Faculty of Mathematics and Computer Science

Emanuel Bîscă

Knowledge based systems and language technology

Natural Language Processing in Medicine

2021

Department of Computer Science

Abstract

The most notable challenges in NLP are involving, most of the time, the generation of the natural language as well as understanding it. In this paper, I will discuss about KD-NLP – *Kawasaki disease language processing*, a tool which will be used in Medicine to gain more knowledge about various patients and their chance of suffering from Kawasaki syndrome.

The objective of my paper, is to show that NLP can help clinicians in their attempt of differential diagnosis, being used when investigating the patients files. NLP using straight pattern recognition could be successfully utilized to highlight the signs of the disease, by evaluating narrative texts provided from the Emergency Department notes. This attempt is possible because the studied disease is characterized by fever and a set of 5 known symptoms, so the desired tool must identify these symptoms in the patient files to conclude if there is a low or a high likelihood of KD.

Contents

1. Background and Significance

Natural language processing or NLP is considered to be a field of artificial intelligence and computer science, which uses linguistic analysis. NLP is concerned with the study of different kind of interactions between human language and computers, with a special attention to the analysis of the way in which a system can be programmed so that it'll be capable of processing and analysing a great amount of natural language data. The resulting systems makes a computer able to *understand* a document content, with respect to its contextual nuances of the used language. The capabilities extends to the accurate extraction of meaningful information or various insights included in the selected document.

The variety of NLP algorithms extends to machine learning and even pattern recognition. They can include probabilistic models which would have different weights on specific variables or decision trees, with the usual if-then branches. The extraction of phenotypic information provided by electronic health records, with the help of natural language processing, is being seen as an area of great interest for many scientists, throughout both computer science and medicine. This fact is happening due to the increasing use of EMRs (stands for electronic medical records), delivering huge amounts of clinical data waiting to be studied. In this approach, there have been used notes of clinicians, as narrative forms.

I believe it is ideal to apply NLP techniques when diagnosing the Kawasaki disease (KD), because it rests mainly on acknowledging the presence of the fever together with the five traditional signs identified by Dr. Kawasaki back in the 60s. NLP can help clinicians in their attempt of differential diagnosis, being used when investigating the patients. Consequently, I think it's clear why NLP using straight pattern recognition could be successfully utilized to highlight the signs of the disease, by evaluating narrative texts provided from the Emergency Department (ED) notes.

A great prevention encountered by the ED health workers, to correctly diagnose a patient with KD, is overlooking the disease from the differential diagnosis of children experiencing mucocutaneous symptoms and fever. The concerns are growing higher, because a missed or a tardy diagnosis of KD is often the cause of acquired heart disorder in kids, in most developed countries (Newburger et al. 2004). If the KD is identified on the early stages and the treatment begins with intravenous administration of immunoglobulin, the risks of suffering a coronary artery aneurysms, which is the most dangerous sequela of Kawasaki disease, drops dramatically with 20%, to 5%, while a ten days late diagnosis increases the chance of coronary aneurysms by 3% to 7%.

There does not exist a final laboratory test that can diagnose the KD, the progress consists of newly improved algorithms that can help the health workers to correctly diagnose the KD. Ling et al. (2011) developed an algorithm implemented to identify the disorder utilizing 12 different laboratory variables that performs properly when differentiating among usual febrile episodes and KD. Still, using such an approach, health providers should take into account Kawasaki disorder as part of this differential diagnosing in the children experiencing febrile episodes. But statistically speaking, this is not feasible, as KD evolves continuously over time causing the patient to present episodic disappearances of key clinical signs (Wilder et al. 2007).

Looking at the subjects definitively diagnosed with Kawasaki syndrome, a stunning 95% of them have been consulted by a specialist in the first 5 days of illness. Yet, the percents of the patients properly diagnosed on the first medical visit are around 4.7%, with a mean of 3 visits to the hospital, prior to diagnosing. In a nutshell, the delayed proper diagnostic is increasing the chances of cardiovascular complications both in newborns and children with KD.

1.1 Investigation target

The target of my research study, is to theoretically investigate the design and the performances of a NLP tool, which will be referred to as KD-NLP, utilized in discovering the high suspicious subjects of Kawasaki disease. The input is consisted of texts from clinical notes in the EMR. The KD-NLP is developed as a screening tool to suggest further tests required to precisely diagnose the KD, while the KD-NLP is intended to discern the subjects which are experiencing fever in combination with three or more clinical clues of Kawasaki disease and in which the possibility of diagnosing the KD must be considered.

2. NLP in Biosurveillance

Doan et al. (2016) conducted a study with subjects from two hospitals from the United States, between January 2010 and December 2014. The current section is based on this study. The two sites were as follows, the EMR from Rady Children's Hospital San Diego and Children's Healthcare of Atlanta. In the next sections, the study uses data about patients with Kawasaki disease, whom where diagnosed among the third day of illness and the tenth one , where first day of sickness was considered the first one with fever signs. Beside fever, other four or five clinical criterion from Figure 2.1 had to be present for a eventual diagnostic, while if the patient had abnormalities documented by echocardiogram, fewer criterion had to be simultaneously fulfilled.

Clinical Criterion	Tag name	Description and examples
Fever	FEVER	Fever or temperature at least 100.4°F or 38°C *Examples:* Fever Any mention of temperature 100.4°F (38°C) or above
Bilateral conjunctival injection	CONJUNCTIVAL_INJECTION	Bilateral bulbar conjunctival injection without exudate *Examples:* Redness of eyes Eyes: positive for redness
Changes of the oropharynx: injected pharynx, injected, fissured lips, strawberry tongue	ORAL_CHANGES	Changes in lips and oral cavity, including erythema, cracked lips, strawberry tongue, diffuse injection of oral and pharyngeal mucosae *Examples:* Red cracked lips Strawberry like tongue
Changes of the peripheral extremities: peripheral edema, palm/sole erythema, periungual desquamation	EXTREMITY_CHANGES	Changes in extremities: palms, soles, hands, feet, or periungual peeling of fingers or toes *Examples:* Redness of hands or feet Swelling of hands or feet
Polymorphous rash	POLYMORPHOUS_EXANTHEMA	Polymorphous exanthema *Examples:* Skin rash Pink blanching patches scattered on body
Cervical adenopathy > 1.5 cm	CERVICAL_LYMPHADENOPATHY	Cervical lymphadenopathy (≥1.5 cm diameter), usually unilateral *Examples:* Neck adenopathy Neck swelling

KD = Kawasaki disease.

Figure 2.1: List of Semantic Tags for KD Tagger

2.1 Modelling a standard of criteria

The goal of the investigation is to analyze the NLP tool that would become proper for use in the emergency departments, and the tool is using a single clinical note provided by the emergency department, for each patient considered in this investigation. For further ado, we can say that a number of children with fever (FC) are considered part of this study, febrile children who were consulted by health workers from the ED and in their differential diagnosis had been specifically taken into account the evaluation for KD, but the FC have not ultimately been diagnosed with KD.

To develop the NLP tool, there has been chosen a convenience sample of notes belonging to 22 children from the Site 1, whom were diagnosed with Kawasaki syndrome, during a three months period; that's the case since there are no standard numbers of patients to be considered, nor other instances of documentation, such as the number of times we found a sign of KD in the notes. Another convenience sample of subjects was used to evaluate the performance of the NLP tool (166 notes from Site 1 and 87 notes from Site 2).

After identifying the necessity of a standard of criteria, all of the 253 notes have been manually reviewed by a group of medical specialists. They analyzed individually each note provided by the ED to validate the absence or presence of one or more of the KD signs, without having access to subjects information, outside what they already found in the corresponding notes. Then, the assessments were compared and discussed until an uniformly concluded standard has been stated.

2.2 KD tagger development process

The NLP tool contains as the most important element, a KD tagger. Its work is basically to discover a predefined sign, to help us discover if there is a trace of Kawasaki related symptoms, within a clinical note corresponding to a patient. To conceive the tagger and to make it to find out the KD signs, some annotation benchmark has bee created from those 22 initial notes. The goal of this benchmark was the establishing of a list with KD signs, which would be translated into linguistic reports easy to comprehend by the NLP tagger, from medical language. More details about the KD clinical criteria and their corresponding semantic tags together with some examples from the actual notes can be found in the Figure 2.1.

The benchmark has been built on four components, definitions, patterns, examples and negations. American Hearth Association offered guidelines to establish the definitions. The training notes together with the experience of the clinical specialists were used to generate patterns from a word or phrase describing a Kawasaki disease signal, with respect to the clinical note texts. Sometimes in these texts, there were words or phrases which shouldn't have been annotated, the negations explain what not to annotate.

Then, an environment in compliance with Health Insurance Portability and Accountability Act (HIPPA), has been elaborated within iDASH (Ohno-Machado et al. 2012), with the use of ***Brat***, a visual tool used when annotating the notes (Stenetorp et al. 2012).

2.3 KD-NLP overview

The KD-NLP pipeline consists of three main modules:

- ***Preprocessing:*** *ED provider notes were copied in their entirety from the EMR (electronic medical records) and saved in Microsoft Word format. Then a doc2txt package was used to convert text from MS Word format to plain-text format. After that, a Perl-based sentence splitter program developed by the group of experts was utilized to divide text into individual sentences.*
- ***KD tagger:*** *The KD tagger recognizes fever and KD signs from clinical text.*
- ***KD classifier:*** *This module assigns a subject as high suspicion for KD if the KD tagger detects fever and three or more KD signs; otherwise, it assigns the subject as low suspicion for KD.*

Figure 2.2: The KD-NLP outline

The ***doc2text*** package used in the preprocessing step is a Perl based command line utility.

2.3.1 The KD-NLP tagger

In the next figure, we can see the scheme which explain the KD-NLP. As it can be observed, the tagger has three parts.

Lexicon look-up – using the initial 22 notes, after annotation and training, there were obtained 313 keywords or key phrases as part of the preliminary lexicon (related KD). This list was then increased to include various synonyms, in accordance with the Unified Medical Language System (known as UMLS) dictionary Lindberg et al. (1993). The final lexicon contained a number of 28 580 keywords.

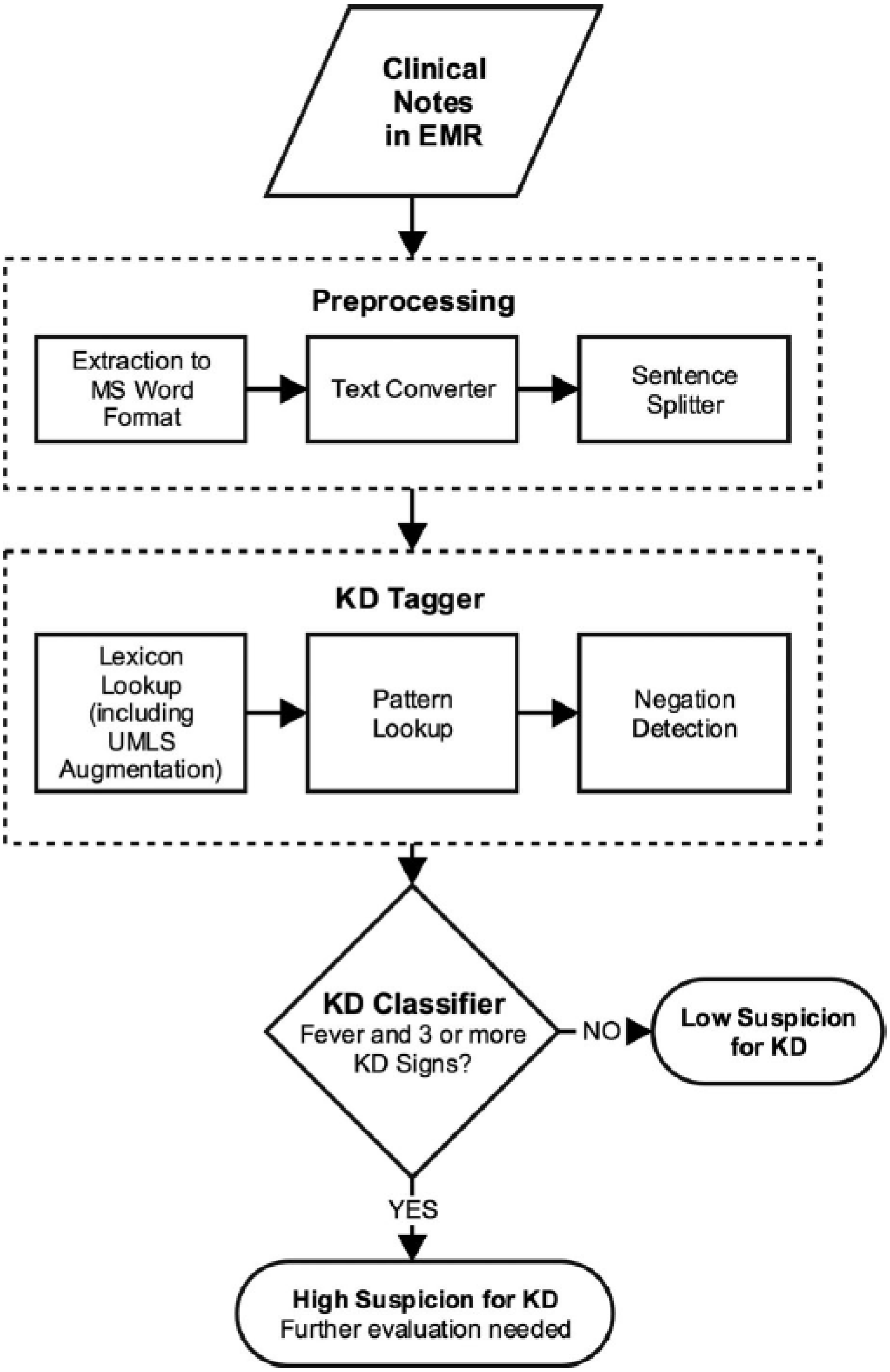

Figure 2.3: The KD-NLP framework

Pattern look-up – patterns must be seen as some regular expressions. As an example: for the changes in the extremities, where we have the tag EXTREMITIES_CHANGES, if we consider the given sentence "hands and feet appeared red and swollen", changes in extremities goes with the common statement *hands* or *feet* or *hands and feet* or *hands or feet* (are) *red* or *swollen* or *red and swollen.*

Negation detection – is a recognition software which distinguishes a denial linked between a Kawasaki disease signal from usual notes, i.e. for example "no swollen feet". This negation detection control gives us the means to not annotate some wrong sign of KD as a KD semantic tag.

2.3.2 The KD-NLP classifier

The purpose of this classifier is detection of subject with a high probability of suffering from Kawasaki syndrome, based on their symptoms as previously specified. Conducting a research with another group of Kawasaki diagnosed patients, it was outlined that from a number of 455 people:

- around 18% shown 5 medical proofs of the disease (83 subjects);
- around 56% shown 4 medical proofs (257 subjects);
- around 18% shown 3 medical proofs (80 subjects);
- around 7% shown 2 medical proofs (30 subjects);
- around 1% presented 1 clinical evidence of having the disease (5 subjects);

The meaningful part here, is that we observe 92 percents of the patients are describing at least three medical signals of KD. Even if a negligible sensitivity can be earned by adding the patients presenting at most two KD signs, there would be a dramatic decrease of specificity taking into account those patients. Consequently, only the subjects with febrile episodes and at least 3 signs of KD would be considered with a high likelihood of having KD.

2.4 Results

The sensitivity together with the specificity have been computed for the performances of the KD-NLP to properly outline the suspicious patients, identifying at least 3 individual signs and fever. The 95% confidence intervals (CIs) have been also computed for the sake of a greater image. In the next table, more information about these statistics can be found.

Sites	Sensitivity	Specificity
Site 1	92.7 (87.8 – 97.6)	79.0 (68.4 – 89.5)
Site 2	95.3 (90.1 – 100.0	73.9 (56.0 – 91.9)
Site 1 and Site 2 combined	93.6 (90.9 – 97.3)	77.5 (68.4 – 86.7)
Data is reported as % (CI) – confidence intervals		

Figure 2.4: Site performances of KD-NLP

When identifying individual proofs of Kawasaki syndrome, the KD-NLP tool performed really well, with the highest sensitivity (> 95%) for *cervical lymphadenopathy* (the meaning of lymphadenopathy is usually the given name of an enlargement syndrome of the lymph nodes), *conjunctival injection* (it basically refers to the inflamation of the conjunctival vessels) and *rash*. Also, the NLP tool was capable to properly distinguish all the patients suffering from *fever*.

2.4.1 Origins of fallacy

The KD-NLP tool had some difficulties of misclassification. There were detected 11 false-negative low suspicious cases of KD patients (11 from a total of 173, i.e. 6.4%) and 18 false-negative high suspicious cases of KD subjects (18 from a total of 180, i.e. 10%).

False-negative cases 11	False-positive cases 18
9 – (81.8%) **due to omission or misclassification of string patterns or keywords,** e.g., *"erythema of palms and soles"* was classified as POLYMORPHOUS_ EXANTHEMA rather than EXTREMITY_CHANGES due to the word erythema being common to both tags	**12 – (66.6%)** **due to assigning the wrong KD sign to a pattern** e.g., "erythema of pharynx" or "erythematous pharynx" as POLYMORPHOUS_EXANTHEMA
1 – (9.1%) **due to misspelling,** e.g., *"midl swelling to hands"*	**5 – (27.8%)** **due to failing to recognize the negation** e.g., *"neck without rigidity or adenopathy"*
1 – (9.1%) **due to preprocessing,** e.g., line break in sentence splitting	**1 – (5.6%)** **due to a hypothetical sentence in the discharge instructions** e.g., *"monitor at home for peeling of hands and feet"*

Figure 2.5: Fallacies explained

3. **Final Remarks**

The design of the KD-NLP has been discussed in my theoretical report. The KD-NLP – *Kawasaki disease natural language processing* tool performed comparably with the human reviewers of the same clinical notes, in detecting the Emergency Department subjects presenting three or more signs of KD, which has been considered a high likelihood of Kawasaki syndrome.

I find it essential to mention once more, that the tool is considered to be a reliable source of knowledge utilized to identify a low or a high likelihood of a patient to suffer from Kawasaki disease, **not** as a proper tool to diagnose Kawasaki disease. NLP system similar to this one have been studied in the past few years. Some such tools have properly identify fever (Chapman et al. 2004) and medication information (Uzuner et al. 2010).

There has been observed, in the last few years, a tremendous increase of interest in the Natural Language Processing methods with application in the medical field. The progress known by the NLP methods in the clinical area is great without a doubt; see Névéol and Zweigenbaum (2015) and Velupillai et al. (2015).

References

Chapman, W.W., Dowling, J.N., and Wagner, M.M. (2004). “Fever detection from free-text clinical records for biosurveillance”. In: *Journal of Biomedical Informatics* 37.2, pp. 183–193.

Doan, S., Maehara, C.K., and Chaparro, J.D. (2016). “Building a Natural Language Processing Tool to Identify Patients With High Clinical Suspicion for Kawasaki Disease from Emergency Department Notes”. In: *Official Journal of the Society for Academic Emergency Medicine* 23.5, pp. 628–636.

Lindberg, D.A., Humphreys, B.L., and McCray, A.T. (1993). “The Unified Medical Language System”. In: *Methods of Information in Medicine* 32.4, pp. 281–291.

Ling, X.B., Lau, K., and Kanegaye, J.T. (2011). “A diagnostic algorithm combining clinical and molecular data distinguishes Kawasaki disease from other febrile illnesses”. In: *BMC Medicine* 9.130.

Névéol, A. and Zweigenbaum, P. (2015). “Clinical Natural Language Processing in 2014: Foundational Methods Supporting Efficient Healthcare”. In: *Yearbook of Medical Informatics*, pp. 194–198.

Newburger, J.W., Takahashi, M., and Gerber, M.A. (2004). “Diagnosis, Treatment, and Long-Term Management of Kawasaki Disease: A Statement for Health Professionals From the Committee on Rheumatic Fever, Endocarditis and Kawasaki Disease, Council on Cardiovascular Disease in the Young, American Heart Association”. In: *Circulation* 110.17, pp. 2747–2771.

Ohno-Machado, L, Bafna, V., and Boxwala, A.A. (2012). "iDASH: integrating data for analysis, anonymization, and sharing". In: *Journal of the American Medical Informatics Association* 19.2, pp. 196–201.

Stenetorp, P., Pyysalo, S., Topic, G., Ohta, T., Ananiadou, S., and Tsujii, J. (2012). "BRAT: a web-based tool for NLP-assisted text annotation". In: *The proceedings of the 13th Conference of the European Chapter of the Association for computational Linguistics.*

Uzuner, O., Solti, I., and Cadag, E. (2010). "Fever detection from free-text clinical records for biosurveillance". In: *Journal of the American Medical Informatics Association* 17.5, pp. 514–518.

Velupillai, S., Mowery, D., South, B.R., Kvist, M., and Dalianis, H. (2015). "Recent Advances in Clinical Natural Language Processing in Support of Semantic Analysis". In: *Yearbook of Medical Informatics*, pp. 183–193.

Wilder, M.S., Palinkas, L.A., Kao, A.S., Bastian, J.F., Turner, C.L., and Burns, J.C. (2007). "Delayed Diagnosis by Physicians Contributes to the Development of Coronary Artery Aneurysms in Children With Kawasaki Syndrome". In: *The Pediatric Infectious Disease Journal* 26.3, pp. 256–260.