Chapter 4
On Analyzing Complex Data Within Clinical Decision Support Systems

Jan Kalina
https://orcid.org/0000-0002-8491-0364
Institute of Computer Science, The Czech Academy of Sciences, Czech Republic

ABSTRACT

Clinical decision support systems (CDSSs) represent digital health tools applicable to important tasks within the clinical decision-making process. Training data-driven CDSSs requires extracting medical knowledge from the available information by means of machine learning. The analysis of the complex (possibly big or high-dimensional) training data allows knowledge relevant to be obtained for clinical decisions related to the diagnosis, therapy, or prognosis. This chapter is devoted to training CDSSs by machine learning based on complex data. Remarkable recent examples of CDSSs including those based on deep learning are recalled here. Principles, challenges, or ethical aspects of machine learning are discussed here in the context of CDSSs. Attention is paid to dimensionality reduction, deep learning methods for big data, or explainability of the data analysis methods. Data analysis issues are discussed also for two particular CDSSs on which the author of this chapter participated.

INTRODUCTION

Clinical decision making can be described from a broader perspective as a process of selecting an activity or series of activities among several alternatives related to medical care for an individual patient (Kashyap, 2021). The clinical decision making tasks solved by physicians are focused on the cognition and determination of diagnosis, therapy, and prognosis based on data and knowledge (Berner, 2016). Clinical decision support systems (CDSSs) represent tools of artificial intelligence (AI) helpful for determining the diagnosis, recommending a suitable therapy, or constructing prediction models for individual prognosis. Most of the available CDSSs are designed to find the diagnosis within individual clinical fields. Only some others are specialized to assist with prescription of medicaments, because the

DOI: 10.4018/978-1-6684-5092-5.ch004

search for an appropriate therapy seems to be more complex than determining the diagnosis. Decision making related to therapy depends on many factors so that only a few decision support systems aiming at therapy have been sufficiently evaluated up to now. There has been even less attention paid to decision support systems for prognosis.

Efficient and reliable clinical decision making requires to analyze biomedical data under uncertainty. The (possibly big or high-dimensional) biomedical data may also be accompanied by demographic, behavioral, genetic, or environmental information. The amount of such available data increases hand in hand with the omnipresent digitalization of healthcare. This, together with the increasing understanding of complex diseases (Dagliati et al., 2018), contributes to the increasing complexity of clinical decision making. Therefore, analyzing the data requires sophisticated methods in order to learn the medical knowledge relevant for clinical decision making; advanced methods are also necessary to apply the obtained knowledge within the care for an individual patient.

CDSSs are (often very intricate) systems with a potential to become very useful and powerful digital health tools for the clinical decision making process with the ability to compare different possibilities in terms of their risk. CDSSs are capable to solve a variety of practical tasks such as to combine information components, to extract information under uncertainty due to various sources (individual variability of patients, different levels of severity of diseases, additional factors not considered in the study), and deduce conclusions and recommendations, which are then considered by the physician that carries the full responsibility for the decisions. The training of CDSSs is inconceivable without machine learning methods for data synthesis and formulation of knowledge relevant for clinical care by classification (supervised learning) or regression (e.g. for predicting the survival). While the first systems based on expert knowledge (i.e. knowledge-based, rule-based) were not sufficiently successful, improved performance has been obtained with systems learning the knowledge from clinically relevant data.

CDSSs have established their place as inherent part of e-health with an indisputable ability to improve clinical decision making and with a facility to contribute to a transformation of clinical medicine (Greenes & Del Fiol, 2022). Decision support systems are nowadays often integrated within telemedicine care (Cerrato & Halamka, 2020). In a hospital, a CDSS may run online in real-time and may automatically process the data flow and search for inconsistencies or irregularities (anomalies). Still, there are still obstacles to apply decision support systems in healthcare routinely, although everyday healthcare would greatly benefit from reliable interdisciplinary and multidisciplinary systems; thus, intensive attention has been to approaches for overcoming barriers that prevent from a routine analysis of CDSSs in the everyday care (Laka et al., 2021). So far, the literature has intensively focused on technical and implementation aspects of CDSSs overviewed e.g. in recent monographs on medical informatics (Tenenbaum & Ranallo, 2021); nevertheless, much less attention has been paid to data analysis issues.

This chapter starts with a review of remarkable recent CDSSs; the presented review is oriented on the machine learning tools used for acquiring medical knowledge from training data, with a special focus on deep learning and explainable learning for classification or regression. Then, principles and challenges of analyzing big medical data within CDSSs are discussed. A section devoted to dimensionality reduction follows. After that, various sources of bias of algorithms used within CDSSs are discussed. Finally, data analysis issues are discussed for two particular CDSSs on which the author of this chapter participated.

Related Content