

Abstract

Dr. Yasser A Yakout Shehata

A COMPUTATIONAL APPROACH TO CORRECTLY ASSESS SIGNIFICANCE IN BEST SUBSET REGRESSION

Best subsets regression is often used to identify a good regression model. The standard approach to assess statistical significance for a best subsets regression model is flawed. A computationally intensive randomization algorithm which corrects the problem is outlined and implemented. Simulation studies show that this procedure corrects a non-trivial problem that exists independent of sample size; is a procedure that is robust to the presence of influential observations. This procedure leads to a simple decision rule even with correlated predictors unlike the use of a single probe. The proposed method is shown to retain power in a non-null situation.