Journal of Statistical Software, vol.102, no.7, pp.1-36, April, 2022.

NeuralSens: sensitivity analysis of neural networks

J. Pizarroso Gonzalo; J. Portela González; A. Muñoz San Roque

Abstract-

This article presents the NeuralSens package that can be used to perform sensitivity analysis of neural networks using the partial derivatives method. The main function of the package calculates the partial derivatives of the output with regard to the input variables of a multi-layer perceptron model, which can be used to evaluate variable importance based on sensitivity measures and characterize relationships between input and output variables. Methods to calculate partial derivatives are provided for objects trained using common neural network packages in R, and a 'numeric' method is provided for objects from packages which are not included. The package also includes functions to plot the information obtained from the sensitivity analysis. The article contains an overview of techniques for obtaining information from neural network models, a theoretical foundation of how partial derivatives are calculated, a description of the package functions, and applied examples to compare NeuralSens functions with analogous functions from other available R packages.

Index Terms- neural networks, sensitivity, analysis, variable importance, R, NeuralSens.

Citation:
Pizarroso, J.; Portela, J.; Muñoz, A. "NeuralSens: sensitivity analysis of neural networks", Journal of Statistical Software, vol.102, no.7, pp.1-36, April, 2022.