Artificial Intelligence in Medicine 64 (2015) 17–27

Contents lists available at ScienceDirect

Artificial Intelligence in Medicine

journal homepage: www.elsevier.com/locate/aiim

Position Paper

Sentiment analysis in medical settings: New opportunities and challenges

Kerstin Denecke*, Yihan Deng

Innovation Center Computer Assisted Surgery, University of Leipzig, Semmelweisstr. 14, 04103 Leipzig, Germany

ARTICLE INFO

Article history:
Received 2 September 2014
Received in revised form 18 March 2015
Accepted 25 March 2015

Keywords:
Sentiment analysis
Clinical text mining
Medical language processing
Health status analysis

ABSTRACT

Objective: Clinical documents reflect a patient's health status in terms of observations and contain objective information such as descriptions of examination results, diagnoses and interventions. To evaluate this information properly, assessing positive or negative clinical outcomes or judging the impact of a medical condition on patient's well being are essential. Although methods of sentiment analysis have been developed to address these tasks, they have not yet found broad application in the medical domain.
Methods and material: In this work, we characterize the facets of sentiment in the medical sphere and identify potential use cases. Through a literature review, we summarize the state of the art in healthcare settings. To determine the linguistic peculiarities of sentiment in medical texts and to collect open research questions of sentiment analysis in medicine, we perform a quantitative assessment with respect to word usage and sentiment distribution of a dataset of clinical narratives and medical social media derived from six different sources.
Results: Word usage in clinical narratives differs from that in medical social media: Nouns predominate. Even though adjectives are also frequently used, they mainly describe body locations. Between 12% and 15% of sentiment terms are determined in medical social media datasets when applying existing sentiment lexicons. In contrast, in clinical narratives only between 5% and 11% opinionated terms were identified. This proves the less subjective use of language in clinical narratives, requiring adaptations to existing methods for sentiment analysis.
Conclusions: Medical sentiment concerns the patient's health status, medical conditions and treatment. Its analysis and extraction from texts has multiple applications, even for clinical narratives that remained so far unconsidered. Given the varying usage and meanings of terms, sentiment analysis from medical documents requires a domain-specific sentiment source and complementary context-dependent features to be able to correctly interpret the implicit sentiment.

1. Introduction

In the event of complex medical conditions and associated treatments, patients are examined by different types of specialists. Because they are specialized in a particular medical field, e.g. radiology, surgery, paediatrics etc. this leads to a restricted view of each patient's status and their medical conditions. In their documentations, physicians describe their personal views and observations. This might be a judgement or evaluation, an affective state, or be intended to provoke some emotion in the reader. Sentiment analysis methods aim to determine the attitude expressed with respect to some theme or the overall contextual polarity of a document. Originating in the field of web mining, the development of sentiment analysis methods often concentrates on processing very subjective texts such as customer reviews [1,2]. Limited work has shown that less subjectively written texts can be analysed using sentiment analysis methods, enabling better insights into semantics [3].

In this work, we describe facets and potentials of sentiment analysis in the context of medicine and healthcare for several reasons. A treatment process often involves various persons, including physicians of different specialities, nurses, therapists etc. Since the personal observations and attitudes of a physician influence clinical decision-making, it is crucial to identify them in medical records so that a complete view of a patient's health status can be achieved and presented to other treating healthcare professionals. While examination results are often reported in a structured manner, observations or experiences are communicated in an unstructured way in finding reports or other clinical documents. Accordingly, extracting opinions and intentions from medical narratives can be

* Corresponding author. Tel.: +49 341 97 12002; fax: +49 341 97 12009.
E-mail address: kerstin.denecke@iccas.de (K. Denecke).

http://dx.doi.org/10.1016/j.artmed.2015.03.006
0933-3657/

Table 1
Entities and events in the medical domain and possible sentiment characteristics.

Entity	Possible sentiment values
Health status	Improve, worsen
Medical condition	Present, improve, worsen
Diagnosis	Certain, uncertain, preliminary
Effect of a medical event	Critical, non-critical
Medical procedure	Positive or negative outcome, successful or unsuccessful
Medication	Helpful, useless, serious adverse events

important for assessing clinical data, monitoring a patient's health status or providing automated decision support for physicians. Even though clinical documents are often written in an objective manner, medical conditions impact a patient's life. Sentiment analysis can help determine this impact from written documents.

Natural language processing of clinical narratives has sparked increasing attentions in recent years, resulting in effective algorithms for named entity recognition and relation extraction methods [4]. Based on recognized entities and relations among them, the analysis of opinion and sentiment in clinical narratives can offer a higher-level text understanding [4]. However, the sentiments expressed in clinical narratives have not been well analysed and exploited yet. This paper examines this newly emerging research topic, identifies potential use cases, and summarizes the main open research questions by:

- Determining the current state of the art in sentiment analysis in healthcare settings
- Describing uses and potential of sentiment analysis in medicine
- Characterizing facets of sentiment in the domain of medicine and drawing conclusions for technical developments in this field

In the following section, we describe the facets of sentiment in the medical context. Section 3 summarizes existing work on sentiment analysis. To characterize sentiment, we analysed and compared the linguistic peculiarities of clinical and medical social media texts with respect to subjectivity and opinions (Section 4). Possible use cases were collected in discussions with physicians and are described in Section 5. Additionally, research challenges for the future development of medical sentiment analysis methods are outlined. The paper finishes with conclusions in Section 6.

2. The notions of sentiment in medicine

Textual information can be broadly categorized into facts and opinions [2]. Facts are objective expressions about entities or events. Opinions are usually subjective expressions describing people's attitudes, sentiments or feelings about entities. However, the concept of sentiment or opinion is quite broad, encompassing subjectivity, polarity, emotion or even comparison. While it is straightforward for a person to like or dislike a movie or product, sentiment in the context of medicine is difficult to capture in a few words. Facets of sentiment in health-related texts may concern (see Table 1):

- A **change** in health status (e.g a patient can suddenly feel better or worse)
- **Critical events, unexpected situations or specific medical conditions** that impact the patient's life (e.g. *tumour is malignant* as such is a fact, but this medical condition is negative for the patient since it might lead to health problems or death)
- The **outcome or effectiveness of a treatment** (e.g. surgery may be successfully completed)

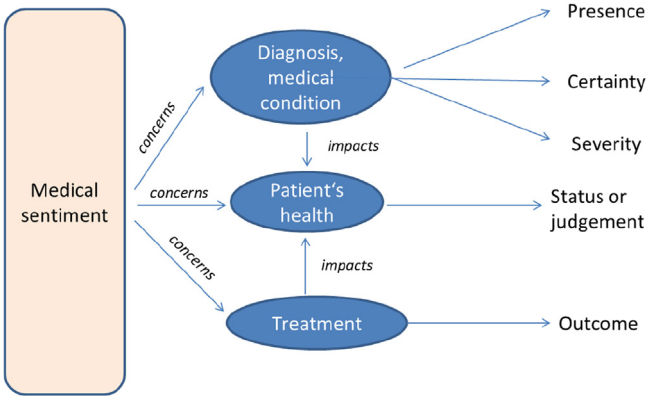

Fig. 1. Facets of sentiment in medical contexts. Sentiment can concern patient's health, a medical condition or treatment. For each aspect, sentiment can occur in different forms.

- **Experiences or opinions of a treatment or a sort of drug** (e.g. a patient or a physician can describe serious adverse events after drug consumption)
- The **certainty of a diagnosis** (e.g. a physician may be certain of some diagnosis)

The examples show that *good* and *bad* or *positive* and *negative* in the context of medicine concern the health status, a medical condition or a form of treatment. It is manifested in improvements or deteriorations of certain medical or physical conditions or in the success or failure of a treatment (see Fig. 1). Below these facets are described in more detail.

Sentiment can be seen as a **reflection of the health status** of a patient which can be *good*, *bad* or *normal* at some point in time. Thus, the health status impacts the quality of life of a patient (e.g. a *severe pain* affects the life of a patient much more than a *slight pain*). By analysing health status over time, improvements or deteriorations in the status can be recognized. In clinical narratives, the health status is expressed either implicitly or explicitly. Implicit descriptions of the health status concern the mentioning of symptoms (e.g. *severe pain, extreme weight loss, high blood pressure*). They require additional content information for correct interpretation. An explicit description of the health status is reflected in phrases such as *the patient recovered well* or *normal*.

A medical condition can exist, improve or worsen. Thus, sentiment can also be considered as the **presence or change of a medical condition**. Sentiment in this context can be implicitly considered as the severity of a disease, which again impacts on life circumstances. A medical condition can have different weights: a chief complaint might affect the health status much more than another symptom. Furthermore, the **certainty of a diagnosis** can be seen as the opinion of a physician. For example, a diagnosis may be a suspicion or it may be assured. This opinion on the certainty of a diagnosis impacts the treatment decision: if a diagnosis is certain, a treatment decision can be made; otherwise additional examinations are necessary. Another interesting facet of sentiment concerns the **judgement of medical conditions**, in particular with respect to their severity. For instance, events such as a *bleeding* can be positive or negative, critical or less critical. The phrase *blood pressure decreased* could express a positive or negative change depending on the previous state of blood pressure. A decrease of blood pressure can be good if it was too high before. This also shows that sentiment in clinical narratives cannot always be manifested in single terms or phrases, and that the context is important.

Additional sentiment aspects concern treatment. It can be complex or less complex, urgent or less urgent. The **outcome of a treatment** may be *positive, negative* (e.g. surgery was successful