Natural Language Processing of Nursing Notes: A Systematic Review

Shazia MITHA[a] , Jessica SCHWARTZ[a], Kenrick CATO[a], Kyungmi WOO[a], Arlene SMALDONE[a] and Maxim TOPAZ[a]
[a] *Columbia University School of Nursing*

1. Introduction

Natural language processing (NLP) is a method that originates in computer science and helps to find meaning in free-text data. NLP is well established for some health data, especially medical data (e.g., radiology and pathology clinical notes). In nursing, NLP have the potential to help with automated data extraction from nursing notes, enabling a diverse range of prediction and clinical decision support tasks. However, the extent of use of NLP in nursing remains unknown. We conducted a systematic review of literature to understand how NLP was applied on nursing data.

2. Methods

We searched PubMed and EMBASE to identify all potentially relevant abstracts related to NLP of nursing notes. We limited our results to articles in English language, without date constraints. Articles were included if they focused on development or implementation of NLP using data generated by nurses (e.g., inpatient or outpatient clinical notes). After excluding duplicates, 234 studies were selected for initial review. After article abstract review, 32 studies met all initial inclusion criteria. A total of 19 articles were included in our final review after independent review by four expert reviewers. For each article, we extracted data related to the study purpose, corpus (e.g., data source, number of narratives), patients (e.g., target population, number of distinct patients), NLP methods (e.g., methodology and/or tools used, performance metrics, standard terminologies used).

3. Results

The majority of the studies (70%) were published in the last four years (2015-19). Most of the studies were conducted using either inpatient (60%) or home health (15%) data. Five studies (25%) used a publically available database of clinical data called Medical Information Mart for Intensive Care (MIMIC). Most common standard vocabularies used were the Unified Medical Language System (UMLS, 50%) and Systematized Nomenclature of Medicine (SNOMED-CT, 40%). Nursing standard terminologies (e.g.,

the International Classification for Nursing Practice, ICNP) were used only in one quarter of the studies. NLP was used to on a variety of topics with most common being: cardiac symptoms (n=4), mortality risk (n=4), falls risk (n=2). Predictive performance metrics (e.g., F-score, sensitivity and specificity) were reported for only 30% of the studies. The majority of studies were published in biomedical informatics journals.

4. Discussion

Although there was a significant increase in the number of published NLP studies in the recent years, the overall number of studies remains relatively small. Vast majority of studies were conducted with hospital or homecare data while little is known about NLP applicability to other settings, such as nursing homes or skilled nursing facilities. One publicly available datasets was frequently used; one in four NLP articles analyzed data from this database. Standard nursing terminologies were not generally applied in nursing NLP studies. NLP with nursing data was conducted on a variety of patient and nursing-related factors, such as symptoms or prediction of patient risk. Only one-third of the studies reported NLP system performance which limits our understanding of NLP applicability in nursing.

5. Conclusions

This systematic review identified a growing trend of NLP with nursing data. However, only one-third of the studies reported NLP system performance and we encourage further NLP projects to use appropriate metrics (e.g., F-score) when reporting results. In addition, nursing NLP projects are encouraged to use exiting standard nursing terminologies to enable future scalability of the methods. Finally, more evidence is needed to understand the applicability of NLP beyond hospital or homecare setting.

References

[1]   Bjarnadottir RI, Lucero RJ. What can we learn about fall risk factors from EHR Nursing notes? A text mining study. 6 (2018) 1–8.
[2]   Bjarnadottir R. Nurse documentation of sexual orientation and gender identity in home healthcare, 37 (2019) 213–221. doi:10.1097/CIN.0000000000000492.
[3]   Conway M, Keyhani S, Christensen L et al. Moonstone: A novel natural language processing system for inferring social risk from clinical narratives. 0 (2019) 1–10.
[4]   Gundlapalli AV, Divita G, Redd A et al. Detecting the presence of an indwelling urinary catheter and urinary symptoms in hospitalized patients using natural language processing. J. Biomed. Inform. 71 (2017) S39–S45. doi: 10.1016/j.jbi.2016.07.012
[5]   Hajihashemi Z, Popescu M. An early illness recognition framework using a temporal smith waterman algorithm and NLP. (n.d.) 548–557.
[6]   Hyun S, Johnson SB, Bakken S. HHS public access, 27 (2015) 215–225. doi: 10.1097/NCN.0b013e3181a91b58.Exploring
[7]   Lehman L, Saeed M, Long W, Lee J, Mark R. Risk stratification of ICU patients using topic models inferred from unstructured progress notes, (2012) 505–511.
[8]   Marafino BJ, Boscardin WJ, Dudley RA. Efficient and sparse feature selection for biomedical text classification via the elastic net: Application to ICU risk stratification from nursing notes. J. Biomed. Inform. 54 (2015) 114–120. doi: 10.1016/j.jbi.2015.02.003

[9]   Neamatullah I, Douglass MM, Lehman LH et al. BMC medical informatics and automated de-identification of free-text medical records. 16 (2008) 1–16. doi: 10.1186/1472-6947-8-32

[10] Popejoy LL, Khalilia MA, Popescu M et al. Quantifying care coordination using natural language processing and domain-specific ontology. (2015) 93–103. doi: 10.1136/amiajnl-2014-002702

[11] Press MJ, Gerber LM, Peng TR et al. Physicians: measurement, quality, and outcomes. (2015) 1299–1305. doi: 10.1111/jgs.13491

[12] Sterling NW, Patzer RE, Di M, Schrager JD. International Journal of Medical Informatics Prediction of emergency department patient disposition based on natural language processing of triage notes. Int. J. Med. Inform. 129 (2019) 184–188. doi: 10.1016/j.ijmedinf.2019.06.008

[13] Topaz M, Radhakrishnan K, Lei V, Zhou L. Mining clinicians' electronic documentation to identify heart failure patients with ineffective self-management: A pilot text-mining study. Stud. Health Technol. Inform. 225 (2016) 856–7. http://www.ncbi.nlm.nih.gov/pubmed/27332377 (accessed September 19, 2019)

[14] Topaz M, Radhakrishnan K, Blackley S, Lei V, Lai K, Zhou L. Studying associations between heart failure self-management and rehospitalizations using natural language processing. West. J. Nurs. Res. 39 (2017) 147–165. doi: 10.1177/0193945916668493

[15] Topaz M, Murga L, Gaddis KM et al. Mining fall-related information in clinical notes: Comparison of rule-based and novel word embedding-based machine learning approach., J. Biomed. Inform. (2019). doi: 10.1016/j.jbi.2019.103103

[16] Travers DA, Haas SW. Using nurses' natural language entries to build a concept-oriented terminology for patients' chief complaints in the emergency department. 36 (2003) 260–270. doi: 10.1016/j.jbi.2003.09.007

[17] Travers D, Haas SW, Waller AE et al. Implementation of emergency medical text classifier for syndromic surveillance public health / biostatistics. (2006) 1365–1374.

[18] Waudby-Smith IE, Tran N, Dubin JA, Lee J. Sentiment in nursing notes as an indicator of out-of-hospital mortality in intensive care patients. (2018) 1–11.

[19] Zhou L, Suominen H, Gedeon T. Adapting state-of-the-art deep language models to clinical information extraction systems: Potentials, challenges, and solutions. J. Med. Internet Res. 21 (2019) 1–15. doi: 10.2196/11499

[20] Dreisbach C, Koleck TA, Bourne PE, Bakken S. A systematic review of natural language processing and text mining of symptoms from electronic patient-authored text data. Int. J. Med. Inform. 125 (2019) 37–46. doi: 10.1016/j.ijmedinf.2019.02.008

[21] Koleck TA, Dreisbach C, Bourne PE, Bakken S. Natural language processing of symptoms documented in free-text narratives of electronic health records: A systematic review. J. Am. Med. Informatics Assoc. 26 (2019) 364–379. doi: 10.1093/jamia/ocy173