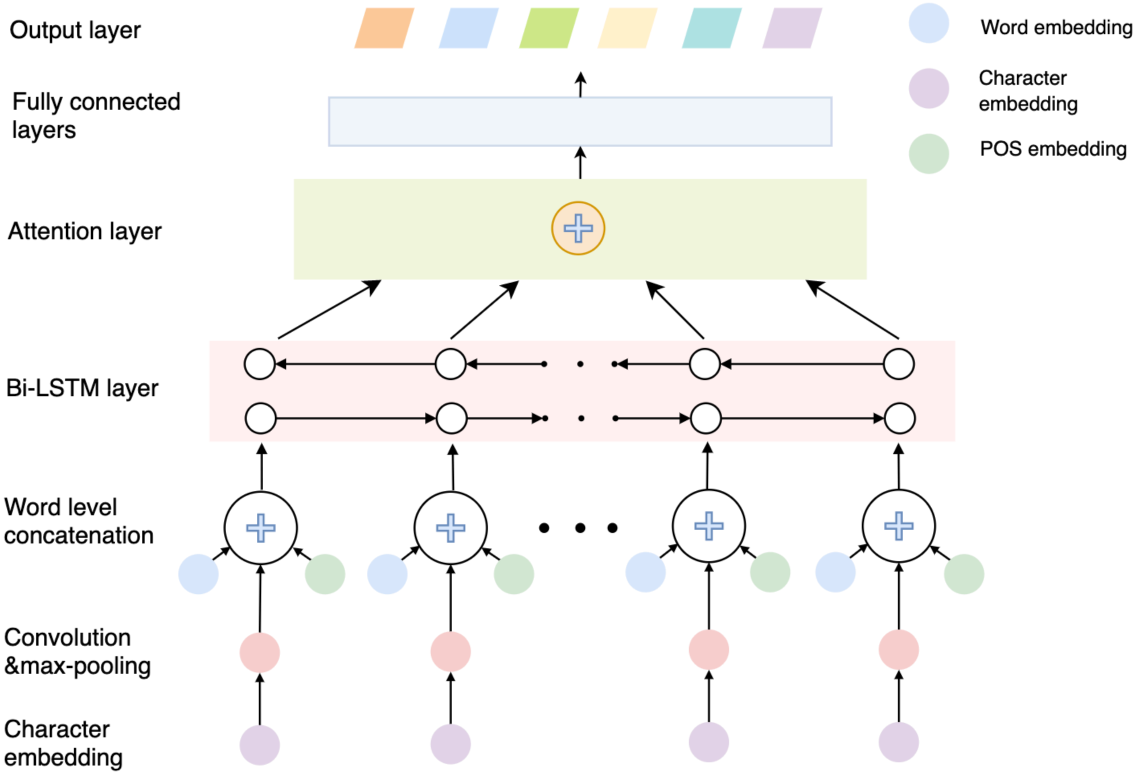


Figure 10. The architecture of character-based recurrent neural network for Twitter user *home location* inference

### 4.2.3. Experiment setting

*Data preprocessing*

The same preprocessing steps were followed as described in section 3.2.3. for Twitter text preprocessing. NLTK TweetTokenizer was leveraged for tweet tokenization and POS Tagging.[192] For each user, all of the historical tweets (up to 100) were combined into a single document, which served as the input data. The ground truth labels were the U.S. states that were extracted from users' profiles using the lexicon-based script.

### *Evaluation of machine learning and deep learning algorithms*

A set of competitive machine learning and deep learning algorithms were chosen for comparison. For machine learning algorithms, term frequency-inverse document frequency (TFIDF) was used as features and several algorithms, including logistic regression (LR), extra trees (ET) and support vector machine (SVM), were evaluated. For the proposed recurrent neural network-based framework, three settings were assessed: 1) use of word embedding, output of character layer, POS embedding together (termed RNN_char_pos); 2) use of word embedding and output of character layer together (termed RNN_char); 3) use of word embedding only (termed RNN). The major hyper-parameters settings for these RNN-based algorithms can be seen in *Table 5*. In addition, the use of convolution neural networks for *home location* tagging tasks was compared. The recurrent neural networks layer and attention layer were replaced with the convolution layer and the max-pooling layer, and the same three embedding fusion settings as RNN-based framework: CNN_char_pos. CNN_char, and CNN were evaluated. The evaluations of these various algorithms were all performed on the open *Lee dataset*.

### *Evaluation of transfer learning on the HPV dataset*

Transfer learning has been shown to be an effective technique to improve the performance of learning on a target task with limited data by using some knowledge learned from a source task.[208–210] A previous study found that the use of transfer learning improved Twitter text mining tasks.[63] Instead of training the model for a target task from a completely blank network, transfer learning can re-use all or some of the parameters trained from a source task. Using the *Lee dataset* as source tasks, transfer learning on the *HPV dataset* were evaluated in three settings: 1) the pre-trained model from the *Lee dataset* directly on the test set of the *HPV dataset*; 2) the train model from scratch on the training set of the *HPV dataset* which was then evaluated on its testing set; 3) a fine-tuned pre-trained model from the *Lee dataset* on the training set of the *HPV dataset* which was then evaluated on its test set.

Table 5. Major hyper-parameters settings for the proposed character-based recurrent neural network for Twitter user *home location* inference

| Hyper-parameter | Setting |
|---|---|
| Learning rate | 0.001 |
| Batch size | 64 |
| POS embedding dimension | 10 |
| Character embedding dimension | 50 |
| Pre-trained word embedding | GloVe Twitter embedding (d=200) |
| Number of hidden RNN/Attention units | 256 |
| Dropout rate | 0.2 |
| Filter size/No. of filters at character level | 3/512 |
| Maximum length of token | 2,000 |
| Maximum length of character | 30 |
| Dimension of hidden fully connected layers | 1,280/128/128 |

### *Evaluation metrics*

Common evaluation metrics for multi-class classification tasks were followed as the *home location* tagging tasks were framed as text classification tasks. The overall accuracy of each algorithm on the test datasets was calculated, as were precision, recall, and F-1 score for each U.S. state.

### 4.2.4. Hybrid approaches for *home location* inference

Three *home location* inference methods were adopted for the present study. The first two methods took the Twitter users' location profiles as the input and generated the U.S. state label. The third method was based partially on the deep learning approach described in section 4.2.2. This approach took Twitter users' historical posts as the input and generated the U.S. state label. There was an accuracy-coverage trade-off among these methods, which is described below.

*Lexicon-based location inference (Profile: lexicon based)*

As mentioned before, among 275,753 active Twitter users in the HPV dataset, 70,858 Twitter users' *home location* could be inferred from their profiles using the lexicon-based script. The location labels of these users were considered as ground truth labels. The lexicon-based method was considered with high accuracy but low coverage. Only 25.70% of the Twitter users' profiles in the dataset could be mapped to a U.S. state. This method was termed, *Profile: lexicon based*.

*Lexicon and Google Map-based location inference (Profile: lexicon plus Google Map)*

There are some cases where the Twitter users' profiles contained the U.S. state location information, but the information couldn't be extracted using the lexicon-based approach. For example, the lexicon-based script was not able to map location strings such as "Yellow City," "Norfolk," "Austin/Oslo," to a U.S. state. Google Map API was leveraged, further, to perform *home location* tagging on the Twitter users' profiles that couldn't be geo-tagged by the lexicon-based script. Using Google Map API, "Yellow City" was able to be identified as "2916 Wolflin Ave, Amarillo, TX 79109, USA." which was then mapped to the U.S. state "TX;" "Norfolk" was mapped to "VA," "Austin/Oslo" was able to be mapped to "TX." However, Google Map API also was able to map some location-irrelevant strings to a U.S. state, which introduced errors into the inference. For example, “heart” was mapped to “Montgomery County, Kansas;” “love” was mapped to “Owasso, Oklahoma.” In total, 31,846 Twitter users’ *home locations* were inferred using Google Map API. Together with 70,858 Twitter users that *home location* was able to infer using the lexicon-based approach, 102,704 (37.24%) Twitter users’ *home location* was inferred from Twitter users’ self-reported profiles. This hybrid method increased *home location* inference coverage but decreased accuracy. This method was termed *Profile: lexicon plus google map*.

*Lexicon and deep learning-based location inference (Profile plus Twitter posts)*

The present study first used the lexicon-based script to infer *home location* from Twitter users’ profiles. For the users whose *home location* couldn’t be inferred using the lexicon-based script, the HPV dataset fine-tuned *home location* inference deep learning model was applied to infer location information from Twitter users. This hybrid method had full coverage but the lowest accuracy among all of the methods, primarily due to the moderate accuracy in deep learning-based *home location* tagging. This was termed *Profile plus Twitter posts*.

### 4.2.5. Chi-square

Using Chi-square, pair-wise interstate variations in public perceptions of the HPV vaccine among three U.S. states, California (CA), Texas (TX) and New York (NY) were investigated. These three states were chosen due to their large populations and volume of Twitter discussions. Differences were measured by the odds ratio between U.S. states versus public perceptions with respect to numbers of tweets mapped to the constructs of HBM and TPB each year, using Twitter data from 2014-2018. Chi-square was used to determine whether there was a significant difference in the frequencies among the three U.S. states. Odds ratios were calculated among the states using three *home location* inferences. To adjust for multiple comparison, Bonferroni correction was used with a nominal significance level of 0.05 and number of comparisons 315 (105 tests for each location pair).

## 4.3. Results

### 4.3.1. Evaluation of various algorithms on the *Lee dataset*

The comparison of various machine learning and deep learning algorithms for *home location* inference on the *Lee dataset* can be seen in *Table 6*. The F-score of various algorithms for each U.S. state can be seen in *Appendix I*. As demonstrated, the deep learning models outperformed traditional machine learning models with a margin in Twitter users' *home location* inference. Recurrent neural network-based models also outperformed convolution neural network-based models on *home location* geo-tagging. Character layer was found to be beneficial in improving the accuracy for both the RNN and CNN-based framework. The use of POS embedding showed improvement for the CNN-based models but not for the RNN-based models. The best model for *home location* inference was the RNN-based model with the concatenation of word embedding and character layer output (RNN_char). At the U.S. state level, RNN_char achieved moderate performance on most of the U.S. states. However, large interstate variations also existed. For some states that had a relatively large number of tweets available, RNN_char was able to achieve F-scores higher than 0.5, for example, 0.5408 for California (CA), and 0.5244 for Texas (TX). For some states with a low tweet distribution, RNN_char achieved lower performance. In total, RNN_char achieved an F-score higher than 0.5 on nine states and higher than 0.4 on 24 states. *Figure 11* shows interstate variations of F-scores on *home location* inference from the best-performing RNN_char model.

Table 6. Accuracy of machine learning and deep learning algorithms on *home location* inference on *Lee dataset*

| | Machine learning | | | Convolution neural network | | | Recurrent neural network | | |
|---|---|---|---|---|---|---|---|---|---|
| | LR | ET | SVM | CNN | CNN_char | CNN_POS_char | RNN | RNN_char | RNN_POS_char |
| Accuracy | 0.3646 | 0.3152 | 0.3715 | 0.4158 | 0.4535 | 0.4622 | 0.4823 | **0.4896** | 0.4857 |

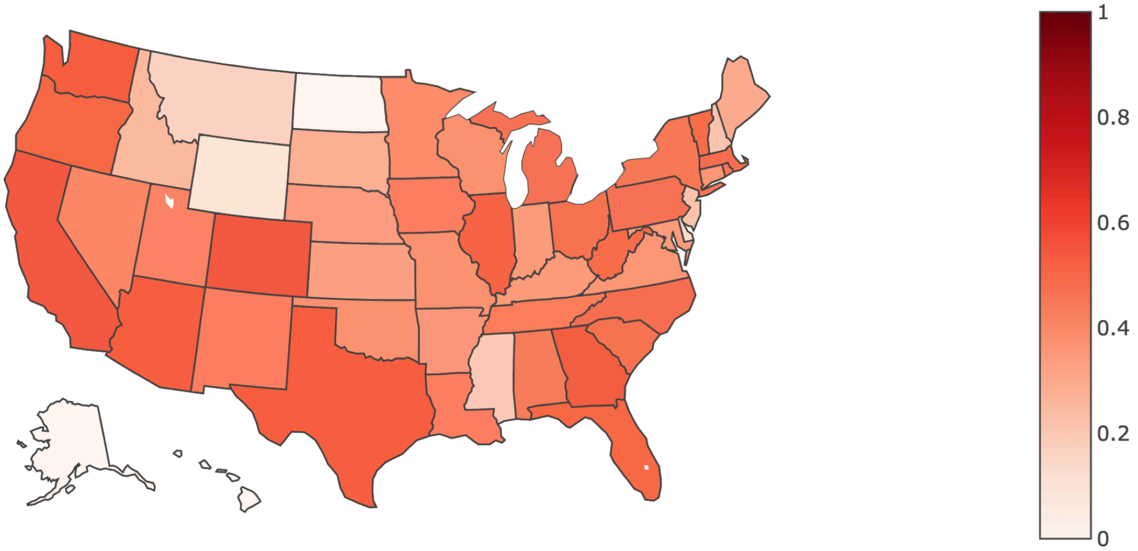


Figure 11. State-wise Twitter user *home location* inference F-score by RNN_char model on the *Lee dataset*

### 4.3.2. Evaluation of transfer learning on the *HPV dataset*

The overall accuracy of different settings for transfer learning on Twitter user *home location* inference for the HPV dataset can be seen in *Table 7*. The detailed F-score for each U.S. state can be seen in *Appendix J.* The pre-trained model from the *Lee dataset* achieved lower accuracy on the *HPV dataset*. The *Lee dataset* was compiled from September 2009 to January 2010, while the HPV dataset was compiled from December 2018 to January 2019. Changes in Twitter textual linguistic style over the past 10 years could be one of the reasons why the pre-trained model did not work well on the more recently collected dataset. The use of the *Lee dataset* for transfer learning improved the overall accuracy of the *HPV dataset*, with a 0.017 increase in overall accuracy compared with the model trained from scratch. Transfer learning was also able to improve the F-score in *home location* inference for 37 U.S. states. In total, the use of transfer learning achieved F-scores higher than 0.5 on 21 states and higher than 0.4 on 44 states. *Figure 12* shows a geographic visualization of F-score on *home location* inference from the fine-tuned model.

Table 7. Accuracy comparison on the use of transfer learning on the *HPV dataset*. Pre-trained: pre-trained model from the *Lee dataset*; training from scratch: model was trained using the *HPV dataset*; fine-tuned: model was fine-tuned on the *HPV dataset* from pre-trained model from the *Lee dataset*.

| | Pre-trained | Training from scratch | Fine-tuned |
|---|---|---|---|
| Accuracy | 0.4276 | 0.4680 | **0.4849** |

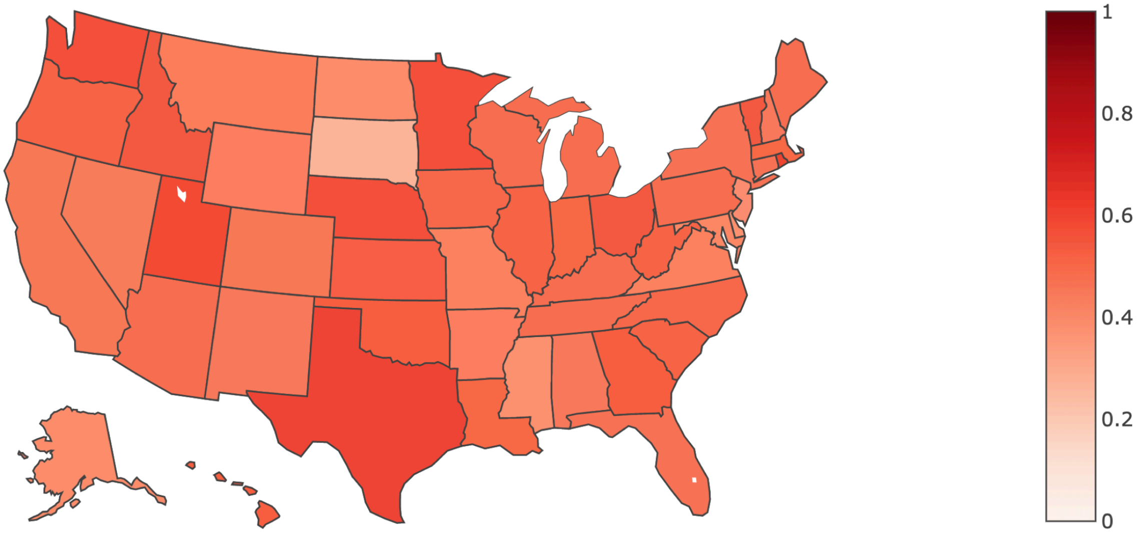


Figure 12. Twitter user *home location* inference F-score by the transfer learning (i.e. fine-tuned model) on the *HPV dataset*

### 4.3.3. Interstate differences

*California versus New York*

The odds ratios for California versus New York's Twitter discussions for the constructs of HBM and TPB for years 2014-2018 can be seen in *Table 8*. Three location inference methods largely showed consistent results. For the health belief model, Twitter discussions in California showed lower prevalence than for New York in perceived susceptibility, severity, and benefits, while showing higher prevalence in perceived barriers. For the theory of planned behavior, California showed higher prevalence in negative attitude and lower prevalence in positive attitude than New York. These results were largely consistent over the years.

Table 8. Difference in frequencies of tweets aligned to HBM and TPB constructs on the HPV vaccine between U.S. states (California versus New York). Odds ratio greater than 1 means California have higher prevalence than New York.

| **Year** | | HBM Susceptibility | HBM severity | HBM benefits | HBM barriers | TPB positive | TPB negative | TPB neutral |
|---|---|---|---|---|---|---|---|---|
| 2014 | Profile: lexicon based | 0.8430 | 0.9793 | 0.5265* | 2.1715* | 0.8288 | 2.0628* | 0.6479* |
| | Profile: lexicon plus google map | 1.0602 | 0.9271 | 0.5787* | 1.9682* | 0.8616 | 1.8578* | 0.6523* |
| | Profile plus Twitter posts | 0.9155 | 0.9280 | 0.7031* | 1.4177* | 0.8344* | 1.2546* | 0.9843 |
| 2015 | Profile: lexicon based | 0.4329* | 0.5613* | 0.5430* | 2.3750* | 0.5615* | 2.2560* | 0.7893* |
| | Profile: lexicon plus google map | 0.4142* | 0.5895* | 0.5539* | 2.2830* | 0.5866* | 2.1744* | 0.7706* |
| | Profile plus Twitter posts | 0.6144* | 0.7035* | 0.6562* | 1.5741* | 0.6395* | 1.5251* | 1.0289 |
| 2016 | Profile: lexicon based | 0.7547* | 0.6305* | 0.5967* | 2.5556* | 0.5425* | 2.4386* | 0.8093* |
| | Profile: lexicon plus google map | 0.8054* | 0.7348* | 0.6280* | 2.0462* | 0.6002* | 1.9757* | 0.8601* |
| | Profile plus Twitter posts | 0.7803* | 0.8154* | 0.7869* | 1.4828* | 0.7557* | 1.4807* | 0.8849* |
| 2017 | Profile: lexicon based | 0.7342* | 0.7059* | 0.8838 | 1.6518* | 0.7234* | 1.5941* | 0.8894 |
| | Profile: lexicon plus google map | 0.8831 | 0.7922* | 0.8749 | 1.4429* | 0.7819* | 1.4054* | 0.9018 |
| | Profile plus Twitter posts | 0.7942* | 0.7961* | 0.8142* | 1.4670* | 0.7566* | 1.4416* | 0.8985* |
| 2018 | Profile: lexicon based | 0.8617 | 0.7348* | 0.9546 | 1.6746* | 0.7603* | 1.4940* | 0.9136 |
| | Profile: lexicon plus google map | 0.8844 | 0.8360* | 0.9642 | 1.3728* | 0.8533 | 1.3269* | 0.8787 |
| | Profile plus Twitter posts | 0.7934* | 0.7324* | 0.7308* | 1.7073* | 0.7031* | 1.5995* | 0.8954* |

** $p < 0.0001$.*

*New York versus Texas*

The odds ratios for New York versus Texas's Twitter discussions for the constructs of HBM and TPB for the years of 2014-2018 can be seen in *Table 9*. New York showed lower prevalence in perceived barriers and negative attitude than Texas from 2014 to 2017, and higher prevalence in perceived benefits and positive attitude in 2014. The prevalence of most constructs for New York and Texas has been moving closer (odds ratio is close to 1). The odds ratios from three location inference models also were consistent for most constructs in most years. However, some variation can be observed, for example in the HBM perceived susceptibility for 2014.

Table 9. Difference in frequencies of tweets aligned to HBM and TPB constructs on the HPV vaccine between U.S. states (New York vs Texas). Odds ratio greater than 1 means New York has higher prevalence than Texas.

| **Year** | | HBM Susceptibility | HBM severity | HBM benefits | HBM barriers | TPB positive | TPB negative | TPB neutral |
|---|---|---|---|---|---|---|---|---|
| 2014 | Profile: lexicon based | 2.2099 | 0.8893 | 1.9904* | 0.4454* | 1.3159 | 0.4699* | 1.4174* |
| | Profile: lexicon plus google map | 1.1795 | 0.9278 | 1.7386* | 0.5037* | 1.2411 | 0.5453* | 1.3705* |
| | Profile plus Twitter posts | 1.2588 | 1.1642 | 1.7561* | 0.5520* | 1.3060* | 0.5967* | 1.2980* |
| 2015 | Profile: lexicon based | 1.1706 | 0.8904 | 1.0364 | 0.8744 | 0.9943 | 0.7988 | 1.2734 |
| | Profile: lexicon plus google map | 1.2645 | 0.8955 | 1.0365 | 0.8763 | 1.0065 | 0.7698* | 1.3103* |
| | Profile plus Twitter posts | 1.3226 | 1.1788 | 1.4693* | 0.6263* | 1.3684* | 0.5632* | 1.3897* |
| 2016 | Profile: lexicon based | 1.0405 | 1.0625 | 1.0248 | 0.7678* | 1.0895 | 0.7566* | 1.1721 |
| | Profile: lexicon plus google map | 1.0165 | 0.9403 | 1.0347 | 0.8410* | 1.0933 | 0.8091* | 1.1182 |
| | Profile plus Twitter posts | 0.9554 | 1.0268 | 0.9993 | 0.8526* | 1.0412 | 0.8030* | 1.2227* |
| 2017 | Profile: lexicon based | 1.2282 | 1.1016 | 0.9095 | 0.8202 | 1.1537 | 0.7912* | 1.0696 |
| | Profile: lexicon plus google map | 1.1322 | 1.0291 | 0.9474 | 0.8728 | 1.1119 | 0.8777 | 1.0191 |
| | Profile plus Twitter posts | 0.9931 | 1.1383* | 1.1617* | 0.7942* | 1.2214* | 0.7018* | 1.1959* |
| 2018 | Profile: lexicon based | 1.0968 | 1.0753 | 0.8531 | 0.9151 | 1.0016 | 0.9950 | 1.0031 |
| | Profile: lexicon plus google map | 1.1623 | 0.9567 | 0.8350* | 1.0603 | 0.9141 | 1.1143 | 1.0015 |
| | Profile plus Twitter posts | 1.0582 | 1.0939 | 1.0777 | 0.8134* | 1.0453 | 0.7377* | 1.3383* |

$^{*}p < 0.0001$.

*California versus Texas*

The odds ratios of California versus Texas's discussions regarding the constructs of HBM and TPB for the years, 2014-2018, can be seen in *Table 10*. Similar to the previous two comparisons, the odds ratios from three location inference methods were largely consistent on most of the constructs for most of the years. In the Twitter space, except for year 2014, California showed higher prevalence in discussions of perceived barriers and negative attitude, and lower prevalence in the discussion of perceived benefits and positive attitude, than Texas. California also showed relatively lower prevalence discussions of perceived susceptibility and severity in since 2015.

Table 10. Difference in frequencies of tweets aligned to HBM and TPB constructs on HPV vaccine between U.S. states (California versus Texas). Odds ratio greater than 1 means California have higher prevalence than Texas.

| **Year** | | HBM Susceptibility | HBM severity | HBM benefits | HBM barriers | TPB positive | TPB negative | TPB neutral |
|---|---|---|---|---|---|---|---|---|
| 2014 | Profile: lexicon based | 1.8630 | 0.8709 | 1.0480 | 0.9672 | 1.0906 | 0.9693 | 0.9183 |
| | Profile: lexicon plus google map | 1.2506 | 0.8601 | 1.0061 | 0.9914 | 1.0693 | 1.0130 | 0.8939 |
| | Profile plus Twitter posts | 1.1524 | 1.0803 | 1.2347* | 0.7826* | 1.0897 | 0.7487* | 1.2776* |
| 2015 | Profile: lexicon based | 0.5068* | 0.4998* | 0.5627* | 2.0766* | 0.5583* | 1.8021* | 1.0050 |
| | Profile: lexicon plus google map | 0.5238* | 0.5279* | 0.5741* | 2.0006* | 0.5904* | 1.6738* | 1.0097 |
| | Profile plus Twitter posts | 0.8127 | 0.8293 | 0.9642 | 0.9858 | 0.8751* | 0.8590* | 1.4299* |
| 2016 | Profile: lexicon based | 0.7852 | 0.6699* | 0.6114* | 1.9621* | 0.5910* | 1.8451* | 0.9486 |
| | Profile: lexicon plus google map | 0.8187 | 0.6909* | 0.6497* | 1.7209* | 0.6562* | 1.5986* | 0.9618 |
| | Profile plus Twitter posts | 0.7455* | 0.8372* | 0.7864* | 1.2643* | 0.7868* | 1.1889* | 1.0820 |
| 2017 | Profile: lexicon based | 0.9017 | 0.7776* | 0.8038* | 1.3548* | 0.8346* | 1.2613* | 0.9514 |
| | Profile: lexicon plus google map | 0.9998 | 0.8153* | 0.8288* | 1.2594* | 0.8694 | 1.2335* | 0.9191 |
| | Profile plus Twitter posts | 0.7887* | 0.9062* | 0.9459 | 1.1650* | 0.9241 | 1.0118 | 1.0746 |
| 2018 | Profile: lexicon based | 0.9451 | 0.7902* | 0.8144* | 1.5324* | 0.7615* | 1.4866* | 0.9165 |
| | Profile: lexicon plus google map | 1.0279 | 0.7997* | 0.8051* | 1.4555* | 0.7800* | 1.4786* | 0.8800 |
| | Profile plus Twitter posts | 0.8395* | 0.8011* | 0.7875* | 1.3888* | 0.7350* | 1.1799* | 1.1983* |

$* p < 0.0001$.

## 4.4. Discussion and Conclusion

In this study, interstate variations in public perceptions regarding HPV vaccines were explored using Twitter space. As information on *home location* is sparse on Twitter users' profiles, the inference of Twitter users' *home location* was framed as text classification tasks and a character-based recurrent neural network model with attention mechanism for the *home location* inference was proposed. An evaluation was performed on a set of competitive machine learning and deep learning algorithms using a publicly available Twitter users location dataset. The proposed model outperformed convolutional neural network models and conventional machine learning baseline models on *home location* inference. The study also found that the use of transfer learning could further improve the inference accuracy of *home location* on the HPV Twitter users' location dataset.

Three hybrid inference methods were adopted to infer the Twitter users' *home location* in the HPV Twitter dataset. However, these methods suffered from the trade-off between accuracy and coverage. Profile-based approaches have higher accuracy but low coverage, while deep learning and historical posts-based approach have decreased accuracy but higher coverage. Chi-square tests were then applied to explore interstate variations in public perceptions of HPV vaccines on U.S. state labels inferred by these three methods. Three U.S. states, California, Texas, and New York were selected for testing. Odds ratios from three inference methods showed largely consistent results in most of the constructs and for most of the years. The difference in public perceptions of health belief and attitude towards HPV vaccination were identified. Notably, this appears to be the first effort to study interstate differences in health beliefs and attitudes regarding HPV vaccines using Twitter space.

There are several limitations to the present study. First, inferred location was treated as a true *home location* in the Chi-square tests, which could have led to information bias due to misclassification rates.[167] Second, it was not possible to filter out non-U.S. Twitter discussions for the deep learning-based approach as country information was not available in the collected Twitter corpus. As a result, Twitter users outside the U.S. could have been misclassified into a U.S. state. In addition, users' historical tweets, only, were used as the input for the proposed deep learning algorithms. Other information, such as Twitter users' network, could also have been used to infer Twitter users' *home location*. Learning from complementary information in multiple modalities was intuitively appealing for improving the performance.[211] Finally, Twitter users are not representative of the general population, which also led to bias in the analysis.

As the proposed deep learning model suffered from suboptimal performance (0.48 ~ 0.49 overall accuracy) in inferring the *home location* of Twitter users, multimodal deep learning should be explored and multiple pieces of Twitter information, including Twitter content, context, and users' networks, should be used to further improve the accuracy of Twitter users' *home location*. More Twitter users' location-related data should be collected, especially Twitter users' data from U.S. states with low Twitter discussion distributions. A larger and more balanced dataset could further improve the location inference accuracy of deep learning models.

# Chapter 5: Develop A Web-Based Visualization System to Monitor Real-Time Twitter Discussions of the HPV Vaccine

## 5.1. Introduction and Related Work

Data visualization is the graphic representation of information, which produces figures that communicate relationships among the represented data to the viewers. Data visualization is able to create meaningful and actional information from data that then can be used by decision makers. Interactive visualization is powered by programming and can further allow users access and opportunities to interact with information. The use of interactive visualization is becoming increasingly popular, especially for real-time data (e.g. social media). Users are able to explore and interact with real-time data and access information that is customized for their needs.

Data visualization has been adopted widely by public health agencies to track outbreaks of infectious diseases and to estimate the coverage of vaccinations. For example, the CDC maintains *FluView* to monitor seasonal influenza activity. *FluView* provides geographical visualizations of statewide influenza-like illnesses.[212] The CDC also developed *FluVaxView*[213] and *TeenVaxView*[98] to allow users to explore statewide vaccine coverage. *FluVaxView* and *TeenVaxView* provide interactive dashboards that permit users to select a specific population and location in order to see trends in vaccination coverage. Vaccination coverage is estimated using data from several national and local representative surveys, including the National Immunization Survey.[103]

Besides the visualizations and dashboards that use data collected by conventional methods (e.g. surveys or clinical data), there also exist some pioneering studies that use internet data, such as those from social media, for public health surveillance. *Vaccine Sentimeter* provides real-time surveillance and trend analyses of vaccination conversations from both mainstream and social media. *Vaccine Sentimeter* leverages automate tools to perform sentiment analysis on vaccine-related discussions.[214] *VaccineWatch* is a monitoring system with visualizations and analytics of vaccine information using Twitter and RSS feeds.[104] *VaccineWatch* allows users to filter social media discussions by specific entities, such as the names of particular diseases and vaccines. An interactive dashboard is provided to visualize the number of relevant social media discussions over time. Similarly, *InSTEDD's Riff* was developed to visualize multiple streams of information (including Twitter data) in order to monitor discussion topics during epidemics and disasters.[215]

Although these aforementioned studies leverage visualizations to offer useful information related to infectious diseases or vaccines by using social media and internet data, the actionable insights generated by these tools are limited by their analysis methods (e.g. number of posts, sentiment of posts). As was discussed before, an understanding of social media discussions from the perspective of behavior change theories can provide more actionable insights for health professionals and policy makers concerned with vaccine promotion. In this study, *VaxInsight*, a web-based dashboard, was built and evaluated. It offers interactive visualizations of public perceptions incorporating the perspectives of multiple behavior change theories on vaccines by using Twitter data. *VaxInsight* allows users to explore the prevalence of a specific construct from the perspective of multiple behavior change theories and compare differences among subpopulations by gender or by location. As part of the present study, a group of faculty and trainees with backgrounds in health promotion were invited to evaluate the usability of *VaxInsight*.

## 5.2. System Design

### 5.2.1. System architecture

The overall architecture of *VaxInsight* is depicted in *Figure 13*. *VaxInsight* comprises several modules that provide stepwise processing and analysis of Twitter data. The data acquisition module collects tweets and users' timelines via Twitter streaming, API, and Tweepy. The data preprocessing module performs preprocessing tasks such as tweet cleaning and normalization. Deep learning-based theoretical constructs mapping modules map Twitter discussions to the primary constructs of the health belief model and the theory of planned behavior (see Chapter 2). The demographic attributes inference module infers two demographic labels, gender (see Chapter 3) and the U.S. state-level location (see Chapter 4). For the current study, the tweets, predicted labels of the tweets, and inferred demographics of Twitter users were stored in MongoDB, which was connected to a web-based dashboard to provide interactive visualizations of public perceptions regarding the HPV vaccine, again, using Twitter data.

The dashboard of *VaxInsight* was built using Python and Plot.ly Dash.[216] Dash is a Python framework built on top of Plotly.js, React, and Flask, that is used for constructing web applications, especially for data visualization. Dash is able to create interactive dashboards that allow multi-view brushing and filtering.

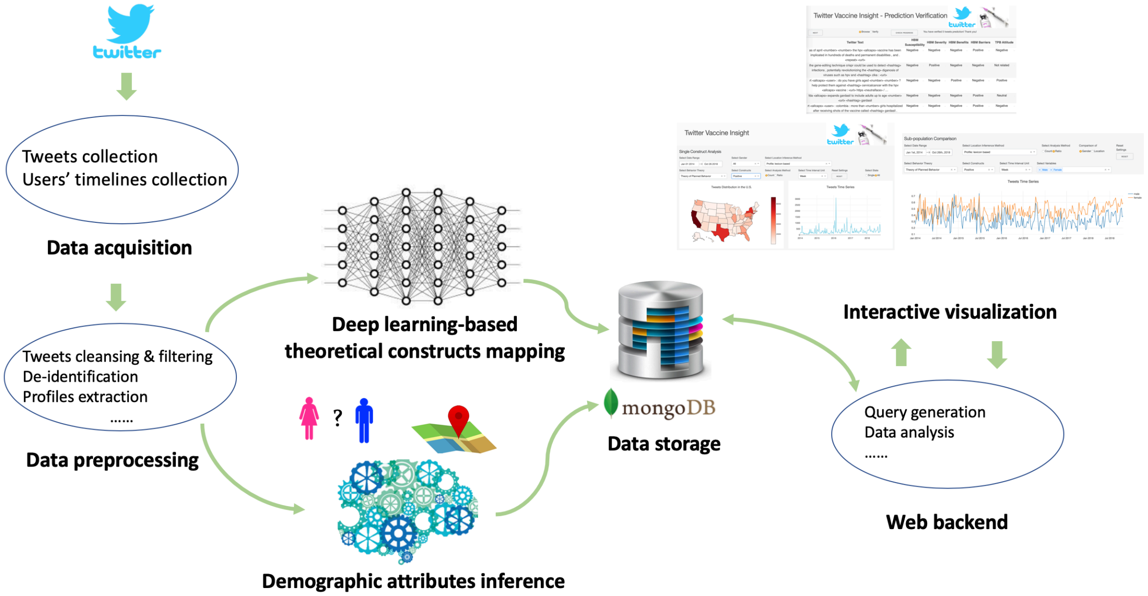


Figure 13. The architecture of *VaxInsight*

### 5.2.2. Functions description

There were three major goals in designing the *VaxInsight* prototype : 1) to explore the trend of a specific construct from a particular behavior change theory; 2) to compare differences among subpopulations in terms of trends related to a specific construct; 3) to browse and verify the predictions of deep learning-based theoretical construct mapping.

*Single construct analysis*

*Single construct analysis* allows users to filter Twitter discussions by time range and demographics, and then to explore trends for a particular construct. The screenshot of the *single construct analysis* page can be seen in *Figure 14*. The descriptions of the options provided on that page are below:

<u>Select Date Range</u>: Select date range of Twitter discussion. Specific start/end dates can be selected using a dropdown list, or a date string can be typed into a date string the text box with a the format such as "Oct 26 2018."

Select Gender: Filter Twitter discussion by gender. The gender label is inferred by deep learning algorithms employing historical user tweets as input. The process was described in Chapter 3.

Select Location Inference Method: Select different method to infer Twitter users' U.S. state-level location information. There are three options available: 1) Profile: lexicon-based, where the location is extracted from users' self-reported location in their profiles using a lexicon-based approach; 2) Profile: lexicon Based + Google Map, where the location is inferred by the lexicon-based approach and Google Map API from users' self-reported profiles. The present study also leveraged Google Map API to extract location information from users whose locations could not be inferred by the lexicon-based approach; 3) Profile plus Twitter posts, where the location is inferred by the lexicon-based approach and deep learning. The present study also leveraged deep learning to infer location information by utilizing users' timelines as the input when user location could not be inferred by the lexicon-based approach in their profiles. Detailed descriptions of these methods are provided in Chapter 4.

Select Behavior Theory: Select the behavior change theory of interest. Currently, *VaxInsight* supports two behavior theories: the health belief model, and the theory of planned behavior. The default option is "All", which returns all of the HPV vaccine-related tweets collected by keywords.

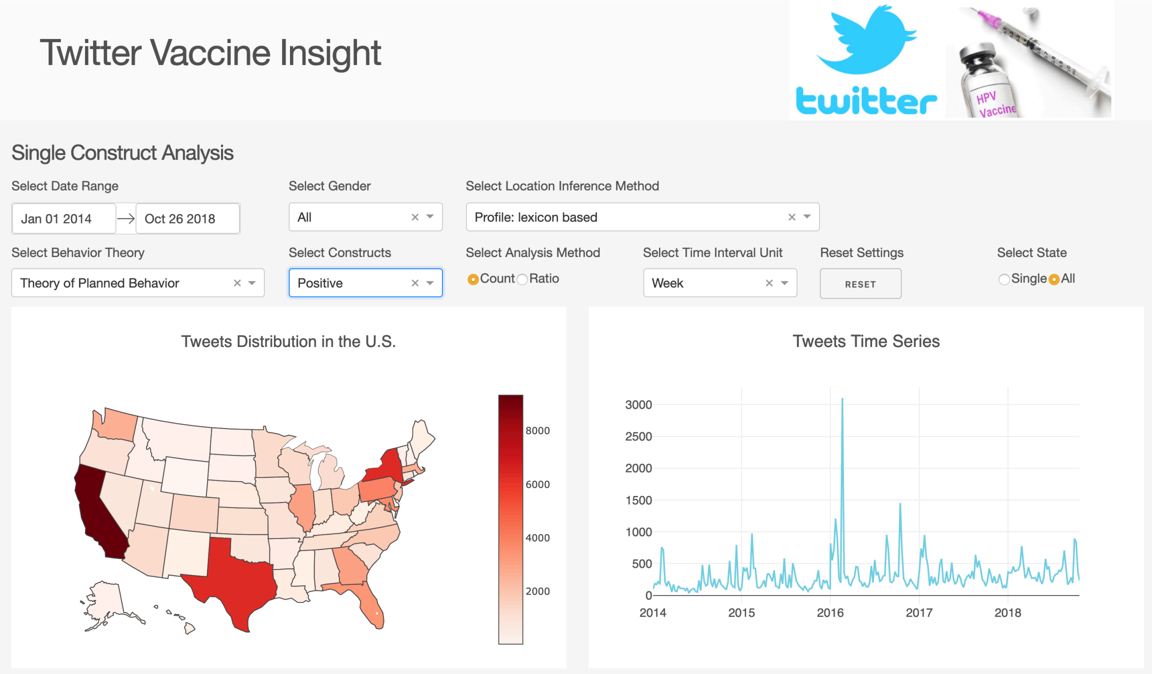
Select Constructs: Select the construct for a particular behavior change theory. Currently, for the health belief model, *VaxInsight* supports four primary constructs, *perceived susceptibility*, *perceived severity*, *perceived benefits*, and *perceived barriers*. For the theory of planned behavior, *VaxInsight* focuses on attitudes, *positive, negative, and neutral*. The default option is "All", which will return all the theory-related tweets.

Select Analysis Method: *VaxInsight* currently supports two analysis options: 1) Count, which calculates the aggregate count of relevant tweets during each time interval unit. 2) Ratio (only available when a particular construct is selected), which calculates the ratio of the aggregate count of construct-related tweets to the aggregate count of theory-related tweets during each time interval unit.

Select Time Interval Unit: Select the unit of interval that is used to calculate the aggregate data. VaxInsight provides options for "Day," "Week," "Month," and "Year."

Reset Settings: Reset all of the settings to default.

Select State: Select a particular state for the analysis. First, it is necessary to click "Single" then click one particular state from the U.S. map. By default, the option is "All," which returns all of the data.


Twitter Vaccine Insight

Single Construct Analysis

Select Date Range: Jan 01 2014 → Oct 26 2018

Select Gender: All

Select Location Inference Method: Profile: lexicon based

Select Behavior Theory: Theory of Planned Behavior

Select Constructs: Positive

Select Analysis Method: ● Count ○ Ratio

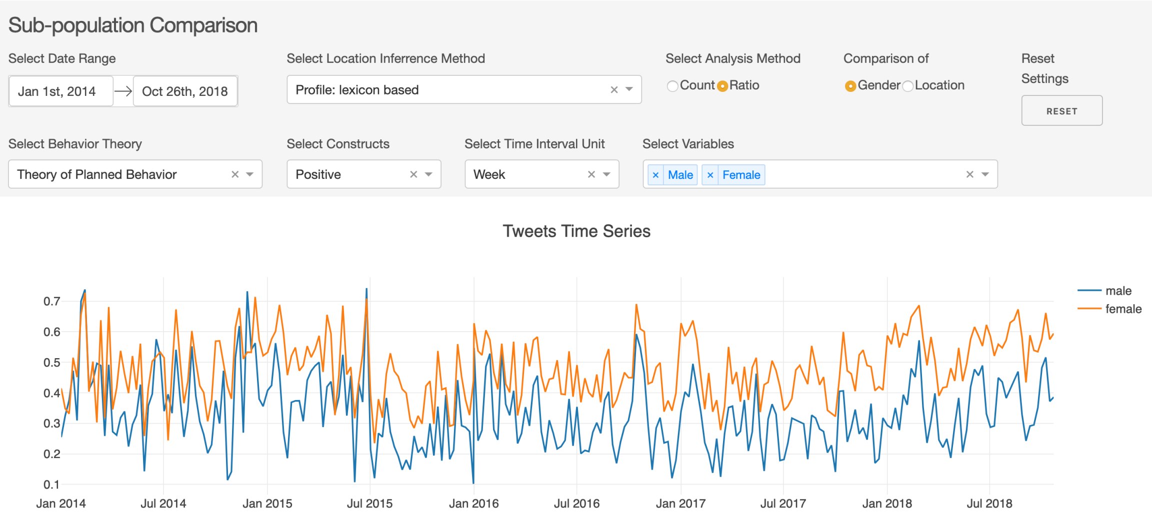
Select Time Interval Unit: Week

Select State: ○ Single ● All

Figure 14. Screenshot of the single construct analysis page in *VaxInsight*

### *Subpopulation Comparison*

*Subpopulation comparison* allows users to compare gender and location subgroup differences in trends related to theoretical constructs. A screenshot of the *subpopulation comparison* page can be seen in *Figure 15*. This page provides functions that are similar to S*ingle Construct Analysis*, such as filtering the data by date range and selecting a particular construct. This page also allows users to select subpopulations for comparison using a combination of the following two options:


Sub-population Comparison

Select Date Range: Jan 1st, 2014 → Oct 26th, 2018

Select Location Inference Method: Profile: lexicon based

Select Analysis Method: ○ Count ● Ratio

Comparison of: ● Gender ○ Location

Select Behavior Theory: Theory of Planned Behavior

Select Constructs: Positive

Select Time Interval Unit: Week

Select Variables: × Male × Female

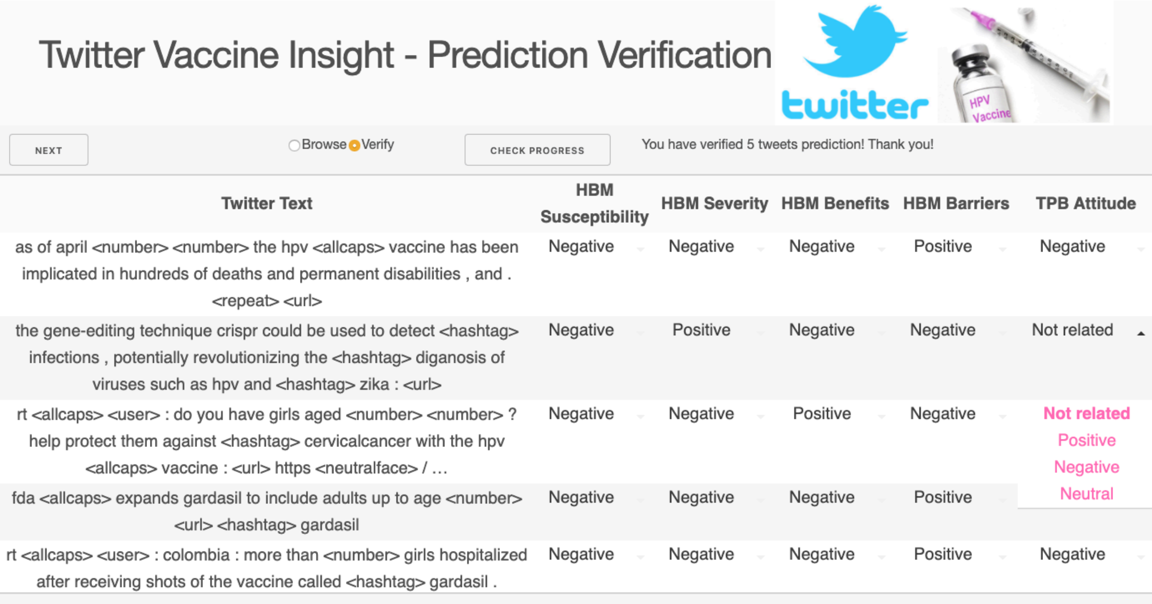
Figure 15. Screenshot of the subpopulation comparison page in *VaxInsight*

Comparison of: Select the type of subpopulation that is to be compared. Currently, *VaxInsight* provides two options: "Gender" and "Location."

Select Variables: Type the subpopulation that is to be compared. For gender, *VaxInsight* has "male" and "female' as the options; for location, *VaxInsight* includes all of the 50 U.S. states.

### *Prediction Verification*

Deep learning-based systems can make mistakes in mapping Twitter discussions to theoretical constructs. Thus, a web page was developed where users can browse and verify the predictions of the system. Users are able to make changes to the predicted labels and the verified changes are stored in the database. The verified labels can be used to fine-tune the deep learning models in order to improve their accuracy. A screenshot of the *prediction verification* page can be seen in *Figure 16*.



You have verified 5 tweets prediction! Thank you!

| Twitter Text | HBM Susceptibility | HBM Severity | HBM Benefits | HBM Barriers | TPB Attitude |
|---|---|---|---|---|---|
| as of april <number> <number> the hpv <allcaps> vaccine has been implicated in hundreds of deaths and permanent disabilities , and . <repeat> <url> | Negative | Negative | Negative | Positive | Negative |
| the gene-editing technique crispr could be used to detect <hashtag> infections , potentially revolutionizing the <hashtag> diganosis of viruses such as hpv and <hashtag> zika : <url> | Negative | Positive | Negative | Negative | Not related |
| rt <allcaps> <user> : do you have girls aged <number> <number> ? help protect them against <hashtag> cervicalcancer with the hpv <allcaps> vaccine : <url> https <neutralface> / … | Negative | Negative | Positive | Negative | Not related |
| fda <allcaps> expands gardasil to include adults up to age <number> <url> <hashtag> gardasil | Negative | Negative | Negative | Positive | |
| rt <allcaps> <user> : colombia : more than <number> girls hospitalized after receiving shots of the vaccine called <hashtag> gardasil . | Negative | Negative | Negative | Positive | Negative |

Figure 16. Screenshot of the prediction verification page in *VaxInsight*

## 5.3. System Usability Evaluation

To evaluate the usability of *VaxInsight*, a System Usability Scale (SUS)[217] was adopted. A small group of researchers with backgrounds in health promotion were recruited from the University of Texas School of Public Health. They then performed the usability test of *VaxInsight*. This study received exempt status from the Committee for the Protection of Human Subjects at The University of Texas Health Science Center at Houston. The approved IRB protocol number is HSC-SBMI-19-0492. The participants were asked to perform four tasks on the website and then fill out the evaluation survey.

### 5.3.1. Tasks

Participants were asked to perform four tasks related to the two major functions of *VaxInsight*. The functions were 1) to explore the trend for a particular theoretical construct; 2) to compare subpopulation differences on the trend for a particular theoretical construct. A description of the tasks are as follows:

Task 1: Select a date range (click dropdown list or type date in the text box), then select a specific construct from one theory of interest (e.g. *positive* in the theory of planned behavior). Identify and describe the peaks of counts or ratios in the trend, using the “Toggle Spike Lines” function in the Tweets Times Series plot.

Task 2: Select a date range (click dropdown list or type date in the text box), then select a specific construct from one theory of interest (e.g. *positive* in the theory of planned behavior). Filter the Twitter data by gender and/or location (select “Single” in “Select State” and then click a state on the U.S. map). Identify and describe the peaks of counts or ratios in the trend, using the “Toggle Spike Lines” function in the Tweets Times Series plot.

Task 3: Select a date range (click dropdown list or type date in the text box), then select a specific construct from one theory of interest (e.g. *positive* in the theory of planned behavior). Select either “Count” or “Ratio” in “Select Analysis Method.” Select “Gender” in “Comparison of,” then type gender labels (e.g. “male”, “female”) in the “Select Variables” text box. Describe the differences among these subpopulations (e.g. which group has relative higher prevalence during what time period).

Task 4: Select a date range (click dropdown list or type date in the text box), then select a specific construct from one theory of interest (e.g. *positive* in the theory of planned behavior). Select either “Count” or “Ratio” in “Select Analysis Method.” Select “Location” in “Comparison of,” then type location labels (e.g. “california”, “texas”, “new york”) in the “Select Variables” text box. Describe the differences among these subpopulation groups (e.g. which group has relative higher prevalence during that time period).

### 5.3.2. Evaluation questions

After the participants performed these tasks, each was asked to complete a survey that contained two questions. One question was related to their educational background (i.e. degree) and the other asked for recommendations to improve the system. The participants were also asked to evaluate 10 statements related to usability. The usability-related statements were as follows:

1. I think that I would like to use this system frequently
2. I found the system unnecessarily complex
3. I thought the system was easy to use
4. I think that I would need the support of a technical person to be able to use this system
5. I found the various functions in this system were well integrated
6. I thought there was too much inconsistency in this system
7. I would imagine that most people would learn to use this system very quickly
8. I found the system very cumbersome to use
9. I felt very confident using the system
10. I needed to learn a lot of things before I could get going with this system

For these statements, five-point scales were provided which ranged from “strongly disagree” to “strongly agree.”

### 5.3.3. Evaluation Results

The evaluation results showed good usability of *VaxInsight*. A majority of the participants (four out of six) agreed or strongly agreed that they would like to use this system frequently. All participants agreed or strongly agreed that the system was easy to use and was not unnecessarily complex. All participants disagreed or strongly disagreed that they would need the support of a technical person to be able to use this system. All participants agreed or strongly agreed that the various functions in the system were well integrated and there wasn't too much inconsistency in it. All participants agreed or strongly agreed that most people could learn to use this system very quickly and strongly disagreed that this system was very cumbersome to use. Five out six participants agreed or strongly agreed that they felt very confident using the system and disagreed or strongly disagreed that they needed to learn a lot of things before they got going with this system. The full responses can be seen in *Appendix H*.

There were also some comments on improving the interactive dashboard of *VaxInsight*. One participant asked for more comparison variables, such as race/ethnicity, education, and city-level location information. Another participant suggested adding more statistical analysis and providing printable summary reports. And another recommendation was to add some tweet examples when data were filtered. Two participants also mentioned some issues with web page format and color.

## 5.4. Discussion and Conclusion

A prototype of *VaxInsight*, a web-based interactive dashboard to synthesize HPV vaccine-related Twitter discussions in a comprehendible format was built. This prototype was based on results from prior research on mapping Twitter discussions to the theoretical constructs of behavior change theories and Twitter demographic attributes inference. The present study demonstrated that *VaxInsight* allowed users not only to filter Twitter discussions and track the trends for particular theoretical constructs, but also to compare differences in public perceptions among subpopulations using visualizations. In addition, *VaxInsight* provided a web-based interface for users to browse and validate predictions of deep learning systems. To evaluate the usability of *VaxInsight*, a group of researchers with backgrounds in health promotion were recruited to perform a usability test on the system. The survey results showed high usability of the system.

*VaxInsight* offers interactive visualizations of Twitter discussions related to HPV vaccines with high usability. However, as a prototype, *VaxInsight* also suffers from some limitations. For example, *VaxInsight* provides limited analysis options (e.g. counts, ratio) of Twitter discussions. Future work will aim to add more statistical analyses (e.g. Chi-square) in order to offer more insights. Some additional functions could also be added. For example, sample tweets related to a particular construct in a certain time frame can be shown. Additionally, the inclusion of printable reports that summarize actionable insights from the visualizations would enhance the utility of the system. Currently, *VaxInsight* continues to be tested, internally, with hopes of releasing it in the near future.

# Chapter 6: Summary

## 6.1. Summary of Key Findings

The present research explored the use of machine learning and deep learning algorithms to understand vaccine-related Twitter discussions. The research was framed using the perspective of behavior change theories and pioneers the use of artificial intelligence algorithms for public health surveillance. The key findings for each chapter are summarized as follows:

Chapter 1 demonstrated the rising trend of vaccine refusal and hesitancy and discussed its consequences for public health. It was then argued that a feasible and accurate method to access opinions from a large sample was needed as a first and critical step in reversing this trend and in promoting vaccination. A literature review of existing studies that used social media data for public health surveillance, with an emphasis on vaccine-related topics, was performed. The rise of social media and advances in machine and deep learning algorithms has opened a new door to accessing public perceptions that sidestep many of the limitations of traditional survey-based methods. The significance of and challenges to the use of social media were discussed and three aims were identified to address these challenges. These aims were: 1) Develop a deep learning-based system to understand public perceptions of HPV vaccination from the perspective of grounded behavior changes theories and by using Twitter data; 2) Develop novel computational algorithms to infer users' demographic attributes and investigate demographic differences in public perceptions regarding the HPV vaccine using Twitter data; 3) Develop a web-based interactive visualization system to monitor real-time Twitter discussions of the HPV vaccine.

Chapter 2 described efforts to evaluate various machine-learning and deep-learning algorithms in mapping HPV vaccine-related Twitter discussions to the constructs of multiple grounded behavior change theories, specifically the health belief model and the theory of planned behavior. Deep-learning algorithms outperformed machine-learning algorithms on all of the tasks. The bi-directional long short-term memory (LSTM) with attention mechanism achieved overall best performance in most of the tasks. Domain-specific embedding led to comparable or higher performance compared with pre-trained embedding from the general domain. Locally estimated scatterplot smoothing (LOESS) revealed changes in the constructs. For example, in the last two years covered by the study, *positive* attitude toward the HPV vaccine showed an increasing trend, while *negative* attitude showed a decreasing trend.

Chapter 3 described efforts to infer gender on Twitter using deep learning. A character-based CNN model using historical English Twitter posts as the input for the Twitter gender inference task was proposed. Gender differences in public perceptions of the HPV vaccine were calculated using Chi-square and such differences were then reported. The performance of the proposed model compared favorably to the state-of-the-art performance on a recent Public Author Profiling challenge. The fusion of character-based CNN, word embedding, and POS embedding led to the best performance in the gender inference task. Furthermore, Twitter gender inference models on the HPV vaccine-related Twitter corpus (described in the chapter) were leveraged and identified a gender-based difference in public perceptions of the HPV vaccine. Men were found to perceive higher *barriers* and *negative* attitude and lower *benefits* and *positive* attitude toward the HPV vaccine. These findings are largely consistent with previous survey-based studies. Thus, the present study demonstrates the potential of using social media and deep learning models to understand differences in the public perceptions of public health-related topics for different population subgroups.

Chapter 4 describes efforts to infer Twitter users' *home location* (at the U.S. state- level) using deep learning. Interstate variations in public perceptions regarding the HPV vaccine were explored using Chi-square and the inference of Twitter users' *home location* was framed as text classification tasks. A character-based recurrent neural network model with attention mechanism for such purpose was proposed. Additionally, a set of competitive machine learning and deep learning algorithms, employing a publicly available Twitter users' location dataset, was evaluated. The proposed model outperformed convolutional neural network models and conventional machine learning baseline models on the Twitter users' *home location* inference. The study also revealed that the use of transfer learning could improve the accuracy of the *home location* inference on our HPV Twitter users' location dataset. Chi-square analyses identified interstate variations in public perceptions of the HPV vaccine.

Chapter 5 described efforts to build a web-based interactive visualization system to surveil public perceptions regarding the HPV vaccine from real-time Twitter discussions. This chapter reports on a prototype of *VaxInsight*, a web-based interactive dashboard to synthesize HPV vaccine-related Twitter discussions in a comprehendible format. The prototype is based on previous Twitter mapping work and uses constructs from behavior change theories (Chapter 2). It also allows for Twitter demographic attributes inference (Chapter 3 and 4) and permits users to filter Twitter discussions and track trends for a particular theoretical construct. Furthermore, it can compare demographic differences in public perceptions using visualizations. In addition, *VaxInsight* provides a web-based interface for users to browse and validate predictions of deep learning systems. A group of researchers with backgrounds in health promotion was recruited and they performed a usability test of *VaxInsight*. The survey results showed high usability of the system.

## 6.2. Innovation and Contribution

Notably, this is the first study to use deep learning algorithms to understand Twitter discussions of the HPV vaccine that is framed in grounded behavior change theories. *VaxInsight* is also the first system that allows users to explore public health beliefs on vaccine-related topics from Twitter. In the following discussion, the major innovations and contributions of this study from two perspectives, medical informatics and methodology, are identified.

### 6.2.1. Medical informatics

This work contributes to medical informatics, especially public health informatics, by proposing a system that uses novel artificial intelligence algorithms (i.e. deep learning) and social media data to monitor public perceptions regarding vaccines for the purpose of vaccination promotion. It is innovative in the following aspects:

1) It is hybrid in nature as it combines a data driven approach from informatics and a theory driven approach from public health.
2) It transforms social media data into aggregate population health level attributes, which is an improvement over existing efforts that focus on the understanding of Twitter content, only.
3) It allows for informed decision making by policy makers as it synthesizes social media data, behavior change theories, and interactive data visualizatons into a comprehendible format.

### 6.2.2. Methodology

The current study proposed and evaluated various machine learning and deep learning algorithms using Twitter data. Specifically, it is innovative from a methodological perspective in the following ways:

1) It proposed a novel deep learning framework that was based on a bi-directional long short-term memory network with attention mechanism for mapping Twitter discussions to the constructs of behavior change theories. The study's algorithm performed better than traditional machine learning and competitive deep learning baselines on a majority of tasks.
2) It proposed a novel character-based convolutional neural network model for a Twitter gender inference task. The CNN model with embedding fusion achieved comparable state-of-the-art performance in an open general-domain Twitter gender inference challenge.
3) It framed Twitter users' *home location* inference as text classification tasks and proposed a novel character-based recurrent neural network model for a geo-tagging task. The proposed model performed better than other competitive machine learning and deep learning algorithms in a general-domain Twitter users' location dataset.

## 6.3. Limitations and Future Work

This current research makes original and systematical contributions to medical informatics. It provides a foundation for the next generation of real-time public health surveillance by combining cutting-edge artificial intelligence algorithms with behavior change theories. However, it only focuses on Twitter users. These users are not representative of the general population. Therefore, future research should test the model developed in this study with other social media platforms and data sources. Particular attention should also be paid to the influence of malicious bots which appear on these platforms and are created to spread negative and unfounded information about vaccinations. Removing bots (both benign and malicious) is an essential step in order to more accurately reflect public opinions. Accordingly, future research should adopt bot detection methods and remove discussions which are not generated by humans.

A furthermore limitation of the present study is that the algorithms suffered suboptimal performance on certain tasks. This may be overcome by incorporating the rapid advances in natural language processing and deep learning algorithms that are taking place, currently. Additionally, future research should integrate feedback from researchers in the field of vaccine promotion that may then enhance the model by adding relevant theoretical constructs and demographic attributes.  Additional detailed information could provide a more comprehensive understanding of the users, their perceptions, and their intentions. Ultimately, the goal of research in this area and of the present study, is to use social media platforms to promote vaccination behavior. Therefore, an evaluation of tailored and personalized messages aimed at promoting vaccinations by using artificial intelligence systems will need to take place.

## References

1 Omer SB, Salmon DA, Orenstein WA, *et al.* Vaccine refusal, mandatory immunization, and the risks of vaccine-preventable diseases. *N Engl J Med* 2009;**360**:1981–8. doi:10.1056/NEJMsa0806477

2 Kwok R. Vaccines: The real issues in vaccine safety. *Nature* 2011;**473**:436–8. doi:10.1038/473436a

3 Maglione MA, Das L, Raaen L, *et al.* Safety of Vaccines Used for Routine Immunization of US Children: A Systematic Review. *Pediatrics* 2014;**134**:325–37. doi:10.1542/peds.2014-1079

4 Vaccine Safety. https://www.vaccines.gov/basics/safety

5 'Centers for Disease Control and Prevention'. Achievements in public health, 1990-1999: impact of vaccines universally recommended for children. *United States* 1990;**1998**:243–8.

6 Milestones in Public Health. http://www.aspph.org/app/uploads/2014/02/Milestones_in_Public_Health.pdf

7 Dubé E, Vivion M, MacDonald NE. Vaccine hesitancy, vaccine refusal and the anti-vaccine movement: influence, impact and implications. *Expert Rev Vaccines* 2015;**14**:99–117. doi:10.1586/14760584.2015.964212

8 Lieu TA, Ray GT, Klein NP, *et al.* Geographic clusters in underimmunization and vaccine refusal. *Pediatrics* 2015;**135**:280–9.

9 Omer SB, Pan WKY, Halsey NA, *et al.* Nonmedical exemptions to school immunization requirements: secular trends and association of state policies with pertussis incidence. *Jama* 2006;**296**:1757–63.

10 Dempsey AF, Schaffer S, Singer D, *et al.* Alternative vaccination schedule preferences among parents of young children. *Pediatrics* 2011;:peds--2011.

11 'American Academy of of Pediatrics'. Documenting parental refusal to have their children vaccinated. 2014.

12 Parents Distrust, Delay Vaccines, Survey Finds. https://www.livescience.com/16346-parents-distrust-delay-vaccines.html

13 1 in 10 Parents Skip or Delay Vaccines. http://healthland.time.com/2011/10/03/more-than-1-in-10-parents-skip-or-delay-vaccines/

14 Richards JL, Wagenaar BH, Van Otterloo J, *et al.* Nonmedical exemptions to immunization requirements in California: a 16-year longitudinal analysis of trends and associated community factors. *Vaccine* 2013;**31**:3009–13.

15 Sadaf A, Richards JL, Glanz J, *et al.* A systematic review of interventions for reducing parental vaccine refusal and vaccine hesitancy. *Vaccine* 2013;**31**:4293–304. doi:10.1016/j.vaccine.2013.07.013

16 Phadke VK, Bednarczyk RA, Salmon DA, *et al.* Association between vaccine refusal and vaccine-preventable diseases in the United States A review of measles and pertussis. JAMA - J. Am. Med. Assoc. 2016;**315**:1149–58. doi:10.1001/jama.2016.1353

17 Antona D, Lévy-Bruhl D, Baudon C, *et al.* Measles elimination efforts and 2008--2011

outbreak, France. *Emerg Infect Dis* 2013;**19**:357.
18 Zipprich J, Winter K, Hacker J, *et al.* Measles outbreak—California, December 2014-February 2015. *MMWR Morb Mortal Wkly Rep* 2015;**64**:153–4.
19 Paules CI, Marston HD, Fauci AS. Measles in 2019 — Going Backward. *N Engl J Med* Published Online First: 2019. doi:10.1056/nejmp1905099
20 McClure CC, Cataldi JR, O'Leary ST. Vaccine Hesitancy: Where We Are and Where We Are Going. *Clin Ther* 2017;**39**:1550–62. doi:10.1016/j.clinthera.2017.07.003
21 Cooper LZ, Larson HJ, Katz SL. Protecting public trust in immunization. *Pediatrics* 2008;**122**:149–53. doi:DOI 10.1542/peds.2008-0987
22 Zhao Z, Luman ET. Progress toward eliminating disparities in vaccination coverage among US children, 2000--2008. *Am J Prev Med* 2010;**38**:127–37.
23 Zimet GD, Weiss TW, Rosenthal SL, *et al.* Reasons for non-vaccination against HPV and future vaccination intentions among 19-26 year-old women. *BMC Womens Health* 2010;**10**:27.
24 Dredze M, Broniatowski DA, Smith MC, *et al.* Understanding Vaccine Refusal. *Am J Prev Med* 2016;**50**:550–2.
25 Poland GA, Jacobson RM. The clinician's guide to the anti-vaccinationists' galaxy. *Hum Immunol* 2012;**73**:859–66.
26 Kata A. Anti-vaccine activists, Web 2.0, and the postmodern paradigm - An overview of tactics and tropes used online by the anti-vaccination movement. *Vaccine* 2012;**30**:3778–89. doi:10.1016/j.vaccine.2011.11.112
27 Jolley D, Douglas KM. The effects of anti-vaccine conspiracy theories on vaccination intentions. *PLoS One* 2014;**9**. doi:10.1371/journal.pone.0089177
28 Wakefield AJ, Murch SH, Anthony A, *et al.* RETRACTED: Ileal-lymphoid-nodular hyperplasia, non-specific colitis, and pervasive developmental disorder in children. 1998.
29 Taylor B, Miller E, Farrington Cp, *et al.* Autism and measles, mumps, and rubella vaccine: no epidemiological evidence for a causal association. *Lancet* 1999;**353**:2026–9.
30 Godlee F, Smith J, Marcovitch H. Wakefield's article linking MMR vaccine and autism was fraudulent. 2011.
31 Larson HJ, Smith DMDD, Paterson P, *et al.* Measuring vaccine confidence: Analysis of data obtained by a media surveillance system used to analyse public concerns about vaccines. *Lancet Infect Dis* 2013;**13**:606–13. doi:10.1016/S1473-3099(13)70108-7
32 Lawrence HY, Hausman BL, Dannenberg CJ. Reframing medicine's publics: The local as a public of vaccine refusal. *J Med Humanit* 2014;**35**:111–29.
33 Organization WH, others. The guide to tailoring immunization programmes (TIP). *World Heal Organ Copenhagen, Denmark* 2013.
34 Yaqub O, Castle-Clarke S, Sevdalis N, *et al.* Attitudes to vaccination: a critical review. *Soc Sci Med* 2014;**112**:1–11.
35 Cox DS, Cox AD, Sturm L, *et al.* Behavioral interventions to increase HPV vaccination acceptability among mothers of young girls. *Heal Psychol* 2010;**29**:29.
36 Cates JR, Ortiz R, Shafer A, *et al.* Designing messages to motivate parents to get their preteenage sons vaccinated against human papillomavirus. *Perspect Sex Reprod Health* 2012;**44**:39–47.

37 Clayton EW, Hickson GB, Miller CS. Parents’ responses to vaccine information pamphlets. *Pediatrics* 1994;**93**:369–72.

38 Kagashe I, Yan Z, Suheryani I. Enhancing Seasonal Influenza Surveillance: Topic Analysis of Widely Used Medicinal Drugs Using Twitter Data. *J Med Internet Res* 2017;**19**:315–3151. doi:10.2196/jmir.7393

39 Labarthe DR, Merchant RM, Jha S, *et al.* Psychological language on Twitter predicts county-level heart disease mortality. *Psychol Sci* 2015;**26**:159–69. doi:10.1177/0956797614557867.Psychological

40 Chan B, Lopez A, Sarkar U. The canary in the coal mine tweets: social media reveals public perceptions of non-medical use of opioids. *PLoS One* 2015;**10**:e0135072. doi:10.1371/journal.pone.0135072

41 Mitra T, Counts S, Pennebaker J. Understanding anti-vaccination attitudes in social media. *Proc 10th Int Conf Web Soc Media, ICWSM 2016* 2016;:269–78.

42 Grimm P. Social desirability bias. *Wiley Int Encycl Mark* 2010.

43 Tomeny TS, Vargo CJ, El-Toukhy S. Geographic and demographic correlates of autism-related anti-vaccine beliefs on Twitter, 2009-15. *Soc Sci Med* 2017;**191**:168–75. doi:10.1016/j.socscimed.2017.08.041

44 Kaplan AM, Haenlein M. Users of the world, unite! The challenges and opportunities of Social Media. *Bus Horiz* 2010;**53**:59–68.

45 Moorhead SA, Hazlett DE, Harrison L, *et al.* A new dimension of health care: systematic review of the uses, benefits, and limitations of social media for health communication. *J Med Internet Res* 2013;**15**:1–16. doi:10.2196/jmir.1933

46 Global Digital Report 2019. https://hootsuite.com/pages/digital-in-2019

47 Social Media Fact Sheet. http://www.pewinternet.org/fact-sheet/social-media/

48 Patel R, Chang T, Greysen SR, *et al.* Social media use in chronic disease: A systematic review and novel taxonomy. *Am J Med* 2015;**128**:1335–50. doi:10.1016/j.amjmed.2015.06.015

49 Househ M, Borycki E, Kushniruk A. Empowering patients through social media: The benefits and challenges. *Health Informatics J* 2014;**20**:50–8. doi:10.1177/1460458213476969

50 Taggart T, Grewe ME, Conserve DF, *et al.* Social media and HIV: A systematic review of uses of social media in HIV communication. *J Med Internet Res* 2015;**17**. doi:10.2196/jmir.4387

51 Sarker A, Ginn R, Nikfarjam A, *et al.* Utilizing social media data for pharmacovigilance: A review. *J Biomed Inform* 2015;**54**:202–12. doi:10.1016/j.jbi.2015.02.004

52 Nikfarjam A, Sarker A, O’Connor K, *et al.* Pharmacovigilance from social media: mining adverse drug reaction mentions using sequence labeling with word embedding cluster features. *J Am Med Informatics Assoc* 2015;**22**:ocu041. doi:10.1093/jamia/ocu041

53 Golder S, Norman G, Loke YK. Systematic review on the prevalence, frequency and comparative value of adverse events data in social media. *Br J Clin Pharmacol* 2015;**80**:878–88. doi:10.1111/bcp.12746

54 Lardon J, Abdellaoui R, Bellet F, *et al.* Adverse drug reaction identification and extraction in social media: A scoping review. *J Med Internet Res* 2015;**17**:1–16.

doi:10.2196/jmir.4304

55 Grindrod K, Forgione A, Tsuyuki RT, *et al.* Pharmacy 2.0: A scoping review of social media use in pharmacy. *Res Soc Adm Pharm* 2014;**10**:256–70. doi:10.1016/j.sapharm.2013.05.004

56 Korda H, Itani Z. Harnessing Social Media for Health Promotion and Behavior Change. *Health Promot Pract* 2013;**14**:15–23. doi:10.1177/1524839911405850

57 Webb TL, Joseph J, Yardley L, *et al.* Using the Internet to promote health behavior change: A systematic review and meta-analysis of the impact of theoretical basis, use of behavior change techniques, and mode of delivery on efficacy. *J Med Internet Res* 2010;**12**:1–18. doi:10.2196/jmir.1376

58 Signorini A, Segre AM, Polgreen PM. The use of Twitter to track levels of disease activity and public concern in the U.S. during the influenza A H1N1 pandemic. *PLoS One* 2011;**6**. doi:10.1371/journal.pone.0019467

59 Scotch MM, Yip KY, Cheung K-HK-H. Development of grid-like applications for public health using Web 2.0 mashup techniques. *J Am Med Informatics Assoc* 2008;**15**:783–6. doi:10.1197/jamia.M2731.Introduction

60 Bernardo TM, Rajic A, Young I, *et al.* Scoping review on search queries and social media for disease surveillance: A chronology of innovation. *J Med Internet Res* 2013;**15**:1–13. doi:10.2196/jmir.2740

61 Guntuku SC, Yaden DB, Kern ML, *et al.* Detecting depression and mental illness on social media: an integrative review. *Curr Opin Behav Sci* 2017;**18**:43–9. doi:10.1016/j.cobeha.2017.07.005

62 Braithwaite SR, Giraud-Carrier C, West J, *et al.* Validating Machine Learning Algorithms for Twitter Data Against Established Measures of Suicidality. *JMIR Ment Heal* 2016;**3**:e21. doi:10.2196/mental.4822

63 Du J, Zhang Y, Luo J, *et al.* Extracting psychiatric stressors for suicide from social media using deep learning. *BMC Med Inform Decis Mak* 2018;**18**:43.

64 Ortiz RR, Smith A, Coyne-Beasley T. A systematic literature review to examine the potential for social media to impact HPV vaccine uptake and awareness, knowledge, and attitudes about HPV and HPV vaccination. *Hum Vaccines Immunother* Published Online First: 2019. doi:10.1080/21645515.2019.1581543

65 Smith PJ, Humiston SG, Marcuse EK, *et al.* Parental delay or refusal of vaccine doses, childhood vaccination coverage at 24 months of age, and the Health Belief Model. *Public Health Rep* 2011;**126**:135–46. doi:10.1177/00333549111260S215

66 Dube E, Bettinger JA, Halperin B, *et al.* Determinants of parents' decision to vaccinate their children against rotavirus: results of a longitudinal study. *Health Educ Res* 2012;**27**:1069–80.

67 Shapiro GK, Surian D, Dunn AG, *et al.* Comparing human papillomavirus vaccine concerns on Twitter: a cross-sectional study of users in Australia, Canada and the UK. *BMJ Open* 2017;**7**:e016869. doi:10.1136/bmjopen-2017-016869

68 Massey PM, Leader A, Yom-Tov E, *et al.* Applying multiple data collection tools to quantify human papillomavirus vaccine communication on twitter. *J Med Internet Res* 2016;**18**:1–11. doi:10.2196/jmir.6670

69 Ache KA, Wallace LS. Human Papillomavirus Vaccination Coverage on YouTube. *Am J Prev Med* 2008;**35**:389–92. doi:10.1016/j.amepre.2008.06.029

70 Keim-Malpass J, Mitchell EM, Sun E, *et al.* Using Twitter to Understand Public Perceptions Regarding the #HPV Vaccine: Opportunities for Public Health Nurses to Engage in Social Marketing. *Public Health Nurs* 2017;**34**:316–23. doi:10.1111/phn.12318

71 Robichaud P, Hawken S, Beard L, *et al.* Vaccine-critical videos on YouTube and their impact on medical students' attitudes about seasonal influenza immunization: A pre and post study. *Vaccine* 2012;**30**:3763–70. doi:10.1016/j.vaccine.2012.03.074

72 Lehmann BA, Ruiter RAC, Kok G. A qualitative study of the coverage of influenza vaccination on Dutch news sites and social media websites. *BMC Public Health* 2013;**13**:1–13. doi:10.1186/1471-2458-13-547

73 Krittanawong C, Tunhasiriwet A, Chirapongsathorn S, *et al.* Tweeting influenza vaccine to cardiovascular health community. *Eur J Cardiovasc Nurs* 2017;:147451511770786. doi:10.1177/1474515117707867

74 McNeill A, Harris PR, Briggs P. Twitter Influence on UK Vaccination and Antiviral Uptake during the 2009 H1N1 Pandemic. *Front public Heal* 2016;**4**:26. doi:10.3389/fpubh.2016.00026

75 Huang X, Smith MC, Paul MJ, *et al.* Examining Patterns of Influenza Vaccination in Social Media. *Work Thirty-First AAAI Conf Artif Intell* 2017;:1–5.http://www.cs.jhu.edu/~mdredze/publications/2017_w3phi_vaccines.pdf

76 Dredze M, Broniatowski DA, Hilyard KM, *et al.* Zika vaccine misconceptions: A social media analysis. *Vaccine* 2016;**34**:3441. doi:10.2196/jmir.1933

77 Ghenai A, Mejova Y. Catching Zika fever: Application of crowdsourcing and machine learning for tracking health misinformation on Twitter. *arXiv Prepr arXiv170703778* 2017.

78 Seltzer EK, Horst-Martz E, Lu M, *et al.* Public sentiment and discourse about Zika virus on Instagram. *Public Health* 2017;**150**:170–5. doi:10.1016/j.puhe.2017.07.015

79 Chen B, Zhang JM, Jiang Z, *et al.* Media and public reactions toward vaccination during the 'hepatitis B vaccine crisis' in China. *Vaccine* 2015;**33**:1780–5. doi:10.1016/j.vaccine.2015.02.046

80 Becker BFHH, Larson HJ, Bonhoeffer J, *et al.* Evaluation of a multinational, multilingual vaccine debate on Twitter. *Vaccine* 2016;**34**:6166–71. doi:10.1016/j.vaccine.2016.11.007

81 Orr D, Baram-Tsabari A, Landsman K. Social media as a platform for health-related public debates and discussions: the Polio vaccine on Facebook. *Isr J Health Policy Res* 2016;**5**:34. doi:10.1186/s13584-016-0093-4

82 Radzikowski J, Stefanidis A, Jacobsen KH, *et al.* The Measles Vaccination Narrative in Twitter: A Quantitative Analysis. *JMIR Public Heal Surveill* 2016;**2**:e1. doi:10.2196/publichealth.5059

83 Kang GJ, Ewing-Nelson SR, Mackey L, *et al.* Semantic network analysis of vaccine sentiment in online social media. *Vaccine* 2017;**35**:3621–38. doi:10.1016/j.vaccine.2017.05.052

84 Basch CH, Zybert P, Reeves R, *et al.* What do popular YouTubeTM videos say about vaccines? *Child Care Health Dev* 2017;**43**:499–503. doi:10.1111/cch.12442

85 Guidry JPD, Carlyle K, Messner M, *et al.* On pins and needles: How vaccines are portrayed on Pinterest. *Vaccine* 2015;**33**:5051–6. doi:10.1016/j.vaccine.2015.08.064
86 Love B, Himelboim I, Holton A, *et al.* Twitter as a source of vaccination information: content drivers and what they are saying. *Am J Infect Control* 2013;**41**:568–70. doi:10.1016/j.ajic.2012.10.016
87 Myneni S, Fujimoto K, Cobb N, *et al.* Content-driven analysis of an online community for smoking cessation: integration of qualitative techniques, automated text analysis, and affiliation networks. *J Inf* 2015;**105**.
88 Centola D. The spread of behavior in an online social network experiment. *Science (80- )* 2010;**329**:1194–7.
89 Cobb NK, Graham AL, Abrams DB. Social network structure of a large online community for smoking cessation. *Am J Public Health* 2010;**100**:1282–9.
90 Keelan J, Pavri V, Balakrishnan R, *et al.* An analysis of the Human Papilloma Virus vaccine debate on MySpace blogs. *Vaccine* 2010;**28**:1535–40.
91 Gelaye B, Rondon M, Araya PR, *et al.* Comparison of different recruitment methods for sexual and reproductive health research: Social media-based versus conventional methods. *Vaccine* 2017;**35**:e016869. doi:10.1016/j.vaccine.2017.06.066
92 Dunn AG, Surian D, Leask J, *et al.* Mapping information exposure on social media to explain differences in HPV vaccine coverage in the United States. *Vaccine* 2017;**35**:3033–40. doi:10.1016/j.vaccine.2017.04.060
93 Surian D, Nguyen DQ, Kennedy G, *et al.* Characterizing Twitter Discussions About HPV Vaccines Using Topic Modeling and Community Detection. *J Med Internet Res* 2016;**18**.
94 Brooks B. *Using Twitter data to identify geographic clustering of anti-vaccination sentiments*. ProQuest Diss. Theses. 2014.
95 Zhou X, Coiera E, Tsafnat G, *et al.* Using social connection information to improve opinion mining: Identifying negative sentiment about HPV vaccines on Twitter. *Stud Health Technol Inform* 2015;**216**:761–5. doi:10.3233/978-1-61499-564-7-761
96 Salathé M, Khandelwal S. Assessing vaccination sentiments with online social media: Implications for infectious disease dynamics and control. *PLoS Comput Biol* 2011;**7**. doi:10.1371/journal.pcbi.1002199
97 Flu Vaccination Coverage, United States, 2016-17 Influenza Season. https://www.cdc.gov/flu/fluvaxview/coverage-1617estimates.htm#data
98 TeenVaxView. https://www.cdc.gov/vaccines/imz-managers/coverage/teenvaxview/index.html
99 Myers LB, Goodwin R. Determinants of adults' intention to vaccinate against pandemic swine flu. *BMC Public Health* 2011;**11**:15.
100 Bynum SA, Brandt HM, Friedman DB, *et al.* Knowledge, beliefs, and behaviors: Examining human papillomavirus-related gender differences among african American College Students. *J Am Coll Heal* Published Online First: 2011. doi:10.1080/07448481.2010.503725
101 Kim HW. Gender differences in knowledge and health beliefs related to behavioral intentions to prevent human papillomavirus infection. *Asia-Pacific J Public Heal* Published Online First: 2013. doi:10.1177/1010539512444307

102 Siddiqui M, Salmon DA, Omer SB. Epidemiology of vaccine hesitancy in the United States. *Hum Vaccin Immunother* 2013;**9**:2643–8.

103 National Immunization Surveys (NIS). https://www.cdc.gov/vaccines/imz-managers/nis/index.html

104 Powell GA, Zinszer K, Verma A, *et al.* Media content about vaccines in the United States and Canada, 2012-2014: An analysis using data from the Vaccine Sentimeter. *Vaccine* 2016;**34**:6229–35. doi:10.1016/j.vaccine.2016.10.067

105 Satterwhite CL, Torrone E, Meites E, *et al.* Sexually transmitted infections among US women and men: Prevalence and incidence estimates, 2008. *Sex Transm Dis* Published Online First: 2013. doi:10.1097/OLQ.0b013e318286bb53

106 'Centers for Disease Control and prevention'. 6 Reasons to Get HPV Vaccine for Your Child. https://www.cdc.gov/hpv/infographics/vacc-six-reasons.pdf

107 Saraiya M, Unger ER, Thompson TD, *et al.* US assessment of HPV types in cancers: implications for current and 9-valent HPV vaccines. *JNCI J Natl Cancer Inst* 2015;**107**.

108 'Centers for Disease Control and Prevention'. HPV Vaccination Coverage Data. https://www.cdc.gov/hpv/hcp/vacc-coverage/index.html

109 Gilkey MB, Calo WA, Marciniak MW, *et al.* Parents who refuse or delay HPV vaccine: Differences in vaccination behavior, beliefs, and clinical communication preferences. *Hum Vaccines Immunother* 2017;**13**:680–6. doi:10.1080/21645515.2016.1247134

110 Brown S. The new deficit model. *Nat Nanotechnol* 2009;**4**:609.

111 Corace KM, Srigley JA, Hargadon DP, *et al.* Using behavior change frameworks to improve healthcare worker influenza vaccination rates: a systematic review. *Vaccine* 2016;**34**:3235–42.

112 Kessels SJMM, Marshall HS, Watson M, *et al.* Factors associated with HPV vaccine uptake in teenage girls: a systematic review. *Vaccine* 2012;**30**:3546–56. doi:10.1016/j.vaccine.2012.03.063

113 Hughes CC, Jones AL, Feemster KA, *et al.* HPV vaccine decision making in pediatric primary care: a semi-structured interview study. *BMC Pediatr* 2011;**11**:74.

114 Thompson EL, Rosen BL, Vamos CA, *et al.* Human papillomavirus vaccination: what are the reasons for nonvaccination among US adolescents? *J Adolesc Heal* 2017;**61**:288–93.

115 Wong LP. Knowledge and attitudes about HPV infection, HPV vaccination, and cervical cancer among rural southeast Asian women. *Int J Behav Med* 2011;**18**:105–11.

116 Reiter PL, Brewer NT, Gottlieb SL, *et al.* Parents' health beliefs and HPV vaccination of their adolescent daughters. *Soc Sci Med* 2009;**69**:475–80.

117 Donadiki EM, Jiménez-García R, Hernández-Barrera V, *et al.* Health Belief Model applied to non-compliance with HPV vaccine among female university students. *Public Health* 2014;**128**:268–73. doi:10.1016/j.puhe.2013.12.004

118 Skinner CS, Tiro J, Champion VL. The health belief model. *Heal Behav theory, Res Pract 5th ed San Fr Jossey-Bass* 2015;:75–94.

119 Mehta P, Sharma M, Lee RC. Designing and evaluating a health belief model-based intervention to increase intent of HPV vaccination among college males. *Int Q Community Health Educ* 2014;**34**:101–17.

120 Gerend MA, Shepherd JE. Predicting human papillomavirus vaccine uptake in young

adult women: Comparing the health belief model and theory of planned behavior. *Ann Behav Med* 2012;**44**:171–80. doi:10.1007/s12160-012-9366-5

121 Askelson NM, Campo S, Lowe JB, *et al.* Using the theory of planned behavior to predict mothers' intentions to vaccinate their daughters against HPV. *J Sch Nurs* 2010;**26**:194–202.

122 Kahn JA, Rosenthal SL, Hamann T, *et al.* Attitudes about human papillomavirus vaccine in young women. *Int J STD AIDS* 2003;**14**:300–6.

123 Ajzen I. The theory of planned behavior. *Organ Behav Hum Decis Process* 1991;**50**:179–211.

124 McDonald L, Malcolm B, Ramagopalan S, *et al.* Real-world data and the patient perspective: The PROmise of social media? BMC Med. 2019. doi:10.1186/s12916-018-1247-8

125 Du J, Xu J, Song H-Y, *et al.* Leveraging machine learning-based approaches to assess human papillomavirus vaccination sentiment trends with Twitter data. *BMC Med Inform Decis Mak* 2017;**17**:69.

126 Dunn AG, Leask J, Zhou X, *et al.* Associations between exposure to and expression of negative opinions about human papillomavirus vaccines on social media: An observational study. J. Med. Internet Res. 2015;**17**:e144. doi:10.2196/jmir.4343

127 Nakov P, Ritter A, Rosenthal S, *et al.* SemEval-2016 task 4: Sentiment analysis in Twitter. In: *Proceedings of SemEval*. 2016. 1–18.

128 Mohammad S, Bravo-Marquez F, Salameh M, *et al.* Semeval-2018 task 1: Affect in tweets. In: *Proceedings of The 12th International Workshop on Semantic Evaluation*. 2018. 1–17.

129 Rosenthal S, Farra N, Nakov P. SemEval-2017 task 4: Sentiment analysis in Twitter. In: *Proceedings of the 11th International Workshop on Semantic Evaluation (SemEval-2017)*. 2017. 502–18.

130 Du J, Tang L, Xiang Y, *et al.* Public Perception Analysis of Tweets During the 2015 Measles Outbreak: Comparative Study Using Convolutional Neural Network Models. *J Med Internet Res* 2018;**20**.

131 Du J, Cunningham RM, Xiang Y, *et al.* Leveraging deep learning to understand health beliefs about the Human Papillomavirus Vaccine from social media. *npj Digit Med* 2019;**2**:27.

132 Goodfellow I, Bengio Y, Courville A. *Deep learning*. MIT Press 2016.

133 LeCun Y, Bengio Y, Hinton G. Deep learning. *Nature* 2015;**521**:436–44.

134 Young T, Hazarika D, Poria S, *et al.* Recent trends in deep learning based natural language processing. *IEEE Comput Intell Mag* 2018;**13**:55–75. doi:10.1109/MCI.2018.2840738

135 Twitter by the Numbers: Stats, Demographics & Fun Facts. https://www.omnicoreagency.com/twitter-statistics/

136 O'Connor B, Balasubramanyan R, Routledge BR, *et al.* From tweets to polls: Linking text sentiment to public opinion time series. *Icwsm* 2010;**11**:1–2.

137 Paul MJ, Dredze M. You are what you Tweet: Analyzing Twitter for public health. In: *5th International AAAI Conference on Weblogs and Social Media*. 2011. 265–72.

138 Go A, Huang L, Bhayani R. Twitter sentiment analysis. *Entropy* 2009;**17**:252.

139 Du J, Xu J, Song H, *et al.* Optimization on machine learning based approaches for sentiment analysis on HPV vaccines related tweets. *J Biomed Semantics* 2017;**8**:9. doi:10.1186/s13326-017-0120-6

140 Mikolov T, Sutskever I, Chen K, *et al.* Distributed representations of words and phrases and their compositionality. In: *Advances in neural information processing systems*. 2013. 3111–9.

141 Pennington J, Socher R, Manning CD. GloVe: Global Vectors for Word Representation. Empir. Methods Nat. Lang. Process. 2014;:1532–43.https://nlp.stanford.edu/projects/glove/

142 Bojanowski P, Grave E, Joulin A, *et al.* Enriching word vectors with subword information. *arXiv Prepr arXiv160704606* 2016.

143 Google Code Archive: word2vec. https://code.google.com/archive/p/word2vec/

144 Go A, Bhayani R, Huang L. Twitter sentiment classification using distant supervision. *CS224N Proj Report, Stanford* 2009;**1**:1–6. doi:10.1016/j.sedgeo.2006.07.004

145 Kumar A, Sebastian TM. Sentiment analysis on twitter. *Int J Comput Sci Issues* 2012;**9**:372.

146 Kim Y. Convolutional neural networks for sentence classification. *arXiv Prepr arXiv14085882* 2014.

147 Xu H, Kotov A, Dong M, *et al.* Text classification with topic-based word embedding and Convolutional Neural Networks. In: *7th ACM Conference on Bioinformatics, Computational Biology, and Health Informatics, ACM-BCB 2016*. Department of Computer Science, Wayne State University, Detroit, MI 48202, United States Pediatric Prevention Research Center, Wayne State University, Detroit, MI 48201, United States: : Association for Computing Machinery, Inc 88–97.

148 Liu P, Qiu X, Huang X. Recurrent neural network for text classification with multi-task learning. *arXiv Prepr arXiv160505101* 2016;:2873–9.

149 Yao L, Mao C, Luo Y. Graph Convolutional Networks for Text Classification. *Proc AAAI Conf Artif Intell* 2019;**33**:7370–7. doi:10.1609/aaai.v33i01.33017370

150 Howard J, Ruder S. Universal Language Model Fine-tuning for Text Classification. http://nlp.fast.ai/ulmfit.

151 Zhou P, Shi W, Tian J, *et al.* Attention-based bidirectional long short-term memory networks for relation classification. In: *Proceedings of the 54th Annual Meeting of the Association for Computational Linguistics (Volume 2: Short Papers)*. 2016. 207–12. doi:10.18653/v1/P16-2034

152 Peters ME, Neumann M, Iyyer M, *et al.* Deep contextualized word representations. *arXiv Prepr arXiv180205365* 2018.

153 Devlin J, Chang M-W, Lee K, *et al.* Bert: Pre-training of deep bidirectional transformers for language understanding. *arXiv Prepr arXiv181004805* 2018.

154 Vaswani A, Shazeer N, Parmar N, *et al.* Attention is all you need. In: *Advances in Neural Information Processing Systems*. 2017. 5998–6008.

155 Geurts P, Ernst D, Wehenkel L. Extremely randomized trees. *Mach Learn* 2006;**63**:3–42.

156 Ramos J, others. Using tf-idf to determine word relevance in document queries. In:

*Proceedings of the first instructional conference on machine learning*. 2003. 133–42.

157 Valentini G, Masulli F. Ensembles of learning machines. In: *Italian workshop on neural nets*. 2002. 3–20.

158 Cleveland RB, Cleveland WS, McRae JE, *et al.* STL: a seasonal-trend decomposition. *J Off Stat* 1990;**6**:3–73.

159 R Foundation for Statistical Computing. R: A languge and environment for statistical computing.

160 Stojanovski D, Strezoski G, Madjarov G, *et al.* Twitter sentiment analysis using deep convolutional neural network. In: *International Conference on Hybrid Artificial Intelligence Systems*. 2015. 726–37. doi:10.1007/978-3-319-19644-2_60

161 Chatterjee A, Narahari KN, Joshi M, *et al.* SemEval-2019 Task 3: EmoContext Contextual Emotion Detection in Text. In: *Proceedings of the 13th International Workshop on Semantic Evaluation*. 2019. 39–48.

162 Roberts K. Assessing the Corpus Size vs. Similarity Trade-off for Word Embeddings in Clinical NLP. *Clin 2016* 2016;:54.

163 Nikolov A, Radivchev V. Nikolov-Radivchev at SemEval-2019 Task 6: Offensive Tweet Classification with BERT and Ensembles. In: *Proceedings of the 13th International Workshop on Semantic Evaluation*. 2019. 691–5.

164 Ma G. Tweets Classification with BERT in the Field of Disaster Management. http://web.stanford.edu/class/cs224n/reports/custom/15785631.pdf (accessed 18 Jun 2019).

165 Peng Y, Yan S, Lu Z. Transfer Learning in Biomedical Natural Language Processing: An Evaluation of BERT and ELMo on Ten Benchmarking Datasets. *arXiv Prepr arXiv190605474* 2019.

166 Duan R, Cao M, Wu Y, *et al.* An Empirical Study for Impacts of Measurement Errors on EHR based Association Studies. *AMIA . Annu Symp proceedings AMIA Symp* 2016.

167 Chen Y, Wang J, Chubak J, *et al.* Inflation of type I error rates due to differential misclassification in EHR-derived outcomes: Empirical illustration using breast cancer recurrence. *Pharmacoepidemiol Drug Saf* Published Online First: 2019. doi:10.1002/pds.4680

168 Ganin Y, Lempitsky V. Unsupervised Domain Adaptation by Backpropagation. Published Online First: 26 September 2014.http://arxiv.org/abs/1409.7495 (accessed 20 May 2019).

169 D'Arrigo T. HPV vaccination rates not meeting target goals, gender gap remains. *Pharm Today* 2018;**24**:19.

170 Reimer RA, Schommer JA, Houlihan AE, *et al.* Ethnic and gender differences in HPV knowledge, awareness, and vaccine acceptability among white and hispanic men and women. *J Community Health* Published Online First: 2014. doi:10.1007/s10900-013-9773-y

171 Christy SM, Winger JG, Raffanello EW, *et al.* The role of anticipated regret and health beliefs in HPV vaccination intentions among young adults. *J Behav Med* Published Online First: 2016. doi:10.1007/s10865-016-9716-z

172 Kim HW. Awareness of human papillomavirus and factors associated with intention to obtain HPV vaccination among Korean youth: Quasi experimental study. *BMC Int Health*

*Hum Rights* Published Online First: 2015. doi:10.1186/s12914-015-0042-2
173 Knowles R, Carroll J, Dredze M. Demographer: Extremely simple name demographics. *NLP+ CSS 2016* 2016;:108.
174 gender-detector 0.1.0. https://pypi.org/project/gender-detector/
175 Genderize.io | Determine the gender of a first name. https://genderize.io/ (accessed 26 Apr 2019).
176 Burger JD, Henderson J, Kim G, *et al.* Discriminating gender on Twitter. In: *Proceedings of the Conference on Empirical Methods in Natural Language Processing*. 2011. 1301–9.
177 Basile A, Dwyer G, Medvedeva M, *et al.* N-GRAM: New groningen author-profiling model: Notebook for PAN at CLEF 2017. In: *CEUR Workshop Proceedings*. 2017.
178 Sap M, Park G, Eichstaedt J, *et al.* Developing age and gender predictive lexica over social media. In: *Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP)*. 2014. 1146–51. doi:10.3115/v1/d14-1121
179 Daneshvar S, Inkpen D. Gender identification in Twitter using N-grams and LSA: Notebook for PAN at CLEF 2018. In: *CEUR Workshop Proceedings*. 2018.
180 Volkova S. *Predicting Demographics and Affect in Social Networks*. ProQuest Diss. Theses. 2015.
181 Zheng X, Han J, Sun A. A survey of location prediction on Twitter. *IEEE Trans Knowl Data Eng* 2018.
182 Rangel F, Rosso P, Montes-y-Gómez M, *et al.* Overview of the 6th author profiling task at pan 2018: multimodal gender identification in Twitter. *Work Notes Pap CLEF* 2018.
183 Sierra S, González FA. Combining textual and visual representations for multimodal author profiling: Notebook for PAN at CLEF 2018. In: *CEUR Workshop Proceedings*. 2018.
184 Takahashi T, Tahara T, Nagatani K, *et al.* Text and image synergy with feature cross technique for gender identification: Notebook for PAN at CLEF 2018. In: *CEUR Workshop Proceedings*. 2018.
185 Rangel F, Rosso P, Potthast M, *et al.* Overview of the 5th author profiling task at PAN 2017: Gender and language variety identification in Twitter. In: *CEUR Workshop Proceedings*. 2017.
186 Miura Y, Taniguchi T, Taniguchi M, *et al.* Author profiling with word+character neural attention network: Notebook for PAN at CLEF 2017. In: *CEUR Workshop Proceedings*. 2017.
187 Tweepy. https://www.tweepy.org/
188 LeCun Y, Kavukcuoglu K, Farabet C. Convolutional networks and applications in vision. In: *Circuits and Systems (ISCAS), Proceedings of 2010 IEEE International Symposium on*. 2010. 253–6.
189 Dos Santos CN, Gatti M. Deep Convolutional Neural Networks for Sentiment Analysis of Short Texts. In: *COLING*. 2014. 69–78.
190 Conneau A, Schwenk H, Le Cun Y, *et al.* Very Deep Convolutional Networks for Text Classification. *arXiv Prepr arXiv160601781* 2016.
191 Paulus R. Script for preprocessing tweets. http://nlp.stanford.edu/projects/glove/preprocess-twitter.rb

192 Natural Language Toolkit. https://www.nltk.org/
193 State Vaccination Exemptions for Children Entering Public Schools. https://vaccines.procon.org/state-vaccination-exemptions-for-children-entering-public-schools/
194 2017 Childhood Measles, Mumps, and Rubella (MMR) Vaccination Coverage Report. https://www.cdc.gov/vaccines/imz-managers/coverage/childvaxview/data-reports/mmr/reports/2017.html
195 2017 Adolescent Human Papillomavirus (HPV) Vaccination Coverage Report. https://www.cdc.gov/vaccines/imz-managers/coverage/teenvaxview/data-reports/hpv/reports/2017.html
196 Xie T, Grady C, Cacciatore M, *et al.* Understanding Flu Vaccination Acceptance Among US Adults: The Health Belief Model and Media Sources. Published Online First: 2019.https://stars.library.ucf.edu/icrcc/2019/proceedings/11/ (accessed 24 Sep 2019).
197 Mahmud J, Nichols J, Drews C. Home location identification of twitter users. *ACM Trans Intell Syst Technol* 2014;**5**:47.
198 Mahmud J, Nichols J, Drews C. Where Is This Tweet From? Inferring Home Locations of Twitter Users. *ICWSM* 2012;**12**:511–4.
199 Cheng Z, Caverlee J, Lee K. You are where you tweet: a content-based approach to geo-locating twitter users. In: *Proceedings of the 19th ACM international conference on Information and knowledge management*. 2010. 759–68.
200 Cheng Z, Caverlee J, Lee K. A content-driven framework for geolocating microblog users. *ACM Trans Intell Syst Technol* 2013;**4**:2.
201 Han B, Cook P, Baldwin T. Text-based twitter user geolocation prediction. *J Artif Intell Res* 2014;**49**:451–500.
202 Wing B, Baldridge J. Hierarchical discriminative classification for text-based geolocation. In: *Proceedings of the 2014 conference on empirical methods in natural language processing (EMNLP)*. 2014. 336–48.
203 Rahimi A, Cohn T, Baldwin T. A neural model for user geolocation and lexical dialectology. *arXiv Prepr arXiv170404008* 2017.
204 Miura Y, Taniguchi M, Taniguchi T, *et al.* Unifying text, metadata, and user network representations with a neural network for geolocation prediction. In: *Proceedings of the 55th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*. 2017. 1260–72.
205 twitter-user-geocoder. https://github.com/bianjiang/twitter-user-geocoder
206 Yang Z, Yang D, Dyer C, *et al.* Hierarchical attention networks for document classification. In: *Proceedings of the 2016 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*. 2016. 1480–9.
207 Wang X, Peng Y, Lu L, *et al.* Tienet: Text-image embedding network for common thorax disease classification and reporting in chest x-rays. In: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. 2018. 9049–58.
208 Pan SJ, Fellow QY. A Survey on Transfer Learning. 2009;:1–15.
209 Yosinski J, Clune J, Bengio Y, *et al.* How transferable are features in deep neural

networks? In: *Advances in neural information processing systems*. 2014. 3320–8.
210 Kermany DS, Goldbaum M, Cai W, *et al.* Identifying medical diagnoses and treatable diseases by image-based deep learning. *Cell* 2018;**172**:1122–31. doi:10.1016/j.cell.2018.02.010
211 Liu K, Li Y, Xu N, *et al.* Learn to Combine Modalities in Multimodal Deep Learning. *arXiv Prepr arXiv180511730* 2018.
212 FluView. https://www.cdc.gov/flu/weekly/
213 FluVaxView. https://www.cdc.gov/flu/fluvaxview/reportshtml/trends/index.html
214 Bahk CY, Cumming M, Paushter L, *et al.* Publicly available online tool facilitates real-time monitoring of vaccine conversations and sentiments. *Health Aff* 2016;**35**:341–7. doi:10.1377/hlthaff.2015.1092
215 InSTEDD’s Riff. https://instedd.org/technologies/riff/
216 Dash User Guide. https://dash.plot.ly/
217 system usability scale. https://www.usability.gov/how-to-and-tools/methods/system-usability-scale.html

## Appendix A: Definitions and Examples in Twitter of Key Constructs of Behavior Change Theories

| **Behavior Change Theory** | **Construct** | **Definition** | **Sample Raw Tweets** |
|---|---|---|---|
| Health Belief Model | Perceived susceptibility | The assessment of the risk of getting HPV infections | • hpv is so common almost everyone will be infected with the virus. but it can cause cancer. so why wait? vaccinate!<br>• men equally at risk of hpv infection: boys should also be vaccinated for the human papillomavirus |
| | Perceived severity | The assessment of whether an HPV infection is a sufficient health concern | • learn about the human papillomavirus (hpv), which causes almost all cases of cervical cancer<br>• hpv is more badass than hiv. keep yourself |
| | Perceived benefits | Benefits of the HPV vaccine in protecting against HPV infection, HPV infection-induced cancers, and so forth | • health lifestyle \| here's how the hpv vaccine can help cut the risk of cancer in gay men \| news &gt;<br>• benefits of hpv vaccine can be seen in high school girls, study says |
| | Perceived barriers | Side effects of the HPV vaccine, cost of getting an HPV vaccine, negative news and reports on the HPV vaccine, and so forth. | • hpv vaccine is associated with serious health risks<br>• hpv vax kills a much higher % of young athletic girls |
| Theory of Planned Behavior | Attitude - Positive | Show positive opinion or prompt HPV vaccine. | • Save lives by getting children the HPV vaccine<br>• the hpv vax prevents cancers later in life. |
| | Attitude - Neutral | Related to HPV vaccine topic but contains no sentiment, or sentiment is unclear. | • About I of 10 NJ boys received all 3 doses of HPV vaccine<br>• the myths and facts about hpv vaccines |
| | Attitude - Negative | Concerns or doubts about the HPV vaccine. | • Study reveals 'unaoindable' danger of HPV vaccines<br>• just seen a story, girl got her hpv vaccination, starting having seizures so bad she couldn't stay in school anymore, her junior year! |

# Appendix B: The Impact of Word Embedding Measured by Sensitivity, Specificity, and Accuracy at Levels of Relevance to The Theory (HBM- And TPB-Related) and Theoretical Constructs

| | **Sensitivity** | | | | | **Specificity** | | | | | **Accuracy** | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | GloVe General | FT HPV | GloVe HPV | W2V HPV | Random | GloVe General | FT HPV | GloVe HPV | W2V HPV | Random | GloVe General | FT HPV | GloVe HPV | W2V HPV | Random |
| HBM-related | 0.8061 | **0.8072** | 0.7991 | 0.7969 | 0.7975 | 0.7985 | 0.7954 | 0.7892 | **0.8079** | 0.7583 | **0.8027** | 0.8018 | 0.7946 | 0.8019 | 0.7796 |
| HBM Susceptibility | 0.7263 | 0.6889 | 0.7071 | **0.7606** | 0.5980 | 0.9380 | **0.9396** | 0.9333 | 0.9165 | 0.9295 | **0.9058** | 0.9015 | 0.8989 | 0.8928 | 0.8791 |
| HBM Severity | 0.7597 | 0.7620 | 0.7938 | 0.7767 | **0.7845** | 0.9337 | **0.9419** | 0.9272 | 0.9323 | 0.9130 | 0.8992 | **0.9063** | 0.9008 | 0.9015 | 0.8876 |
| HBM Benefits | 0.7122 | **0.7305** | 0.6860 | 0.7280 | 0.6713 | 0.9121 | **0.9197** | 0.9148 | 0.9133 | 0.8816 | 0.8618 | **0.8721** | 0.8572 | 0.8667 | 0.8287 |
| HBM Barriers | 0.8867 | 0.8890 | 0.8909 | **0.8913** | 0.8744 | 0.9041 | **0.9219** | 0.8983 | 0.9181 | 0.8746 | 0.8959 | **0.9063** | 0.8948 | 0.9054 | 0.8745 |
| TPB-related | **0.9514** | 0.9487 | 0.9471 | 0.9506 | 0.9328 | 0.8690 | 0.8710 | **0.8769** | 0.8583 | 0.8514 | 0.9237 | 0.9226 | 0.9235 | 0.9196 | 0.9054 |
| TPB Attitude | N/A | | | | | | | | | | 0.7509 | **0.7538** | 0.7398 | 0.7534 | 0.7259 |

Note: average score from 10 experiments

# Appendix C: The Impact of Word Embedding Measured by Precision, Recall, and F Score at Levels Of Relevance to The Theory (HBM- And TPB-Related) and Theoretical Constructs

| | **Precision** | | | | | **Recall** | | | | | **F score** | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | GloVe General | FT HPV | GloVe HPV | W2V HPV | Random | GloVe General | FT HPV | GloVe HPV | W2V HPV | Random | GloVe General | FT HPV | GloVe HPV | W2V HPV | Random |
| HBM-related | 0.8274 | 0.8254 | 0.8190 | **0.8321** | 0.7986 | 0.8061 | **0.8072** | 0.7991 | 0.7969 | 0.7975 | **0.8162** | 0.8156 | 0.8087 | 0.8140 | 0.7973 |
| HBM Susceptibility | **0.6846** | 0.6784 | 0.6654 | 0.6252 | 0.6483 | 0.7263 | 0.6889 | 0.7071 | **0.7606** | 0.5980 | **0.7021** | 0.6805 | 0.6814 | 0.6837 | 0.5747 |
| HBM Severity | 0.7437 | **0.7681** | 0.7333 | 0.7422 | 0.6936 | 0.7597 | 0.7620 | **0.7938** | 0.7767 | 0.7845 | 0.7484 | **0.7626** | 0.7593 | 0.7571 | 0.7345 |
| HBM Benefits | 0.7345 | **0.7564** | 0.7320 | 0.7395 | 0.6571 | 0.7122 | **0.7305** | 0.6860 | 0.7280 | 0.6713 | 0.7217 | **0.7407** | 0.7056 | 0.7323 | 0.6618 |
| HBM Barriers | 0.8941 | **0.9123** | 0.8882 | 0.9077 | 0.8633 | 0.8867 | 0.8890 | 0.8909 | **0.8913** | 0.8744 | 0.8898 | **0.8999** | 0.8890 | 0.8992 | 0.8685 |
| TPB-related | 0.9350 | 0.9357 | **0.9383** | 0.9300 | 0.9255 | **0.9514** | 0.9487 | 0.9471 | 0.9506 | 0.9328 | **0.9430** | 0.9421 | 0.9427 | 0.9401 | 0.9291 |
| TPB Attitude (micro-ave) | N/A | | | | | | | | | | 0.7483 | **0.7515** | 0.7383 | 0.7510 | 0.7240 |

Note: average score from 10 experiments

## Appendix D: The Impact of Word Embedding Measured by Precision, Recall, and F-Score on TPB Attitude Classification

| | Positive | | | Negative | | | Neutral | | | Micro-average F-score |
|---|---|---|---|---|---|---|---|---|---|---|
| | Precision | Recall | F-score | Precision | Recall | F-score | Precision | Recall | F-score | |
| GloVe General | **0.7587** | 0.7101 | 0.7326 | 0.7856 | 0.8242 | **0.8028** | 0.7134 | 0.7077 | **0.7093** | 0.7483 |
| FT HPV | 0.7425 | **0.7500** | **0.7447** | 0.7987 | 0.8235 | 0.8103 | **0.7172** | 0.6843 | 0.6996 | **0.7515** |
| GloVe HPV | 0.7457 | 0.7078 | 0.7257 | **0.8016** | 0.7871 | 0.7925 | 0.6822 | **0.7154** | 0.6966 | 0.7383 |
| W2V HPV | 0.7460 | 0.7345 | 0.7396 | 0.7977 | **0.8264** | 0.8108 | 0.7151 | 0.6928 | 0.7025 | 0.7510 |
| Random | 0.7308 | 0.7028 | 0.7157 | 0.7651 | 0.7892 | 0.7759 | 0.6831 | 0.6794 | 0.6805 | 0.7240 |

Note: average score from 10 experiments

# Appendix E: The Comparison of Deep Learning and Machine Learning Algorithms Measured by Sensitivity, Specificity, and Accuracy at Levels of Relevance to The Theory (HBM- And TPB-Related) and Theoretical Constructs

| | **Sensitivity** | | | | | **Specificity** | | | | | **Accuracy** | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Deep learning | | | Machine learning | | Deep learning | | | Machine learning | | Deep learning | | | Machine learning | |
| | Att-RNN | BERT | Att-ELMo | TFIDF | Mean-emb | Att-RNN | BERT | Att-ELMo | TFIDF | Mean-emb | Att-RNN | BERT | Att-ELMo | TFIDF | Mean-emb |
| HBM-related | 0.8072 | 0.8193 | **0.8230** | 0.7962 | 0.8176 | **0.7954** | 0.7678 | 0.7707 | 0.7684 | 0.5378 | **0.8018** | 0.7958 | 0.7992 | 0.7835 | 0.6900 |
| HBM Susceptibility | 0.6889 | 0.8364 | 0.8152 | **0.8646** | 0.7525 | **0.9396** | 0.8331 | 0.8651 | 0.8752 | 0.7727 | **0.9015** | 0.8336 | 0.8575 | 0.8736 | 0.7696 |
| HBM Severity | 0.7620 | 0.8085 | 0.7651 | **0.8612** | 0.7264 | **0.9419** | 0.9065 | 0.9247 | 0.8967 | 0.7252 | **0.9063** | 0.8871 | 0.8931 | 0.8897 | 0.7255 |
| HBM Benefits | 0.7305 | 0.8006 | 0.6689 | **0.8152** | 0.6848 | 0.9197 | 0.8869 | **0.9375** | 0.8145 | 0.7559 | **0.8721** | 0.8652 | 0.8699 | 0.8147 | 0.7380 |
| HBM Barriers | 0.8890 | **0.9003** | 0.8922 | 0.8874 | 0.7327 | **0.9219** | 0.9099 | 0.8831 | 0.9058 | 0.7764 | **0.9063** | 0.9054 | 0.8874 | 0.8971 | 0.7557 |
| TPB-related | 0.9487 | 0.9489 | 0.9294 | **0.9526** | 0.9521 | 0.8710 | **0.8851** | 0.8591 | 0.8035 | 0.4092 | 0.9226 | **0.9274** | 0.9058 | 0.9025 | 0.7696 |
| TPB Attitude | N/A | | | | | | | | | | **0.7538** | 0.7423 | 0.7420 | 0.7261 | 0.6178 |

Note: average score from 10 experiments

# Appendix F: The Comparison of Deep Learning and Machine Learning Algorithms Measured by Precision, Recall, and F Score at Levels of Relevance to The Theory (HBM- And TPB-Related) and Theoretical Constructs

| | **Precision** | | | | | **Recall** | | | | | **F score** | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Deep learning | | | Machine learning | | Deep learning | | | Machine learning | | Deep learning | | | Machine learning | |
| | Att-RNN | BERT | Att-ELMo | TFIDF | Mean-emb | Att-RNN | BERT | Att-ELMo | TFIDF | Mean-emb | Att-RNN | BERT | Att-ELMo | TFIDF | Mean-emb |
| HBM-related | **0.8254** | 0.8084 | 0.8108 | 0.8039 | 0.6784 | 0.8072 | 0.8193 | **0.8230** | 0.7962 | 0.8176 | 0.8156 | 0.8136 | **0.8166** | 0.8000 | 0.7414 |
| HBM Susceptibility | **0.6784** | 0.4782 | 0.5229 | 0.5555 | 0.3727 | 0.6889 | 0.8364 | 0.8152 | **0.8646** | 0.7525 | **0.6805** | 0.6062 | 0.6359 | 0.6758 | 0.4981 |
| HBM Severity | **0.7681** | 0.6837 | 0.7221 | 0.6738 | 0.3951 | 0.7620 | 0.8085 | 0.7651 | **0.8612** | 0.7264 | **0.7626** | 0.7394 | 0.7393 | 0.7554 | 0.5115 |
| HBM Benefits | 0.7564 | 0.7070 | **0.7876** | 0.5968 | 0.4858 | 0.7305 | 0.8006 | 0.6689 | **0.8152** | 0.6848 | 0.7407 | **0.7487** | 0.7212 | 0.6888 | 0.5680 |
| HBM Barriers | **0.9123** | 0.9004 | 0.8734 | 0.8947 | 0.7473 | 0.8890 | **0.9003** | 0.8922 | 0.8874 | 0.7327 | 0.8999 | **0.9002** | 0.8825 | 0.8909 | 0.7396 |
| TPB-related | 0.9357 | **0.9424** | 0.9288 | 0.9055 | 0.7610 | 0.9487 | 0.9489 | 0.9294 | **0.9526** | 0.9521 | 0.9421 | **0.9455** | 0.9291 | 0.9284 | 0.8459 |
| TPB Attitude (micro-ave) | N/A | | | | | | | | | | **0.7515** | 0.7409 | 0.7399 | 0.7263 | 0.6138 |

Note: average score from 10 experiments

## Appendix G: The Comparison of Deep Learning and Machine Learning Algorithms Measured by Precision, Recall, and F-Score on TPB Attitude Classification

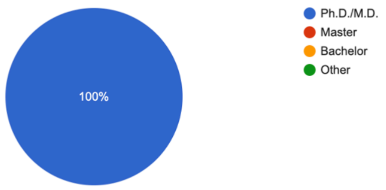
| | | Positive | | | Negative | | | Neutral | | | Micro-average F-score |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | | Precision | Recall | F-score | Precision | Recall | F-score | Precision | Recall | F-score | |
| Deep learning | Att-RNN | 0.7425 | **0.7500** | **0.7447** | 0.7987 | **0.8235** | **0.8103** | **0.7172** | 0.6843 | 0.6996 | **0.7515** |
| | BERT | 0.7470 | 0.7303 | 0.7372 | 0.7880 | 0.8048 | 0.7958 | 0.6944 | 0.6872 | 0.6898 | 0.7409 |
| | Att-ELMo | **0.7487** | 0.7171 | 0.7320 | 0.7750 | 0.7937 | 0.7830 | 0.7053 | **0.7070** | **0.7048** | 0.7399 |
| Machine learning | TFIDF | 0.7315 | 0.7132 | 0.7218 | **0.8087** | 0.7615 | 0.7838 | 0.6489 | 0.7004 | 0.6732 | 0.7263 |
| | Mean-emb | 0.6144 | 0.5745 | 0.5934 | 0.6475 | 0.7093 | 0.6768 | 0.5858 | 0.5581 | 0.5712 | 0.6138 |

Note: average score from 10 experiments

# Appendix H: Result Summary of Usability Test on Vaxinsight

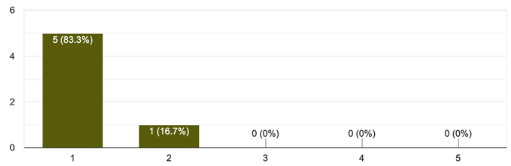
Degree (or expected degree)

6 responses



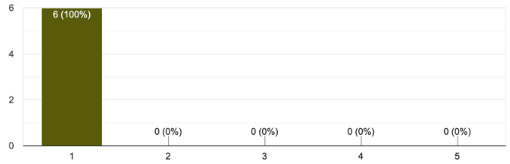
I think that I would need the support of a technical person to be able to use this system

6 responses



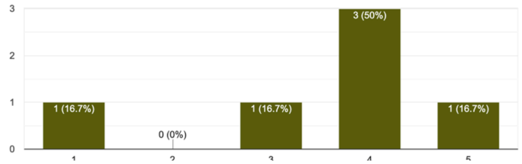
I found the system very cumbersome to use

6 responses



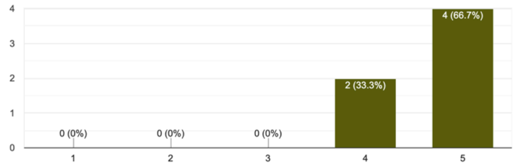
I think that I would like to use this system frequently

6 responses



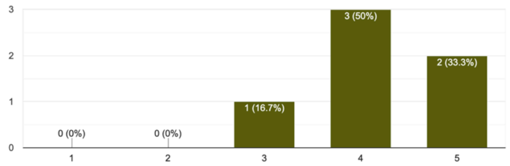
I found the various functions in this system were well integrated

6 responses



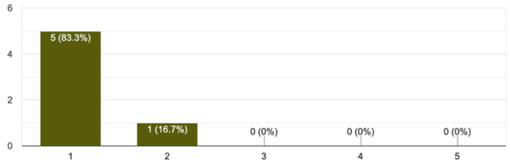
I felt very confident using the system

6 responses



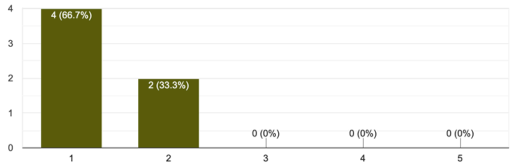
I found the system unnecessarily complex

6 responses



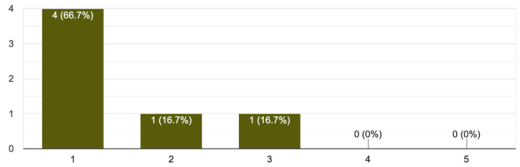
I thought there was too much inconsistency in this system

6 responses



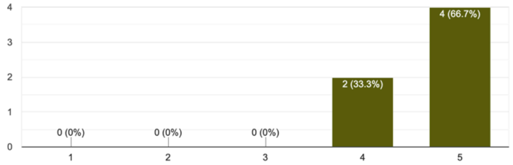
I needed to learn a lot of things before I could get going with this system

6 responses



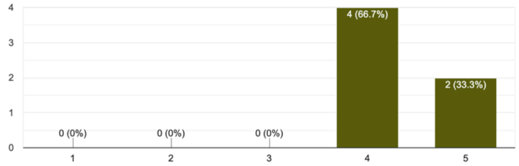
I thought the system was easy to use

6 responses



I would imagine that most people would learn to use this system very quickly

6 responses



# Appendix I: The Performance Comparison of Machine Learning and Deep Learning Algorithms on U.S. State Level Home Location Tagging on Lee Dataset

| States | F-measure | | | | | | | | | Support |
|---|---|---|---|---|---|---|---|---|---|---|
| | Machine learning | | | Convolution neural network | | | Recurrent neural network | | | |
| | LR | ET | SVM | CNN | CNN_char | CNN_POS_char | RNN | RNN_char | RNN_POS_char | |
| AK | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| AL | 0.1122 | 0.12 | 0.2606 | 0.2215 | 0.3684 | 0.3932 | 0.4211 | 0.439 | 0.3962 | 183 |
| AZ | 0.243 | 0.2519 | 0.3201 | 0.4255 | 0.4862 | 0.4531 | 0.5182 | 0.5263 | 0.5049 | 460 |
| AR | 0.0494 | 0.0488 | 0.2268 | 0.1489 | 0.26 | 0.2969 | 0.2931 | 0.3529 | 0.339 | 79 |
| CA | 0.4417 | 0.3939 | 0.4985 | 0.5637 | 0.5623 | 0.5503 | 0.5396 | 0.5408 | 0.5434 | 4289 |
| CO | 0.3554 | 0.276 | 0.3808 | 0.4763 | 0.488 | 0.5049 | 0.5371 | 0.5397 | 0.5506 | 446 |
| CT | 0 | 0.02 | 0.1897 | 0.048 | 0.1368 | 0 | 0.2969 | 0.3582 | 0.3077 | 98 |
| DC | 0.2286 | 0.0161 | 0.3498 | 0.2456 | 0.32 | 0.3235 | 0.3617 | 0.3559 | 0.3301 | 245 |
| DE | 0 | 0 | 0.0364 | 0 | 0 | 0 | 0.1429 | 0.1356 | 0.1429 | 52 |
| FL | 0.3782 | 0.2994 | 0.3223 | 0.3923 | 0.4481 | 0.4639 | 0.5109 | 0.5042 | 0.5049 | 1170 |
| GA | 0.4104 | 0.3089 | 0.3854 | 0.4416 | 0.4831 | 0.5141 | 0.5507 | 0.5282 | 0.5354 | 795 |
| HI | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| ID | 0 | 0 | 0.0741 | 0 | 0.0769 | 0 | 0.303 | 0.25 | 0.4 | 25 |
| IL | 0.3911 | 0.2585 | 0.3519 | 0.404 | 0.4483 | 0.478 | 0.5134 | 0.5175 | 0.5218 | 985 |
| IN | 0.0724 | 0.0543 | 0.2147 | 0.1544 | 0.2882 | 0.2775 | 0.3186 | 0.3434 | 0.3415 | 212 |
| IA | 0.144 | 0.2481 | 0.3462 | 0.3743 | 0.377 | 0.3918 | 0.4339 | 0.4311 | 0.3974 | 113 |
| KS | 0.0387 | 0.0886 | 0.1731 | 0.1364 | 0.2432 | 0.2601 | 0.32 | 0.3216 | 0.3468 | 149 |
| KY | 0.0896 | 0.0458 | 0.2722 | 0.1646 | 0.2857 | 0.324 | 0.3684 | 0.3478 | 0.3505 | 127 |
| LA | 0.1359 | 0.1262 | 0.296 | 0.3578 | 0.3681 | 0.4164 | 0.4259 | 0.4353 | 0.4495 | 186 |
| ME | 0 | 0 | 0.1212 | 0.0667 | 0 | 0.0625 | 0.129 | 0.2927 | 0.2286 | 29 |
| MD | 0.0717 | 0.0938 | 0.2051 | 0.1984 | 0.2572 | 0.2946 | 0.327 | 0.3452 | 0.3317 | 240 |
| MA | 0.3228 | 0.2009 | 0.3146 | 0.3515 | 0.4369 | 0.4282 | 0.4574 | 0.4812 | 0.4592 | 555 |
| MI | 0.2287 | 0.1561 | 0.2926 | 0.3694 | 0.3679 | 0.4099 | 0.4598 | 0.4595 | 0.4557 | 372 |
| MN | 0.1031 | 0.1576 | 0.3191 | 0.2569 | 0.3103 | 0.3677 | 0.1395 | 0.3927 | 0.4405 | 180 |
| MS | 0 | 0 | 0.1034 | 0 | 0 | 0 | 0.3235 | 0.1918 | 0.2258 | 51 |
| MO | 0.125 | 0.0985 | 0.2625 | 0.2277 | 0.313 | 0.3317 | 0.399 | 0.3728 | 0.381 | 250 |
| MT | 0.0606 | 0 | 0.3 | 0 | 0.1081 | 0.0606 | 0.2727 | 0.1622 | 0.2051 | 32 |
| NE | 0.0909 | 0.1449 | 0.2716 | 0.25 | 0.3462 | 0.303 | 0.3542 | 0.3393 | 0.383 | 63 |
| NV | 0.242 | 0.1515 | 0.2857 | 0.3077 | 0.3705 | 0.3672 | 0.4203 | 0.403 | 0.4208 | 358 |
| NH | 0 | 0 | 0.1053 | 0 | 0.1111 | 0.1538 | 0.1081 | 0.2083 | 0.1951 | 33 |
| NJ | 0.0406 | 0.0404 | 0.1157 | 0.0982 | 0.1888 | 0.2292 | 0.223 | 0.2158 | 0.292 | 193 |
| NM | 0.0274 | 0.08 | 0.1882 | 0.2273 | 0.4 | 0.3551 | 0.4211 | 0.4314 | 0.42 | 72 |
| NY | 0.3738 | 0.2158 | 0.353 | 0.3886 | 0.4145 | 0.4127 | 0.4396 | 0.4533 | 0.4548 | 1694 |

| | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| NC | 0.2596 | 0.2701 | 0.2698 | 0.3452 | 0.4153 | 0.4591 | 0.478 | 0.4854 | 0.487 | 601 |
| ND | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 16 |
| OH | 0.2722 | 0.2182 | 0.3019 | 0.3663 | 0.3841 | 0.432 | 0.4515 | 0.4683 | 0.4601 | 576 |
| OK | 0.1233 | 0.2358 | 0.2778 | 0.3415 | 0.3425 | 0.3704 | 0.3924 | 0.3754 | 0.3636 | 213 |
| OR | 0.2537 | 0.2431 | 0.3043 | 0.398 | 0.4466 | 0.4489 | 0.4763 | 0.5024 | 0.4762 | 399 |
| PA | 0.2388 | 0.1463 | 0.2637 | 0.3034 | 0.359 | 0.3978 | 0.4428 | 0.4629 | 0.4464 | 569 |
| RI | 0 | 0.0408 | 0.1852 | 0.0727 | 0.2 | 0.0408 | 0.481 | 0.45 | 0.4211 | 48 |
| SC | 0.2025 | 0.1957 | 0.3518 | 0.3363 | 0.3756 | 0.4088 | 0.4237 | 0.4696 | 0.4481 | 210 |
| SD | 0 | 0 | 0 | 0 | 0.1739 | 0 | 0.32 | 0.2759 | 0.1667 | 19 |
| TN | 0.1014 | 0.0291 | 0.2621 | 0.2315 | 0.3012 | 0.3828 | 0.443 | 0.4361 | 0.4149 | 202 |
| TX | 0.4138 | 0.3802 | 0.3819 | 0.4409 | 0.4687 | 0.4895 | 0.5022 | 0.5244 | 0.5069 | 1509 |
| UT | 0.1143 | 0.1135 | 0.2921 | 0.25 | 0.3111 | 0.4062 | 0.4171 | 0.4167 | 0.4038 | 131 |
| VT | 0 | 0 | 0.2632 | 0 | 0.1143 | 0.0571 | 0.3077 | 0.4898 | 0.5 | 33 |
| VA | 0.1562 | 0.1501 | 0.2111 | 0.2489 | 0.3019 | 0.3074 | 0.3531 | 0.3576 | 0.3206 | 449 |
| WA | 0.3615 | 0.2335 | 0.3461 | 0.4505 | 0.4828 | 0.4752 | 0.5115 | 0.5251 | 0.5081 | 667 |
| WV | 0.0519 | 0.1013 | 0.3656 | 0.2022 | 0.422 | 0.4348 | 0.512 | 0.4874 | 0.4865 | 75 |
| WI | 0.1299 | 0.1217 | 0.2789 | 0.2195 | 0.2848 | 0.2828 | 0.3708 | 0.3729 | 0.3929 | 212 |
| WY | 0 | 0 | 0.087 | 0 | 0 | 0 | 0 | 0.0909 | 0 | 21 |

Note: measured by F score

## Appendix J: The Performance Comparison of Transfer Learning on U.S. State Level Home Location Tagging on HPV Dataset

| States | F-measure | | | Support |
|---|---|---|---|---|
| | Pre-trained | Training from scratch | Fine-tuned | |
| AK | 0 | 0.4533 | 0.3881 | 51 |
| AL | 0.4566 | 0.4557 | 0.4444 | 141 |
| AZ | 0.4444 | 0.487 | 0.4786 | 270 |
| AR | 0.3604 | 0.4333 | 0.4299 | 76 |
| CA | 0.4196 | 0.4789 | 0.4575 | 1921 |
| CO | 0.417 | 0.453 | 0.4535 | 270 |
| CT | 0.2889 | 0.4609 | 0.4655 | 141 |
| DC | 0.1927 | 0.3225 | 0.3605 | 335 |
| DE | 0.1538 | 0.3415 | 0.375 | 23 |
| FL | 0.4366 | 0.4098 | 0.4619 | 831 |
| GA | 0.4735 | 0.5055 | 0.5239 | 449 |
| HI | 0 | 0.4828 | 0.525 | 51 |
| ID | 0.3448 | 0.5714 | 0.5429 | 42 |
| IL | 0.4646 | 0.4466 | 0.5174 | 574 |
| IN | 0.4784 | 0.4754 | 0.5066 | 215 |
| IA | 0.4 | 0.4835 | 0.5 | 117 |
| KS | 0.5226 | 0.5368 | 0.5226 | 107 |
| KY | 0.4236 | 0.4494 | 0.4745 | 156 |
| LA | 0.4131 | 0.4854 | 0.5066 | 179 |
| ME | 0.4364 | 0.4375 | 0.4762 | 40 |
| MD | 0.3473 | 0.4293 | 0.4076 | 355 |
| MA | 0.4832 | 0.5148 | 0.5046 | 465 |
| MI | 0.4348 | 0.4725 | 0.4794 | 318 |
| MN | 0.5235 | 0.5351 | 0.5604 | 217 |
| MS | 0.2637 | 0.4228 | 0.3788 | 68 |
| MO | 0.2845 | 0.4186 | 0.4195 | 179 |
| MT | 0.4151 | 0.4643 | 0.4364 | 39 |
| NE | 0.547 | 0.5769 | 0.5614 | 63 |
| NV | 0.3844 | 0.4646 | 0.4397 | 159 |
| NH | 0.24 | 0.4348 | 0.4474 | 43 |
| NJ | 0.2149 | 0.367 | 0.3822 | 312 |
| NM | 0.4112 | 0.3962 | 0.4423 | 71 |

| | | | | |
|---|---|---|---|---|
| NY | 0.3693 | 0.4512 | 0.4632 | 1202 |
| NC | 0.4426 | 0.462 | 0.4984 | 413 |
| ND | 0.1818 | 0.375 | 0.3922 | 36 |
| OH | 0.5462 | 0.5266 | 0.5356 | 448 |
| OK | 0.4845 | 0.5246 | 0.5217 | 106 |
| OR | 0.4812 | 0.4835 | 0.5187 | 215 |
| PA | 0.4547 | 0.449 | 0.4838 | 537 |
| RI | 0.4286 | 0.5625 | 0.5984 | 72 |
| SC | 0.4514 | 0.4714 | 0.5157 | 172 |
| SD | 0.1538 | 0.2581 | 0.2667 | 23 |
| TN | 0.4157 | 0.4222 | 0.4764 | 230 |
| TX | 0.5663 | 0.5221 | 0.5901 | 1297 |
| UT | 0.4815 | 0.5375 | 0.5799 | 104 |
| VT | 0.4255 | 0.5172 | 0.5424 | 36 |
| VA | 0.3648 | 0.4041 | 0.4206 | 259 |
| WA | 0.4625 | 0.533 | 0.5642 | 492 |
| WV | 0.4828 | 0.5 | 0.5111 | 60 |
| WI | 0.4164 | 0.4397 | 0.4721 | 182 |
| WY | 0.3333 | 0.5 | 0.4286 | 10 |

Note: measured by F score

# VaxInsight: an artificial intelligence system to access large-scale public perceptions of vaccination from social media

By

Jingcheng Du, B.S.

APPROVED:

**Cui Tao, PhD, Chair**

**Hua Xu, PhD**

**Yong Chen, PhD**

M. Sahiti

**Sahiti Myneni, PhD**

**Trevor Cohen, PhD**

Date approved: 12-13-19