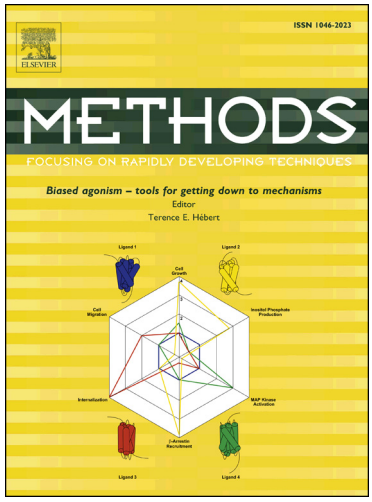
Accepted Manuscript



Discovering Disease Phenotypes from a Large Database of Inpatient Records: A Sepsis Study

Djordje Gligorijevic, Jelena Stojanovic, Zoran Obradovic

PII: S1046-2023(16)30232-8
DOI: http://dx.doi.org/10.1016/j.ymeth.2016.07.021
Reference: YMETH 4047

To appear in: *Methods*

Received Date: 2 May 2016
Revised Date: 26 July 2016
Accepted Date: 26 July 2016

Please cite this article as: D. Gligorijevic, J. Stojanovic, Z. Obradovic, Discovering Disease Phenotypes from a Large Database of Inpatient Records: A Sepsis Study, *Methods* (2016), doi: http://dx.doi.org/10.1016/j.ymeth.2016.07.021

This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

ACCEPTED MANUSCRIPT

Discovering Disease Phenotypes from a Large Database of Inpatient Records: A Sepsis Study

Djordje Gligorijevic[a,b], Jelena Stojanovic[a,b], Zoran Obradovic[a]

[a] *Center for Data Analytics and Biomedical Informatics, Temple University, Philadelphia, PA 19122 USA*
[b] *Authors contributed equally*

Abstract

Data-driven phenotype discoveries on Electronic Health Records (EHR) data have recently drawn benefits across many aspects of clinical practice. In the method described in this paper, we map a very large EHR database containing more than a million inpatient cases into a low dimensional space where diseases with similar phenotypes have similar representation. This embedding allows for an effective segmentation of diseases into more homogeneous categories, an important task of discovering disease types for precision medicine. In particular, many diseases have heterogeneous nature. For instance, sepsis, a systemic and progressive inflammation, can be caused by many factors, and can have multiple manifestations on different human organs. Understanding such heterogeneity of the disease can help in addressing many important issues regarding sepsis, including early diagnosis and treatment, which is of huge importance as sepsis is one of the main causes of in-hospital deaths in the United States. This study analyzes state of the art embedding models that have had huge success in various fields, applying them to disease embedding from EHR databases. Particular interest is given to learning multi-type representation of heterogeneous diseases, which leads to more homogeneous groups. Our results show evidence that such representations have phenotypes of higher quality and also provide benefit when predicting mortality of inpatient visits.

Keywords: EHR databases, Phenotyping, Neural embedding models, Sepsis, Discovery of disease types

ACCEPTED MANUSCRIPT

Download English Version:

https://daneshyari.com/en/article/5513498

Download Persian Version:

https://daneshyari.com/article/5513498

Daneshyari.com