

Accepted Manuscript

A Study of the Suitability of Autoencoders for Preprocessing Data in Breast Cancer Experimentation

Laura Macías-García, José María Luna-Romera, Jorge García-Gutiérrez, María del Mar Martínez-Ballesteros, José C. Riquelme-Santos, Ricardo González-Cámpora

PII: S1532-0464(17)30147-8
DOI: <http://dx.doi.org/10.1016/j.jbi.2017.06.020>
Reference: YJBIN 2816

To appear in: *Journal of Biomedical Informatics*

Received Date: 21 November 2016
Revised Date: 19 May 2017
Accepted Date: 25 June 2017

Please cite this article as: Macías-García, L., Luna-Romera, J.M., García-Gutiérrez, J., Mar Martínez-Ballesteros, M.d., Riquelme-Santos, J.C., González-Cámpora, R., A Study of the Suitability of Autoencoders for Preprocessing Data in Breast Cancer Experimentation, *Journal of Biomedical Informatics* (2017), doi: <http://dx.doi.org/10.1016/j.jbi.2017.06.020>

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.



A Study of the Suitability of Autoencoders for Preprocessing Data in Breast Cancer Experimentation

Laura Macías-García^a, José María Luna-Romera^b, Jorge García-Gutiérrez^b,
María del Mar Martínez-Ballesteros^b, José C. Riquelme-Santos^b, Ricardo
González-Cámpora^c

^a*Department of Anatomic Pathology. Hospital Infanta Elena, Huelva, Spain.*

^b*Department of Computer Languages and Systems, ETSII, University of Seville, Spain.*

^c*Department of Anatomic Pathology, Hospital Universitario Virgen Macarena. Seville, Spain.*

Abstract

Breast cancer is the most common cause of cancer death in women. Today, post-transcriptional protein products of the genes involved in breast cancer can be identified by immunohistochemistry. However, this method has problems arising from the intra-observer and inter-observer variability in the assessment of pathologic variables, which may result in misleading conclusions. Using an optimal selection of preprocessing techniques may help to reduce observer variability. Deep learning has emerged as a powerful technique for any tasks related to machine learning such as classification and regression. The aim of this work is to use autoencoders (neural networks commonly used to feed deep learning architectures) to improve the quality of the data for developing immunohistochemistry signatures with prognostic value in breast cancer. Our testing on data from 222 patients with invasive non-special type breast carcinoma shows that an automatic binarization of experimental data after autoencoding could outperform other classical preprocessing techniques (such as human-dependent or automatic binarization only) when applied to

Download English Version:

<https://daneshyari.com/en/article/4966775>

Download Persian Version:

<https://daneshyari.com/article/4966775>

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