Exploration and Initial Development of Text Classification Models to Identify Health Information Technology Usability-Related Patient Safety Event Reports

Allan Fong[1] Tomilayo Komolafe[2] Katharine T. Adams[1] Arman Cohen[3] Jessica L. Howe[1] Raj M. Ratwani[1,4]

[1] National Center for Human Factors in Healthcare, Washington, District of Columbia, United States
[2] Virginia Polytechnic Institute and State University, Blacksburg, Virginia, United States
[3] Allen Institute for Artificial Intelligence, Seattle, Washington, United States
[4] Georgetown University Medical Center, Washington, District of Columbia, United States

Address for correspondence Allan Fong, MS, National Center for Human Factors in Healthcare, 3007 Tilden Street Northwest, Suite 7L, Washington, DC 20008, United States (e-mail: allan.fong@medstar.net).

Appl Clin Inform 2019;10:521–527.

Abstract

Background With the pervasive use of health information technology (HIT) there has been increased concern over the usability and safety of this technology. Identifying HIT usability and safety hazards, mitigating those hazards to prevent patient harm, and using this knowledge to improve future HIT systems are critical to advancing health care.

Purpose The purpose of this work is to demonstrate the feasibility of a modeling approach to identify HIT usability-related patient safety events (PSEs) from the free-text of safety reports and the utility of such models for supporting patient safety analysts in their analysis of event data.

Methods We evaluated three feature representations (bag-of-words [BOWs], topic modeling, and document embeddings) to classify HIT usability-related PSE reports using 5,911 manually annotated reports. Model results were reviewed with patient safety analysts to gather feedback on their usefulness and integration into workflow.

Results The combination of term frequency-inverse document frequency BOWs and document embedding features modeled with support vector machine (SVM) with radial basis function (RBF) had the highest overall precision-recall area under the curve (AUC) and f1-score, 72 and 66%, respectively. Using only document embedding features achieved a similar precision-recall AUC and f1-score performance with the SVM RBF model, 70 and 66%, respectively. Models generally favored specificity and sensitivity over precision. Patient safety analysts found the model results to be useful and offered three suggestions on how it can be integrated into their workflow at the point of report entry, in a visual dashboard layer, and to support data retrievals.

Conclusion Text mining and document embeddings can support identification of HIT usability-related PSE reports. The positive feedback received on the HIT usability model shows its potential utility in real-world applications.

Keywords

- natural language processing
- interfaces and usability
- incident reporting
- electronic health records and systems
- safety

received
March 4, 2019
accepted after revision
June 6, 2019

DOI https://doi.org/10.1055/s-0039-1693427.
ISSN 1869-0327.

Background and Significance

The widespread use of health information technology (HIT), including electronic health records (EHRs) has improved certain aspects of patient care, but has also resulted in unintended safety consequences. Many of these safety hazards are associated with the usability of the technology and these hazards can lead to patient harm or even death.[1,2] Improving patient safety is a top priority for nearly every health care provider organization. Unsafe care leads to patient harm and unnecessary cost.

Identifying the specific HIT usability and safety hazards associated with patient harm can be challenging given that HIT is now intertwined with the care delivery process and many hazards that may be associated with HIT may not be easily related back to the HIT system. For example, a physician may make an error and enter the incorrect medication dose when ordering through the EHR, because of a confusing display, and that error may be later caught by the nurse attempting to administer the medication. This near-miss event is often documented as a medication error with no mention of the HIT system that may be associated with the safety event. There has been limited published research on identifying HIT usability-related safety issues despite usability being a major concern, as acknowledged by the Institute of Medicine.[3] Given these concerns, we focus on identifying HIT usability-related safety event reports from large databases, evaluating different feature selection techniques.

HIT Usability in Patient Safety Event Reports

One method to identify HIT-related safety events, although incomplete, is to analyze patient safety event (PSE) report data. Nearly every health care system collects PSE reports which are entered by frontline clinicians and are composed of both structured data, such as the event category (e.g., medication, fall) and severity level (e.g., near miss, reached patient, reached patient with harm), as well as an unstructured free-text that provides a description of the event with possible contributing factors. These reports hold tremendous promise for identifying safety hazards and then later monitoring whether a safety risk has been mitigated. However, there are three major challenges that exist when attempting to identify HIT usability-related safety events. First, many frontline clinicians do not categorize HIT-related events in the HIT structured category making it difficult to identify the reports that are actually HIT related. Second, with provider organizations accumulating tens of thousands of these reports it is increasingly difficult for patient safety analysts, who are often responsible for the integration, analysis, dissemination, and management PSE reports,[4] to read each report to determine whether it is related to HIT usability and safety. Third, HIT and usability issues are generally underreported relative to other types of events.

For example, recent research focused on medication events has identified that a large portion of medication events (over 50%) can be attributed to computerized provider order entry (CPOE).[5]

Researchers have explored the use of machine learning and natural language processing (NLP) to analyze the free-text of PSE reports.[6–8] Chai et al developed NLP models to classify HIT events using the Manufacturer and User Facility Device Experience data set made public by the Food and Drug Administration in the United States.[7] PSE reports have also been used in the development of models to categorize tasks associated with HIT such as inadequate handover, incorrect patient identification, and medication errors.[6,8] Building upon this work, we seek to develop models, using different topic modeling, and word embedding to classify PSE reports as being related or unrelated to HIT usability. We then applied the models to classify actual PSE data, presented these results to patient safety analysts, and interviewed them to determine the utility of the algorithms in their workflows.

Topic Modeling and Document Embeddings

Topic modeling and document embeddings are two techniques that extend beyond the standard bag-of-words (BOWs) feature representation and have the potential to capture a more robust representation of terms and phrases.[9–14] Similar to other textual data, PSE reports can have latent topics.[15] These topics may provide additional signals that could be used for better classification. For example, a PSE report classified as a “fall” by a frontline reporter could include latent medication and provider fatigue themes or topics in the free-text. Topic modeling[14] is an unsupervised method to detect these topics, which can be used as additional information for classifying events. Previous work has proved the effectiveness of topic modeling for text classification tasks compared with BOW features.[11–13] Latent Dirichlet allocation (LDA)[16] is a popular generative topic model that discovers topics in textual documents. To improve identification of usability hazards from PSE reports, LDA can be used to generate unsupervised topics to capture the themes of each PSE report.

In addition, existing corpora can be leveraged to train document-level embeddings[17] an extension of word embeddings[9,10] to learn word sequences in paragraphs and documents. Document embeddings is an unsupervised algorithm that learns word sequence features from variable-length documents or reports resulting in dense vector representations of documents.[17] The interesting property of these vectors compared with sparse BOWs features is that they can capture semantic and syntactic relatedness between the words and word sequences in a dense representation. Therefore, document embeddings could represent the dependencies and interactions between word sequences across documents. A recent comparison demonstrated the robust performance of document embeddings when trained on large corpora.[18]

Integrating Machine Learning into Patient Safety Analyst Workflow

The scope and responsibilities of patient safety analysts are rapidly changing as the volume and frequency of reports grow and the need to identify HIT-related events becomes more pressing. With few tools to support their analysis, the cognitive burden on patient safety analysts is increasing.[4] Machine learning models can be integrated to support patient safety analysts; however, the successful integration of machine learning models into real-world health care workflow requires both valid models and proper user-centered integration into existing work practices.[19] Many models and decision support tools fail after transitioning to a real-world health care application due to a variety of reasons including poor workflow or social integration, and a lack of concern for end-user needs.[19] These challenges highlight the importance of understanding the needs and workflow processes of patient safety analysts prior to implementation of machine learning models in real-world health care settings. After developing the HIT usability models, we explored how these models could be effectively used by patient safety analysts by soliciting their feedback on how the models could be integrated to best support their work.

Methods

This study consists of two parts, the development of a HIT usability classification model (Part 1) and the pilot testing of the model with patient safety analysts (Part 2). PSE reports were manually annotated as being likely or unlikely related to usability (►Fig. 1). We then asked patient safety analysts for their feedback on model results applied to their specific data as well as its overall utility in their workflow. This study was approved by the MedStar Health Research Institute Institutional Review Board (protocol #2014-101).

HIT Usability Classification Model

Data Source

Our data source contains deidentified PSE reports from the Pennsylvania Patient Safety Authority's Patient Safety Reporting System and from a large health care system based in the Mid-Atlantic of the United States. The initial data set consists of 1.735 million reports entered between January 2009 and June 2016. We sampled a subset of this data set by selecting all PSE reports indicated as being related to HIT by the reporter at the time of report entry (2,635 out of 1.735 million) and randomly selecting 3,385 PSE reports without such HIT indication. This sampling approach for training and testing was used because the percentage of HIT-related reports in PSE data sets are generally low and similar approach has resulted in good model performance in related prediction tasks.[7]

HIT Usability

Duplicate reports and reports with missing free-text were removed resulting in 5,911 (out of 6,020) reports for expert coding (►Fig. 1). A team of five expert annotators (one physician, one pharmacist, three human factors experts with extensive experience in HIT and usability) first coded the 5,911 into likely HIT or unlikely HIT or need more information. Three hundred fifty reports were first coded individually by all annotators in four separate rounds. An interrater reliability of 0.786 using Fleiss' kappa was achieved before individual coding of the remaining reports. 2,435 reports were identified as likely HIT related and was further categorized into "Likely" or "Unlikely" related to poor HIT usability, which is the extent to which HIT support clinicians in achieving their goals in a satisfying, effective, and efficient manner, was a contributing factor in the report.[3] HIT usability has been categorized to include seven components: data entry, alerting, interoperability, visual display, availability of information, system automation and defaults, and workflow support.[1] "Unlikely" HIT usability reports are those events in which the usability of the system was not likely a contributing factor to the event or the event description did not provide enough information to clearly determine a link to HIT usability. ►Table 1 provides examples of "Likely" and "Unlikely" HIT usability-related reports.

The three human factor experts manually coded all 2,435 HIT reports as being "Likely" or "Unlikely" related to usability with 75% complete agreement. Disagreements were reconciled through group discussion and resulting in a final coding of 982 "Likely" (40% of 2,435) HIT usability-related reports. This ratio is comparable in magnitude to related research showing usability issues as a contributing factor to 63.9% of EHR and medication-related safety events.[2] Reports requiring additional clinical context were discussed with a physician prior to final coding.

Feature Selection

Three feature generation methods were evaluated: BOWs, topic modeling using LDA, and document embedding trained

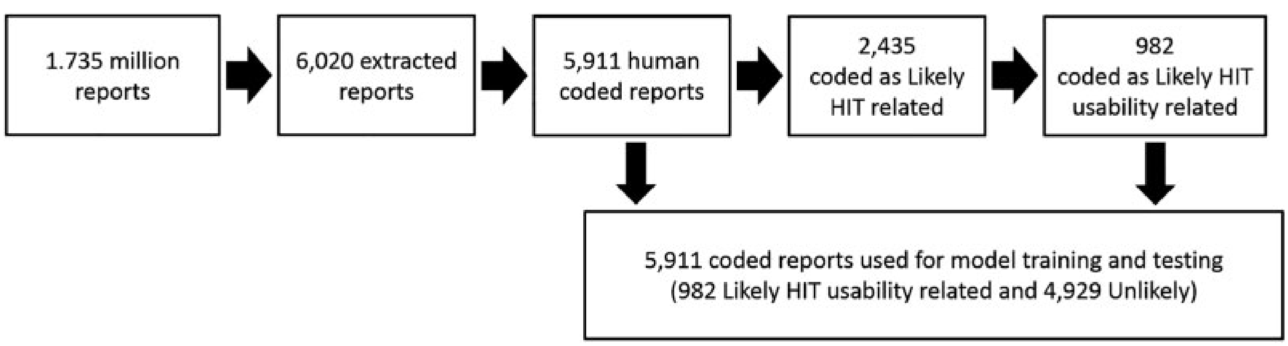

Fig. 1 Report coding workflow for developing the model training and testing data sets.

Table 1 Examples of "Likely" and "Unlikely" HIT usability-related patient safety events

Likely HIT usability related	Unlikely HIT usability related
"…Medication ordered in CPOE, that order does not include all needed data points…" "…Confusing EHR system when prescribing medication. In EHR, I ordered medication 20 mg. I wanted the child to take 3 tabs once daily for 5 days. When I entered 3 tabs it converted it to 60 mg in prescription bar. This made it confusing when decided the quantity to prescribe…"	"…Unable to print reports from computer in pre-admission testing office. Unable to find printer in directory. Information Systems help desk notified. System was dropping printer unexpectedly…" "…No medications were entered into EHR medication history by the 5 PM huddle for patient. This creates limitations on discussing the patient in his entirety…"

Abbreviations: CPOE, computerized provider order entry; EHR, electronic health record; HIT, health information technology.

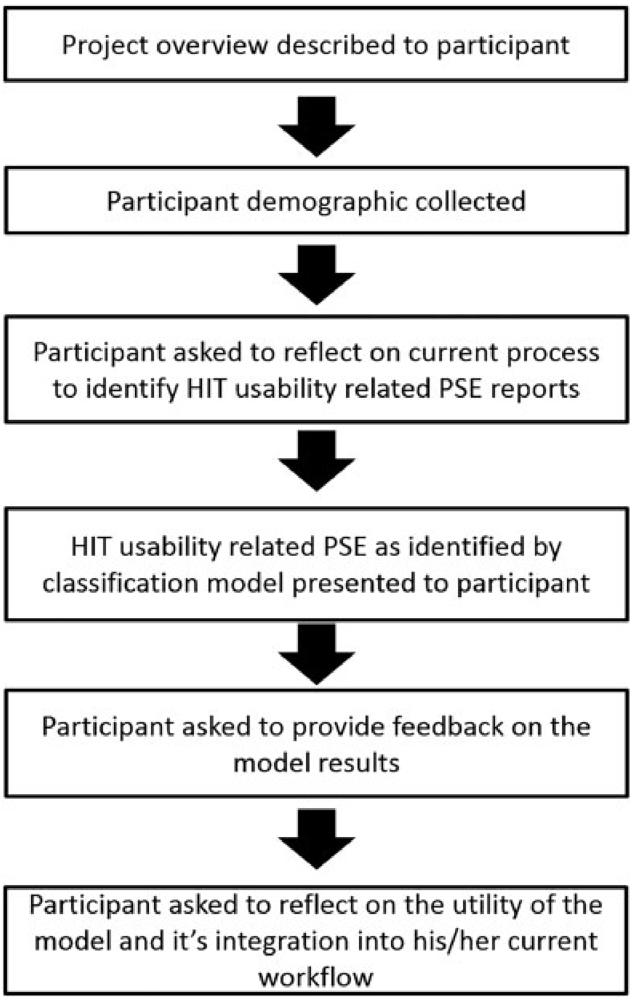

Fig. 2 The flow diagram for each patient safety analyst interview.

with the PSE reports (D2V). In addition, we evaluated the combination of the three feature types.

Bag-of-Words

In order to classify PSE reports, we represented each report as a feature vector. To do so, we used BOW features weighted using the token frequency according to term frequency–inverse document frequency.

Topic Modeling

We trained LDA topic models on the set of PSE reports. We found 700 topics to be sufficient for this task. We used these models to transform the PSE reports into its topic representation. We used the Gensim (https://pypi.org/project/gensim/) tool for training the topic models.

Document Embedding

We trained document embeddings using the unannotated corpus of 1.729 million PSE reports. We used the doc2vec implementation of document embedding generation[18] following the recommendations for hyperparameter settings. The vector size was set at 500 and the window size set to 15. The threshold to down sample high-frequency words was set at 10^{-5} and the number of negative word samples was set to 5. The minimum frequency threshold for word types was 5 and we used 400 training epochs. Using the training doc2vec, we infer an embedding for each of the 5,911 annotated reports. This approach results in a document vector of a PSE report that captures the embedding representations averaged across occurring words and word sequences.

Training and Testing

A total of 5,911 reports were used for training and testing with 20% of the reports reserved for testing. Classifiers were trained using support vector machines (SVMs) with radial basis function (RBF) and linear kernels. Gamma and cost parameters were optimized using grid search on the training data for each respective classification model. SVM has been shown to have high performance in similar classification tasks with PSE reports.[6,7] Each model was evaluated on specificity, sensitivity, precision, f1-score, and precision-recall area under the curve (AUC) using fivefold cross-validation.

Usability Model Application

To evaluate the utility of our model in a real-world context, we interviewed and surveyed 12 patient safety analysts from a multihospital health care system with an average of 4 years of experience in their current patient safety role (range 2–10 years). The flow of each interview is summarized in ►**Fig. 2**. Participants were first asked to describe their current process of identifying HIT usability-related PSE reports. We then applied the model from Part 1 with the best precision-recall AUC and high specificity to their PSE reports extracted from January 2017 to December 2017. We selected 10 reports identified by the model as likely HIT usability related excluding incomplete reports, open reports, and reports with more

Table 2 Specificity, sensitivity/recall, precision, f1-scores, and precision-recall area under the curve (AUC) test results for each feature and classification algorithm

Features	Algorithm	Specificity	Sensitivity	Precision	F1-score	AUC
BOW	SVM Linear	87%	73%	52%	61%	64%
	SVM RBF	88%	73%	55%	63%	65%
LDA	SVM Linear	87%	61%	49%	54%	55%
	SVM RBF	95%	38%	59%	46%	55%
D2V	SVM Linear	86%	77%	52%	62%	67%
	SVM RBF	88%	79%	56%	66%	70%
BOW + LDA	SVM Linear	87%	68%	52%	59%	64%
	SVM RBF	88%	72%	55%	63%	66%
BOW + D2V	SVM Linear	88%	70%	53%	61%	65%
	SVM RBF	89%	76%	59%	66%	72%
LDA + D2V	SVM Linear	87%	74%	53%	62%	66%
	SVM RBF	88%	78%	57%	66%	70%
BOW + LDA + D2V	SVM Linear	88%	68%	54%	60%	65%
	SVM RBF	90%	74%	59%	65%	72%

Abbreviations: BOW, bag-of-word; LDA, latent Dirichlet allocation; RBF, radial basis function; SVM, support vector machine.

than 1,000 words. Participants were asked to provide feedback and overall impressions of the model-processed reports. Lastly, patient safety analysts were asked to comment on the utility of the model as well as how it could be integrated into their workflow.

Results

HIT Usability Model

BOW + D2V had the best precision-recall AUC and f1-score, 72 and 66% using SVM with RBF (►Table 2). Using only D2V features achieved a similar precision-recall AUC and f1-score performance with the SVM RBF model, 70 and 66%, respectively. Models generally favored specificity and sensitivity (or recall) over precision especially for SVM with RBF. LDA had the highest specificity using SVM RBF, 95%, but poor sensitivity, 38%.

In addition, we inspected BOW feature importance ranked by their coefficients in the BOW + D2V linear SVM (►Table 3). Linear SVM results are comparable to RBF SVM with more interpretable coefficient rankings. A combination of HIT systems nouns (e.g., emr, cpoe, and emar) and action verbs (e.g., defaulted, deleted, clicked) were highly ranked features.

Table 3 Top 20 ranked BOW features in the linear SVM using both BOW and D2V features

Top ranked BOW features
emr
cross
duplicate
order
discontinue
cpoe
enter
defaulted
initiate
set
emar
hold
comments
taper
indicate
communication
route
deleted
timed
clicked

Abbreviations: BOW, bag-of-word; SVM, support vector machine.

Patient Safety Analyst Feedback

When describing their current process for identifying safety events that may be HIT and usability related, patient safety analysts said they rely on the free-text narratives of the PSE reports or structured categories. One participant indicated the use of a text search tool to identify reports with keywords. Nine out of 12 participants did not think their current process gave them a good representation of the types of HIT usability events. "There are likely many [HIT usability related] events falling through the cracks that we are not aware of" [P2]. After reviewing the 10 model-identified PSE reports, all participants agreed that the model results identified reports that are generally more fitting for the topic of HIT usability then the reports' structured general categories assigned by the reporters (e.g., medication, diagnosis treatment, imaging). Participants were all very receptive of the model. "This model provides the opportunity to reframe [HIT usability] as a system-based issued" [P9].

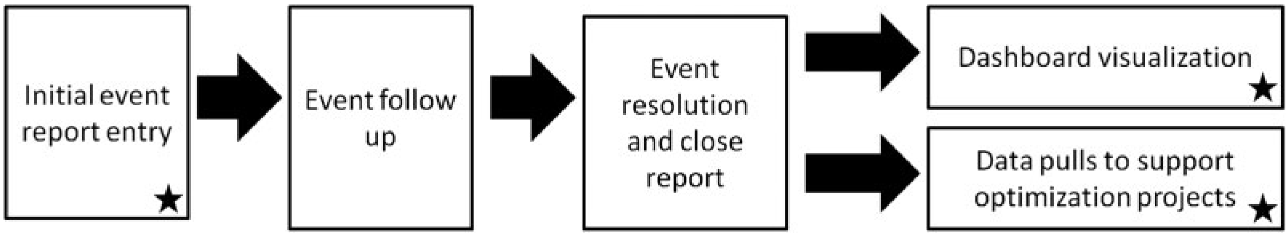

Fig. 3 Stars indicate model integration opportunities in the workflow of patient safety analysts.

Ten out of 12 participants said they would integrate the model into their workflow and provided three main suggestions (►Fig. 3). The most common suggestion (5/10) was to integrate the model classification into a dashboard they are currently using to analyze and trend PSE reports. “[I] would love a dashboard to help highlight HIT issues and share with our system safety groups and informatics leaders” [P6]. Another suggestion was to integrate the model at the point of reporter entry as a mechanism to recommend to reporters which categories to choose. This suggestion resembles a validator functionality recommended by previous studies.[20] In addition, suggestions were made to integrate the model when pulling and reviewing data for different patient safety committees or when working on HIT system optimization projects. “I would like to see the model used at initial data pull” [P5]. Lastly, feedback was given on how this model could be made more useful if the outputs were also tied to “substantive actions” or recommendations that could be taken by the safety team.

Discussion

Of the different feature selections examined, BOW + D2V performed well and demonstrates that machine learning approaches can be used to identify HIT usability-related safety events from large databases of PSE reports. Top ranked action verbs highlight areas of user interactions (e.g., clicked, initiate), workflow communication and interoperability (e.g., comments, communication, indicate, cross), and system automation (e.g., defaulted, timed) that could be more associated with HIT usability challenges. These results also demonstrate the utility of using document embeddings trained on a large corpus of 1.7 million PSE reports to infer dense document vectors for learning, suggesting additional value for continued sharing and learning from PSE reports across health care systems and states.

Given the recognized difficulty in identifying safety events that may have usability as a contributing factor, a more precise model has the potential to support the identification of HIT usability challenges and with this knowledge HIT system usability can be improved to reduce or eliminate these safety hazards. The HIT usability model developed can be beneficial for both health care provider organizations as well as organizations that are charged with collecting and analyzing large databases of PSE reports from multiple institutions, such as patient safety organizations. The Institute of Medicine, the Joint Commission, and other organizations highlight the safety challenges associated with HIT and it is critical that new methods, like the models developed here, are put to use to better identify and mitigate HIT safety hazards.

Workflow Integration

Patient safety analysts agreed that the HIT usability model provided a good representation of the types of HIT usability events and that it would be beneficial to integrate the model into their workflow. It was clear from discussions that for successful adoption the model would have to be discreetly embedded in their current workflow at natural touch points. The three most common suggestions for model integration are at the point of frontline report entry, at the dashboard visualization layer, or during data retrieval. The first two points of integration might benefit more with models tune for higher precision while the data retrieval task might benefit more from higher recall model results. Nevertheless, determining the feasibility of front-end integrating a real-time model to work in the reporting process and evaluating reporter use of this type of feature requires future research. The use of the algorithms includes the potential benefit of identifying events that may be HIT related which would otherwise not be categorized as such, and the benefit of increased efficiency from reducing the amount of manual coding that would otherwise be required. Given the demands placed on patient safety analysts with increasing amounts of data to analyze, the efficiencies are of major benefit.[4]

Limitations

Despite including data from multiple health care systems and states, future work to apply and evaluate a HIT usability model across health care systems will be useful. Furthermore, our model was developed from reports previously identified as HIT related. It would be useful to expand the modeling to include all types of annotated PSE reports, combine multiple features, and explore boosting and bagging modeling techniques to improve the model's overall performance. In addition, it is important to consider how reporters' behavior and free-text language might change after the implementation and/or knowledge of a classification algorithm to analyze reports in real time. This can make reporters more aware of HIT and usability issues which can impact the language used to describe such events. It will be important to periodicity check the performance of any implemented model and update as appropriate. Lastly, the interviews only occurred at one health care system. It would be important to expand this work to more health care systems to capture more patient safety analyst input.

Conclusion

HIT usability can contribute to safety events; however, identifying safety reports that are HIT usability related is challenging. We demonstrated the utility of text mining to identify HIT usability-related PSE reports. Patient safety analysts found the results of the modeling approach powerful and identified ways to better integrate the model with their work practices.

Clinical Relevance Statement

Although document embeddings can be more difficult to implement compared to bag-of-word approaches, document embeddings can capture word sequences that can be used in predictive analytics.

Multiple Choice Question

1. When modeling health information technology usability-related patient safety events, rank the following features in order of decreasing test f1-score when using support vector machine with radial basis function:
   a. D2V > LDA > BOW.
   b. BOW > LDA > D2V.
   c. BOW > D2V > LDA.
   d. D2V > BOW > LDA.

   Correct Answer: The correct answer is option D (D2V > BOW > LDA). Our analysis of health information technology usability-related patient safety event reports demonstrates the utility of document embeddings (D2V) features especially when trained on a large corpus.

Note

These results are the opinions of MedStar Health researchers and do not reflect in any way an analysis or opinions of the Pennsylvania Patient Safety Authority (the "Authority"). This analysis was not prepared by the Authority. This analysis was conducted by researchers from MedStar Health. Neither the Authority nor its agents, and staff bear any responsibility or liability for the results of MedStar Health's analysis, which are solely the opinion of MedStar Health. The opinions expressed in this document are those of the authors and do not necessarily reflect the official position of the Agency for Healthcare Research and Quality or the U.S. Department of Health and Human Services.

Protection of Human and Animal Subjects

This study was approved by the MedStar Health Research Institute Institutional Review Board (protocol #2014-101).

Conflict of Interest

A.F., R.M.R., J.L.H., K.T.A., T.K. report grants from Agency for Healthcare Research and Quality, during the conduct of the study. All other authors declare they have no conflicts of interest related to this study.

References

1 Howe JL, Adams KT, Hettinger AZ, Ratwani RM. Electronic health record usability issues and potential contribution to patient harm. JAMA 2018;319(12):1276–1278

2 Ratwani RM, Savage E, Will A, et al. Identifying electronic health record usability and safety challenges in pediatric settings. Health Aff (Millwood) 2018;37(11):1752–1759

3 Institute of Medicine. Health IT and Patient Safety Building Safer Systems for Better Care. Washington, DC: National Academies Press; 2012

4 Puthumana JS, Fong A, Blumenthal J, Ratwani RM. Making patient safety event data actionable: understanding patient safety analyst needs. J Patient Saf 2017. Doi: 10.1097/PTS.0000000000000400

5 Amato MG, Salazar A, Hickman T-TT, et al. Computerized prescriber order entry–related patient safety reports: analysis of 2522 medication errors. J Am Med Inform Assoc 2017;24(02):316–322

6 Fong A, Harriott N, Walters DM, Foley H, Morrissey R, Ratwani RR. Integrating natural language processing expertise with patient safety event review committees to improve the analysis of medication events. Int J Med Inform 2017;104:120–125

7 Chai KEK, Anthony S, Coiera E, Magrabi F. Using statistical text classification to identify health information technology incidents. J Am Med Inform Assoc 2013;20(05):980–985

8 Ong M-S, Magrabi F, Coiera E. Automated categorisation of clinical incident reports using statistical text classification. Qual Saf Health Care 2010;19(06):e55

9 Pennington J, Socher R, Manning CD. GloVe: Global Vectors for Word Representation. In: Conference on Empirical Methods in Natural Language Processing; 2014:1532–1543

10 Mikolov T, Chen K, Corrado G, Dean J. Distributed Representations of Words and Phrases and their Compositionality. In: Proceedings of the Advances in Neural Information Processing Systems; 2013:3111–3119

11 Chen M, Jin X, Shen D. Short Text Classification Improved by Learning Multi-Granularity Topics. In: International Joint Conference on Artificial Intelligence; 2011:1776–1781

12 Aggarwal CC, Zhai C. A Survey of Text Classification Algorithms. Mining Text Data. Boston, MA: Springer; 2012:163–222

13 Wang Z, Qian X. Text Categorization Based on LDA and SVM. In: International Conference on Computer Science and Software Engineering; 2008:674–677

14 Blei DM. Probabilistic topic models. Commun ACM 2012;55(04):77–84

15 Fong A, Ratwani R. An evaluation of patient safety event report categories using unsupervised topic modeling. Methods Inf Med 2015;54(04):338–345

16 Blei D, Ng A, Jordan M. Latent Dirichlet allocation. J Mach Learn Res 2003;3:993–1022

17 Le Q, Mikolov T. Distributed representations of sentences and documents. In: International Conference on Machine Learning; 2014:1188–1196

18 J.H. Lau, T. Baldwin, An empirical evaluation of doc2vec with practical insights into document embedding generation, ArXiv Prepr. ArXiv1607.05368; 2016

19 Yang Q, Zimmerman J, Steinfeld A, Carey L, Antaki JF. Investigating the heart pump implant decision process: opportunities for decision support tools to help. ACM Trans Comput Hum Interact 2016;2016:4477–4488

20 Gong Y, Kang H, Wu X, Hua L. Enhancing patient safety event reporting. A systematic review of system design features. Appl Clin Inform 2017;8(03):893–909