

Accepted Manuscript

Benchmarking Deep Learning Models on Large Healthcare Datasets

Sanjay Purushotham, Chuizheng Meng, Zhengping Che, Yan Liu

PII: S1532-0464(18)30071-6
DOI: <https://doi.org/10.1016/j.jbi.2018.04.007>
Reference: YJBIN 2964

To appear in: *Journal of Biomedical Informatics*

Received Date: 17 October 2017
Revised Date: 22 March 2018
Accepted Date: 9 April 2018



Please cite this article as: Purushotham, S., Meng, C., Che, Z., Liu, Y., Benchmarking Deep Learning Models on Large Healthcare Datasets, *Journal of Biomedical Informatics* (2018), doi: <https://doi.org/10.1016/j.jbi.2018.04.007>

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.

Benchmarking Deep Learning Models on Large Healthcare Datasets

Sanjay Purushotham^{a,*}, Chuizheng Meng^{b,*}, Zhengping Che^a, Yan Liu^{a,**}

^aUniversity of Southern California, Los Angeles, CA 90089, US

^bTsinghua University, Beijing 100084, China

Abstract

Deep learning models (aka Deep Neural Networks) have revolutionized many fields including computer vision, natural language processing, speech recognition, and is being increasingly used in clinical healthcare applications. However, few works exist which have benchmarked the performance of the deep learning models with respect to the state-of-the-art machine learning models and prognostic scoring systems on publicly available healthcare datasets. In this paper, we present the benchmarking results for several clinical prediction tasks such as mortality prediction, length of stay prediction, and ICD-9 code group prediction using Deep Learning models, ensemble of machine learning models (Super Learner algorithm), SAPS II and SOFA scores. We used the Medical Information Mart for Intensive Care III (MIMIC-III) (v1.4) publicly available dataset, which includes all patients admitted to an ICU at the Beth Israel Deaconess Medical Center from 2001 to 2012, for the benchmarking tasks. Our results show that deep learning models consistently outperform all the other approaches especially when the ‘raw’ clinical time series data is used as input features to the models.

Keywords: deep learning models, super learner algorithm, mortality prediction, length of stay, ICD-9 code group prediction

1. Introduction

Quantifying patient health and predicting future outcomes is an important problem in critical care research. Patient mortality and length of hospital stay are the most important clinical outcomes for an ICU admission, and accurately predicting them can help with the assessment of severity of illness; and determining the value of novel treatments, interventions and health care policies. With the goal of accurately predicting these clinical outcomes, researchers have developed novel machine learning models [1, 2] and scoring systems [3] while measuring the

*Co-first authors.

**Corresponding author.

Email addresses: spurusho@usc.edu (Sanjay Purushotham), mengcz95thu@gmail.com (Chuizheng Meng), zche@usc.edu (Zhengping Che), yanliu.cs@usc.edu (Yan Liu)

Download English Version:

<https://daneshyari.com/en/article/6927435>

Download Persian Version:

<https://daneshyari.com/article/6927435>

[Daneshyari.com](https://daneshyari.com)