Biased Sampling: Solution for Lower Incidence Rate

M Muthu Mangai

Abstract

Given the lower incidence rate, use of decision tree techniques like Classification and Regression Tree (CART) in understanding credit or operational risk becomes quite challenging. A commonly adopted solution is biased sampling approach, where more weights are attached to bad customers to artificially hike the incidence or bad rate. While adopting this type of biased sampling approach, question of the best weight arises. This paper adopts an iterative approach in identifying the best weight. The best weight for bankruptcy (BKO) profiling problem in hand occurred when the incidence rate was around 50 percent where entropy reaches its maximum.