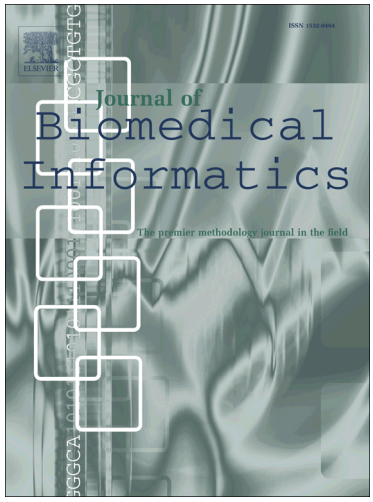
Accepted Manuscript



Embedding of Semantic Predications

Trevor Cohen, Dominic Widdows

PII: S1532-0464(17)30052-7
DOI: http://dx.doi.org/10.1016/j.jbi.2017.03.003
Reference: YJBIN 2737

To appear in: *Journal of Biomedical Informatics*

Received Date: 13 November 2016
Accepted Date: 5 March 2017

Please cite this article as: Cohen, T., Widdows, D., Embedding of Semantic Predications, *Journal of Biomedical Informatics* (2017), doi: http://dx.doi.org/10.1016/j.jbi.2017.03.003

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

Embedding of Semantic Predications

Trevor Cohen[a,*], Dominic Widdows[b]

[a]*School of Biomedical Informatics, The University of Texas Health Science Center, Houston, TX, United States*
[b]*Grab, Inc.*

Abstract

This paper concerns the generation of distributed vector representations of biomedical concepts from structured knowledge, in the form of subject-relation-object triplets known as semantic predications. Specifically, we evaluate the extent to which a representational approach we have developed for this purpose previously, known as Predication-based Semantic Indexing (PSI), might benefit from insights gleaned from neural-probabilistic language models, which have enjoyed a surge in popularity in recent years as a means to generate distributed vector representations of terms from free text. To do so, we develop a novel neural-probabilistic approach to encoding predications, called Embedding of Semantic Predications (ESP), by adapting aspects of the Skipgram with Negative Sampling (SGNS) algorithm to this purpose. We compare ESP and PSI across a number of tasks including recovery of encoded information, estimation of semantic similarity and relatedness, and identification of potentially therapeutic and harmful relationships using both analogical retrieval and supervised learning. We find advantages for ESP in some, but not all of these tasks, revealing the contexts in which the additional computational work of neural-probabilistic modeling is justified.

Keywords: semantic predications, distributional semantics, word embeddings, predication-based semantic indexing, literature-based discovery, pharmacovigilance

*Corresponding author
Email address: trevor.cohen@uth.tmc.edu (Trevor Cohen)

ACCEPTED MANUSCRIPT

Download English Version:

https://daneshyari.com/en/article/4966902

Download Persian Version:

https://daneshyari.com/article/4966902

Daneshyari.com