MEDINFO 2019: Health and Wellbeing e-Networks for All
L. Ohno-Machado and B. Séroussi (Eds.)

doi:10.3233/SHTI190524

Recovery Medication from Free Text to a Structured Form

Humberto Fernán Mandirola Brieux[a], Santiago Orozco[a], Marisa Lanfranconi[a], Jorge De All[a], Nahin Chedresse[b], Sebastián Guillen[c], Alejandro Deporte[c]

[a]Sanatorio Otamendi, Buenos Aires, Argentina,
[b]Hospital Naval Puerto Belgrano (HNPB), Argentina,
[c]BIOCOM Sistemas, Argentina

Abstract

We studied methods to convert medical prescriptions in free text to a structured form for pharmacy insructions and planning nursing activities in hospitalized patients. We compared Natural Language Processing (NLP) with Parsing Process (PP), both for the Spanish language. We studied 87,750 and processed 65,000 prescriptions and recovered 62% and 65% with NLP and PP to a structured format respectively. The difference between the methods is significant (p < 0.001) and further work is needed to determine if combining them will have higher performance.

Keywords:

Natural Language Processing, Prescriptions, Medical Order Entry Systems

Introduction

Physicians usually prescribe medications in free text, however, computer systems need this information in a structured way to organize the data so that the pharmacy can deliver the medications and nurses can organize their tasks [1]. Computerized Provider Order Entry (CPOE) can reduce medication errors; but, its benefits are only achieved when data is entered in a structured format and entries are properly coded [2].

To process the information by computer, the system extracts the data from many fields. The system analyzes the following information: generic drug name, dose, strength, posology, route of administration, care and recommendations that are required to comply with the medication.[3].

FHIR standard (Fast Healthcare Interoperability Resources) has resources to manage pharmaceutical orders but it needs this information to be structured. Thus it is fundamental to perform the data extraction in different fields to match the FHIR interoperability standard [4].

Automated dispensing cabinets for pharmacies use HL7 and the information to construct the messages needed in a structured way [5] [6] .

A prescription is an order for medication which is dispensed to a patient. The medical prescription is not enough for the pharmacy to deliver the medication of the day to different departments. It requires prior processing so that the pharmacist knows the daily amount of each drug to deliver [7]. The pharmacist verifies the legality, safety and appropriateness of the prescription order, checks the patient medication record before dispensing the prescription (when such records are kept in the pharmacy), ensures that the quantities of medication are dispensed accurately and decides whether the medication should be handed to the patient, with appropriate counselling, by a pharmacist [8].

Even if doctors are becoming more accustomed to entering data in a structured way using CPOE, there are prescriptions in free text form that need to be processed.

Our question is: “How good can the medical prescription extraction process be when using a free text input for Spanish language?”

We studied Natural Language Processing (NLP) and Parsing Process (PP) as two possible methods to process unstructured text.

Methods

The objective of this work is to demonstrate that it is possible to recover information about medications in a structured format. Passing the narrative medical prescription through NLP is a helpful tool for doctors to better transform it in structured format for hospitalized patients.

We build the NLP routine for this study in C# .net using the Stanford CoreNLP technology tools.

The PP routine was developed by our programmers. Since the data we have uses a fixed pattern, we use the regular expression technique to extract and control its elements.

A retrospective,,observational, cross-sectional, descriptive study was designed. The sample consisted of 87,750 medical prescriptions collected consecutively during March 1th, 2017 and March 1th, 2018. The medical prescriptions were made by the institution's staff and independent external physicians who assist patients within the institution.

We used the following criteria for the study inclusion:

- Adults pharmacological prescriptions
- Singular generic drugs (for example, Losacor®", that has only one drug (Losartan) but not "Losacor D®" that has two drugs Losartan and Hydrochlorothiazide)
- Commercial names where used (for example, "Losacor®" is acceptable for Losartan)
- One prescription (100 characters per line or field) and one drug per line
- Spelling mistakes and different ways of writing each prescription elements that we analyzed

We used the following criteria for the study exclusion:

- Non-pharmacological prescriptions such as prescription for nutrition, vital signs, studies or any other type of indication that is not pharmacological.
- Pediatric medication

- Parenteral hydration plans
- Medications with more than one drug
- Multiple line medications

We included 65,000 and excluded 22,750 medical prescriptions collected consecutively with this process.The medical prescriptions that met the inclusion criteria were processed by the two algorithms.

The average age of the study population of the samples was 67 years old, with the maximum age of 96 and the minimum age of 22 years. The percentage of males participating in the study was 45.3% and the female participation was 54.7%.

In this study we tested two algorithms:

1. The PP method
2. The NLP algorithm

We considered it a success when the algorithm could complete the five fields and the pharmacist confirmed that it was correct.

Results

The results were the following: we were able to recover 62% and 65% with the NLP and the PP method respectively to a structured format that could be processed by computer systems. Please see tables 1 and 2.

The difference between the two methods was statistically significant when calculated using the Z-test with $p < 0.001$.

Table 1 – Results in Absolute Values

	NLP	PP
Correct	40,300	42,250
Incorrect	24,700	22,750

Table 2 – Results in Percentage

	NLP	PP
Correct	62%	65%
Incorrect	38%	35%

Discussion

The PP routine used the regular expression technique to extract data (i.e., narrative text or speech data) in the context of a specific task. Our results have shown that both algorithms are capable of detecting more than 60% of the medical prescriptions made with free text. However, there is a significant difference between the two methods.

The NLP method has the possibility to learn and improve over time, while the PP cannot improve by itself. It requires more expressions to be loaded by manual programming.

Conclusions

It remains to be seen if these methods can achieve higher performance in future. It is possible to make the system learn from medical prescriptions through a knowledge thesaurus. We will work to incorporate these challenges in future work.

Acknowledgements

We thank Dr. Fernán Gonzalez Bernaldo de Quirós , Dr. Daniel Luna, Dra Paula Otero & Dr. Carlos Otero from Hospital Italiano de Buenos Aires for their asistance and corrections of this work.

References

[1] E.S. Chen, P.D. Stetson, Y.A. Lussier, M. Markatou, G. Hripcsak, and C. Friedman, Detection of practice pattern trends through Natural Language Processing of clinical narratives and biomedical literature, *AMIA Annu Symp Proc*. (2007) 120–124.

[2] L. Zhou, L.M. Mahoney, A. Shakurova, F. Goss, F.Y. Chang, D.W. Bates, and R.A. Rocha, How Many Medication Orders are Entered through Free-text in EHRs? - A Study on Hypoglycemic Agents, *AMIA Annu Symp Proc*. **2012** (2012) 1079–1088.

[3] Ö. Uzuner, I. Solti, and E. Cadag, Extracting medication information from clinical text, *J Am Med Inform Assoc*. **17** (2010) 514–518. doi:10.1136/jamia.2010.003947.

[4] Extension: rateGoal - FHIR v3.0.1, (n.d.). https://www.hl7.org/fhir/extension-pharmacy-core-rategoal.html (accessed November 17, 2018).

[5] Automation Anywhere - Largest Enterprises Have Made Their Choice for RPA, (n.d.). https://www.automationanywhere.com/lp/largest-enterprises-have-made-their-choice?utm_source=google&utm_medium=cpc&utm_campaign=rpabroad_ROW&keyword=Robotic%20Process%20Automation&match=b&network=g&place=&gclid=CjwKCAiAlb_fBRBHEiwAzMeEdn7b6HadbZih5i83jnJhLb1KbcqaR8K5JzQ2EWdB-njKFYFOV1PPiBoC970QAvD_BwE (accessed November 17, 2018).

[6] N.F. Carolina PhD, PharmDDirector, Pharmacy Practice LaboratoriesWingate University School of PharmacyWingate, North, Safety of Automated Dispensing Systems, (n.d.). https://www.uspharmacist.com/article/safety-of-automated-dispensing-systems (accessed November 17, 2018).

[7] H. Liu, Q. Burkhart, and D.S. Bell, Evaluation of the NCPDP Structured and Codified Sig Format for e-prescriptions, *Journal of the American Medical Informatics Association : JAMIA*. **18** (2011) 645. doi:10.1136/amiajnl-2010-000034.

[8] The Role of the Pharmacist in the Health Care System: PART I: THE ROLE OF THE PHARMACIST IN THE HEALTH CARE SYSTEM: 4. The scope of pharmacy and the functions of pharmacists: 4.2 Community pharmacy, (n.d.). http://apps.who.int/medicinedocs/es/d/Jh2995e/1.6.2.html (accessed November 13, 2018).

Address for correspondence

Humberto F Mandirola Brieux. Email hmandirola@gmail.com; Mobile Phone +54911-44174879.