

Study of a well-known importance measure computed via decision trees

Erwan Scornet

Nowadays, machine learning procedures are used in many fields with the notable exception of so-called sensitive areas (health, justice, defense, to name a few) in which decisions to be taken are fraught with consequences. In these fields, it is necessary to obtain a precise decision but, to be effectively applied, these algorithms must provide an explanation of the mechanisms that lead to the decision and, in this sense, be interpretable. Unfortunately, the most accurate algorithms today are often the most complex. A classic technique to try to explain their predictions is to calculate indicators corresponding to the strength of the dependence between each input variable and the output to be predicted. In this talk, we will focus on one measure of importance created for decision trees and we will see how the theoretical study provides explanations on its practical use.