

Text Mining and Automation for Processing of Patient Referrals

James Todd¹ Brent Richards² Bruce James Vanstone¹ Adrian Gepp¹

¹Bond Business School, Bond University, Gold Coast, Queensland, Australia

²Department of Intensive Care, Gold Coast University Hospital, Gold Coast, Queensland, Australia

Address for correspondence Adrian Gepp, PhD, MIT, BCom, BIT, Bond Business School, Bond University, Queensland 4229, Australia (e-mail: adgepp@bond.edu.au).

Appl Clin Inform 2018;9:232–237.

Abstract

Background Various tasks within health care processes are repetitive and time-consuming, requiring personnel who could be better utilized elsewhere. The task of assigning clinical urgency categories to internal patient referrals is one such case of a time-consuming process, which may be amenable to automation through the application of text mining and natural language processing (NLP) techniques.

Objective This article aims to trial and evaluate a pilot study for the first component of the task—determining reasons for referrals.

Methods Text is extracted from scanned patient referrals before being processed to remove nonsensical symbols and identify key information. The processed data are compared against a list of conditions that represent possible reasons for referral. Similarity scores are used as a measure of overlap in terms used in the processed data and the condition list.

Results This pilot study was successful, and results indicate that it would be valuable for future research to develop a more sophisticated classification model for determining reasons for referrals. Issues encountered in the pilot study and methods of addressing them were outlined and should be of use to researchers working on similar problems.

Conclusion This pilot study successfully demonstrated that there is potential for automating the assignment of reasons for referrals and provides a foundation for further work to build on. This study also outlined a potential application of text mining and NLP to automating a manual task in hospitals to save time of human resources.

Keywords

- natural language processing
- process improvement
- efficiency improvement
- cost reduction and return on investment
- quantitative methodology

Background and Significance

Progress in the development of natural language processing (NLP) within clinical areas has been outpaced by growth in the wider NLP field due to obstacles in the clinical space.^{1,2} Despite these obstacles, many papers have focused on specific problems that can be summarized at a high level as being problems of moving from unstructured text (e.g., the content of a letter) in clinical documents to structured data that is of more use in automated processes.^{2–5} A common aim of such

studies is to improve the clinical workflow by better extracting information from medical documentation, but NLP applications remain underutilized in the clinical context.²

This article addresses one such opportunity in applying NLP techniques for determining the reasons for patients being referred using doctor referrals. This builds upon the increasingly common use of electronic referrals,⁶ which are better suited than paper-based, hand-written referrals for optical character recognition (OCR) and NLP applications. OCR software extracts textual information from images, such

received

November 14, 2017

accepted after revision

February 15, 2018

Copyright © 2018 Schattauer

DOI <https://doi.org/10.1055/s-0038-1639482>.

ISSN 1869-0327.

as faxed PDF documents. Similar methods have been used in previous research to extract and identify key information from faxed medical records for care providers as well as to assess the overlap of concepts between different medical records.⁷ Determining the reason for a patient's referral is a vital part of allocating resources efficiently to those patients with the greatest need. However, the current process of manually doing so places a burden on hospital registrars, who could be utilized elsewhere.

The current categorization process involves multiple stages of manual inspection of the information in the referral document to assign each patient to a category of urgency. Category assignment generally involves two parts: (1) identifying the condition or diagnosis of the patient and (2) using condition-specific criteria to determine the appropriate urgency category. The approach currently taken to complete this task is outlined in ►Fig. 1. This figure shows the time-consuming process of multiple staff members manually going through each referral received to assign urgency categories. There appears to be considerable scope for increasing productivity through automation.

This article presents a pilot study into the feasibility of automating the classification task for internal patient referrals. Specifically, the feasibility of the first step of category assignment—assigning a condition to a patient referral—is the focus of the study. The process of preparing referral documents and possible classification outcomes is pre-

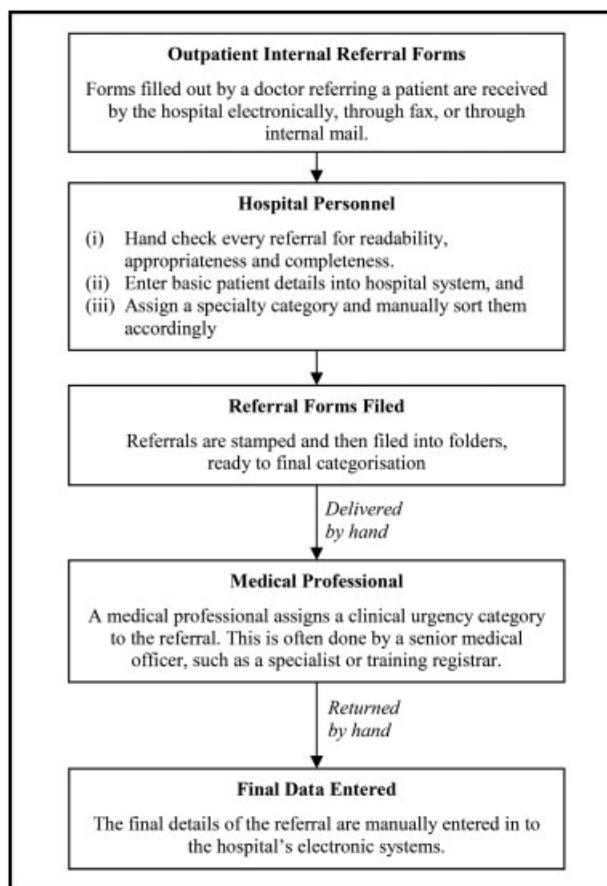


Fig. 1 Current process of determining clinical urgency.

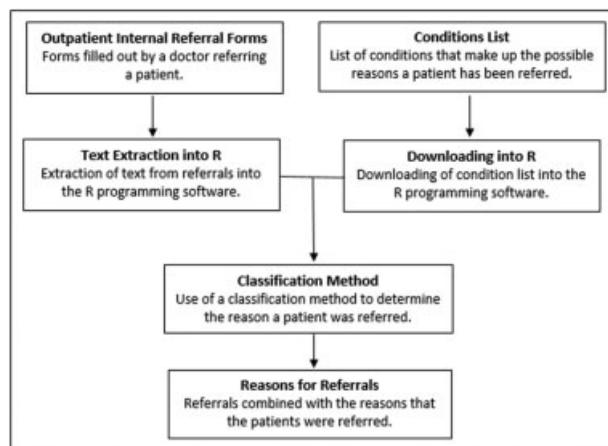


Fig. 2 Proposed process of automating condition assignment.

sented, as well as a simple method for assessing the similarity in terminology used in referrals and condition names. This process is also deployed in an online application.

Methods

The overall approach proposed in this article is shown in ►Fig. 2. The following subsections detail the text extraction from the referrals, the downloading of the condition list, and the classification method employed.

All data extraction, collection, and manipulation tasks were completed using the R programming software, version 3.3.0 for Windows.⁸ R programming software was also used to deploy an online application that performed the steps described in this section. This application allows for the trial of a deployment method throughout the project's life, and thus lets the hospital make deployment decisions in advance of project completion.

Outpatient Internal Referral Forms

The data used in this study consisted of 50 "Outpatient Internal Referral Forms," hereafter referred to as referrals, for patients with neurological conditions. These referrals were faxed to the Gold Coast University Hospital (GCUH), a public teaching and referral hospital with over 500 beds located in Queensland, Australia.⁹ The faxed referrals were stored as individual PDFs, which is the format provided for this project. To ensure the text could be extracted from the patient referrals, they were scanned at 600 dots per inch (dpi), and Adobe Acrobat's OCR was used to convert the files to a format in which the text itself was stored, rather than being an image. The text was then extracted into R programming software. The data were deidentified to obscure sensitive information that could be used to identify the patient, general practitioner, or referring doctor. ►Fig. 3 shows one such referral used in the study.

In most cases, in addition to the original text in the referrals, modifications were made to the hard copy of the referrals. These included handwritten notes, underlining or circling words, and a stamp indicating urgency category.

R-NEW

Queensland Government
Gold Coast Health
Outpatient Internal Referral Form

Internal referrals to Outpatient Specialist clinics are intended for urgent clinical cases requiring specialist input within 30 days (Category 1). Care for other cases (Category 2 & 3) should be transferred back to the GP. Only by direct agreement with the receiving team (according to clinical appropriateness) should Category 2 or 3 cases be referred directly.

Please advise all patients to bring their Medicare Card, any recent x-rays or other reports and current list of medications when presenting to hospital for appointments. Medicare ineligible patients will incur a fee for appointments, payable on presentation.

Referral to:

Reason for Referral:
 2nd seizure leading to MVA ?return to driving

Relevant Clinical Details:
 Known precipitant, sleep deprived and missed 2 by doses of lamotrigine
 Recently changed from valproate to some
 If not Category 1 - Referral has been accepted by Dr:

GENERAL
 SPASTICITY
 NEUROMUSCULAR
 EPILEPSY
 STROKE

Referring Doctor Name: Phone / Speed Dial: Consultant Name: GCH Department: Provider Number:	General Practitioner Details Name: Address: Phone: Fax: Email:
Electronic Signature: _____ Date: _____	

U.R...
 DATE RCVD 27/9/14
 CAT 1 2 3
 SIGNATURE _____
 DATE _____
 DAY PROCEDURE Yes No

Fig. 3 Referral used in the study.

Once the text was extracted, nonsensical elements were removed and key input by the referring doctor was isolated. Punctuation characters were also removed and all letters were recorded in the lower case.

Conditions List and Classification

For the list of potential conditions for referral, the list of conditions from Queensland Health's "Clinical Prioritisation Page"¹⁰ was used, as advised by Dr Brent Richards from the GCUH. This page provides a list of conditions considered for outpatient services, which was downloaded using R programming software.

After downloading the list of conditions, punctuation and stop words (e.g., "to," "the," and "it") were removed to facilitate later investigation of the similarity of terminology in the conditions list and referrals. The set of words treated as stop words was specifically defined for this scenario, given the small data set to which they were applied.

Having stored both the text extracted from patient referrals and the list of possible classifications, a simple method for building associations between these tables was implemented. Limited data, including a lack of reasons for referrals, meant the classification method could not be evaluated directly. Consequently, the following similarity score was

used instead. For each referral, the key input by the referring doctor was compared with the keywords making up each condition. For a given referral, each condition was associated with a score between 0 and 1, calculated as the proportion of keywords for the condition that were also contained in the referral text.

$$\text{Similarity} = \frac{\text{Count}(\text{Key Words in Condition} \cap \text{Unique Words in Referral})}{\text{Count}(\text{Key Words in Condition})}$$

This similarity score is not complex enough for use in a final implementation, but it does provide an indication of the degree to which the terminology employed in referrals and the condition list overlap. This in turn provides an indication of the likely viability of constructing a method for classification using condition names. Development of a reliable classification method is future work that would extend this pilot project, requiring additional data.

Results

The number of words in each referral, after isolating the typewritten content from the referring doctor, is summarized in **Table 1**. For each referral document, similarity

Table 1 The number of words typewritten by referring practitioners in the provided referrals

Minimum	1st quartile	Median	Mean	3rd quartile	Maximum
16.00	35.25	48.50	68.20	87.25	238.00

scores were calculated for each possible condition, and the maximum score was recorded and associated with the referral. ►Fig. 4 shows the distribution of these maximum similarity scores. The prevalence of scores of 50% is evident, being by far the most common maximum similarity score. The prevalence of this score is due to the method employed—conditions with fewer keywords, specifically two keywords in this case, result in very high similarity scores if a single keyword is also found in the referral. Conditions with more keywords may have more cases of keywords being found in the text of a referral, but lower proportions of keywords matched. This is also visible, to a lesser extent, for conditions with three keywords—the next largest groups being 33.3% (1 out of 3) and 66.7% (2 out of 3), respectively.

To evaluate the distribution of similarity scores without distortion from the inclusion of conditions with only two keywords, the process of assigning maximum similarity scores was repeated without such conditions. ►Fig. 5 shows the resulting distribution. While there is still a dispropor-

tionate weighting given to conditions with three keywords, there is a greater variety in similarity scores assigned.

This work was deployed in an online application that allows the GCUH to immediately assess the work done so far and implement a hybrid system whereby personnel manually processing referrals are provided with several suggested conditions based on calculated similarity scores.

Discussion

While the similarity scores being assigned do not appear to be high in many cases and no referrals contain all keywords from any single condition, the lack of cases in which a referral's maximum similarity score is zero is encouraging. No referral completely failed to match at all with a condition, indicating at least some degree of overlap in terminology used and thus some potential for development of a classification method utilizing the text from the condition list. It is important to note that the incidence of high scores is influenced by the word-matching method used. This method was employed because the aim of the study was to assess the viability, rather than implementation of, an automated process. This method is expected to have resulted in both positive and negative bias from various factors, but overall it is believed that the negative bias had a larger effect. As such, the scores observed in this study represent a lower level of performance than could be realistically expected in a final implementation.

Positive bias is expected to have been introduced by the word-matching method not considering the context of matched terms. For example, if the phrases “the patient recently underwent hip surgery” and “the patient does not suffer from chest pain” were both found in one referral, this would be a complete match for “hip pain” as well as “chest pain,” neither of which is desirable. By not considering the context, words from unrelated sentences are matched and negation is ignored. To address this bias in future research, NLP techniques can be employed to break the text into separate sentences and paragraphs as well as to identify and account for negation.

The greater negative bias expected derives from the employed method requiring exact matching of terms in the condition list without considering related words. Related words would include derivative words (e.g., “seizures” is a derivative of “seizure”), abbreviations, and different words referring to the same concept. Several approaches to better considering related words are possible. For derivative words and different words referring to the same concept, the NLP techniques of stemming and lemmatization could be used. Stemming is a process of removing the ends of words to find the root words—for example, “organization” and “organizing” both become “organiz.” Lemmatization is a more sophisticated method of identifying the root meanings of words—for example, “better” has “good” as its lemma. Addressing the use of abbreviations in referrals would involve engaging with clinical professionals to assess commonly used and relevant abbreviations which could then be incorporated in the matching process.

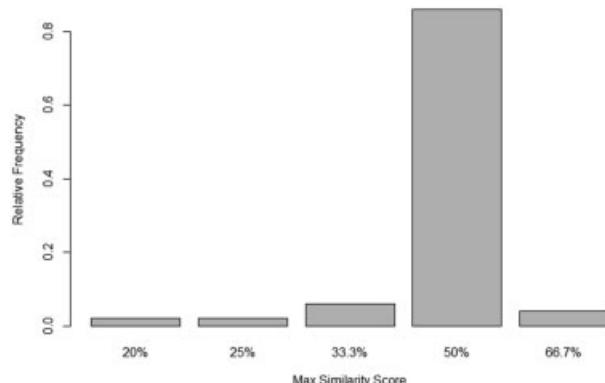


Fig. 4 Relative frequency of maximum similarity scores.

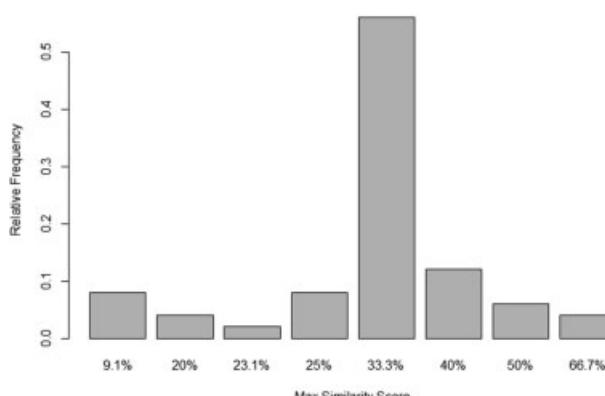


Fig. 5 Relative frequency of maximum similarity scores-conditions with > 2 keywords.

The observed scores suggest that it is feasible to develop a classification method utilizing the free text of the referrals and the condition list, and addressing the identified factors contributing to bias will be the subject of future work.

Lessons Learned

Other takeaways for this study arose from issues encountered when extracting text from the scanned referrals. These issues are caused by the current situation of typed referrals being faxed to a central location, as a step in the move away from handwritten referrals toward an entirely electronic system, which understandably requires a substantial transition period. The lessons learned from working with these referrals were:

1. The use of OCR to recognize text within digital documents is imperfect, and in some cases incorrectly interprets symbols. In most cases, this resulted in letters being recorded as alternate letters or not registered at all, which caused some words to be recorded incorrectly.
2. The conversion of a structured form to a purely textual format causes a loss of contextual information inherent in the original forms. For example, information entered by the referring doctor is less easily differentiated from field names and instructions that were contained in the form prior to any user input.
3. The modification made to the referral documents after they were submitted by the referring doctor prevents some text elements from being extractable or, if it is still extractable, from retaining its original form. Specific examples include underlining, circling, and other forms of emphasis, which tend to prevent the OCR algorithm from properly registering the information. This is particularly concerning since the intention was to emphasize, but the opposite result occurred.
4. The inclusion of stamps to the document, particularly near clinical information, makes retaining sentence structure more difficult. This is because the text is extracted from left-to-right of the document, and so where a sentence should have continued on the next line, the text extraction's interpretation of the stamp is added to the middle of the sentence.

In this study, the first issue was mitigated through scanning documents at a higher resolution, which increased the accuracy of the OCR in recognizing text. The second issue was addressed through the manual identification of terms that often occurred immediately preceding and following information input by the referring doctor, allowing for isolation of this important information. Lastly, the representation of stamps after text extraction was largely meaningless symbols, which were removed using regular expressions within the R code.

To assist future research, it is recommended that future referrals be obtained prior to modifications and in a digital format rather than being faxed. Alternatively, procedures can be put in place to ensure modifications to the referral do not interfere with text extraction, such as stamps being placed away from key information.

Conclusion

This article has put forward an initial methodology for automating the task of assigning customers to clinical urgency categories using text mining of internal patient referrals. Such automation would result in benefits to the hospital due to a reduced burden on time and staff required, with an approach that scales well to higher volumes of referrals if used in other hospitals. This pilot study was successful and found that there is potential for such automation, which should be the focus of future work in this area. An additional contribution of this article was the documentation of the obstacles faced and lessons learned throughout the project, which will also be of assistance to researchers approaching similar problems.

Future work should also address the second stage of the task—to assign clinical urgency categories to referrals, given a reason for the referral being submitted. This would facilitate a further increase in productivity. At a high level, this will involve applying prioritization criteria associated with each possible condition to the free text of the referral.

Clinical Relevance Statement

Automation, in whole or in part, for the processing of patient referrals would save time for hospital staff who otherwise need to manually go through each patient referral received. These staff could then be utilized elsewhere in the hospital.

Multiple Choice Questions

1. Which of the following reduce the effectiveness of automated processing of patient referrals?
 - a. Placing of stamps in a standard location away from the remainder of the text
 - b. Circling of important information on the referral
 - c. Keeping notes added to the referrals in a section separate to the main text
 - d. Use of lemmatization techniques

Correct Answer: The correct answer is option b. A major obstacle to incorporating natural language processing in the processing of referrals are the modifications made to documents by staff, which include circling or underlining of important information. This impedes the ability of text extraction techniques to retain the original meaning of the circled sections, and results in a loss of important information.

2. The slower development of NLP within clinical contexts has been due to which of the following factors?
 - a. Limited sharing of data sets
 - b. A lack of NLP software tailored specifically for medical applications
 - c. Limited collaboration between researchers in health and those in NLP
 - d. All of the above

Correct Answer: The correct answer is option d. All of the factors mentioned have contributed to a slower development of NLP within clinical applications when compared with the progress of the NLP field as a whole, and continuing to overcome these issues would help future development. The lack of shared data sets prevents research from being easily reproduced. The lack of NLP software specifically designed for medical applications reduces the applicability of NLP techniques, as medical terminology tends to differ substantially from the terminology employed in other areas. Consequently, NLP software from other areas is less effective in clinical contexts. Lastly, the limited collaboration between clinical and NLP researchers has impeded the development of NLP in this context, as domain and technical knowledge are both required for effective application.

Note

All codes used in this study are available on BitBucket, and can be accessed through the following link: https://bitbucket.org/JTodd_Bond/patientreferrals/. The actual patient referrals are not available as they contain private information.

Protection of Human and Animal Subjects

Ethics approval has been obtained for this project. The referral documents were deidentified by redacting names and information of medical professionals and patients, and date information before being provided.

Conflict of Interest

None.

References

- 1 Chapman WW, Nadkarni PM, Hirschman L, D'Avolio LW, Savova GK, Uzuner O. Overcoming barriers to NLP for clinical text: the role of shared tasks and the need for additional creative solutions. *J Am Med Inform Assoc* 2011;18(05):540–543
- 2 Wang Y, Wang L, Rastegar-Mojarad M, et al. Clinical information extraction applications: a literature review. *J Biomed Inform* 2018;77:34–49
- 3 Cai T, Giannopoulos AA, Yu S, et al. Natural language processing technologies in radiology research and clinical applications. *Radiographics* 2016;36(01):176–191
- 4 Meystre SM, Haug PJ. Comparing natural language processing tools to extract medical problems from narrative text. *AMIA Annu Symp Proc* 2005:525–529
- 5 Schadow G, McDonald CJ. Extracting structured information from free text pathology reports. *AMIA Annu Symp Proc* 2003:584–588
- 6 Warren J, White S, Day KJ, Gu Y, Pollock M. Introduction of electronic referral from community associated with more timely review by secondary services. *Appl Clin Inform* 2011;2(04): 546–564
- 7 Moon S, Liu S, Kingsbury P, et al. Medical concept intersection between outside medical records and consultant notes—a case study in transferred cardiovascular patients. 2017 IEEE International Conference on Bioinformatics and Biomedicine. Kansas City, MO; November 13–16, 2017:1495–1500
- 8 R Core Team. R: A Language and Environment for Statistical Computing. Vienna, Austria 2016
- 9 Australian Institute of Health and Welfare. *Gold Coast University Hospital*. 2017. Available at: <https://www.myhospitals.gov.au/hospital/310000050/gold-coast-university-hospital>. Accessed September 18, 2017
- 10 Queensland Health. Clinical Prioritisation Criteria. 2016. Available at: <https://cpc.health.qld.gov.au/>. Accessed August 10, 2017