Sample-Efficient Regression Tree and Its Applications to Semiconductor Yield Learning

Argon Chen
Graduate Institute of Industrial Engineering, National Taiwan University, Taiwan

Amos Hong
Odey Ho
Chao-Wen Liu
Yi-His Huang
Department of Mechanical Engineering, National Taiwan University, Taiwan

Regression trees have been known to be an effective data mining tool for semiconductor yield analysis. The regression tree is built by iteratively splitting data set and selecting factors into a hierarchical tree model. The fundamental assumption is that each tree node with a data subset in the hierarchy has its own governing model and the entire tree consists of governing models in different hierarchical levels. There are mainly two shortcomings of regression trees. One is that the sample size reduces sharply after few levels of data splitting and factor selections. With sample size depletion, variable selection and effect inference in the lower hierarchical levels becomes extremely unreliable. For example, suppose the yield loss is caused by a combination effect of four semiconductor tools (represented by four binary variables x1,K,x4) with only 20 lots of wafers available for analysis. The conventional regression tree depletes the sample size quickly and could fail to select all four variables into the model. The other shortcoming is the over-fitting problem due to empirical stopping rules. An over-fitted model selects too many noise factors into the model. These shortcomings make the conventional decision tree an inappropriate analysis tool for yield ramp-up and foundry fabrication where the number of available lots for analysis is usually small. In contrast, the forward regression analysis selects the influential factors using the entire data set with rigorous criteria. However, the forward regression is not capable of splitting data to subsets with different underlying models. In this research, we attempt to develop a Sample-Efficient Regression Tree (SERT) that not only combines the forward regression and regression tree methodologies but also selects combination effects in a much more efficient manner. Semiconductor yield learning example will be given to demonstrate the proposed methodology.

[Argon Chen, 1 Roosevelt Rd. Sec. 4, Taipei 106, Taiwan; achen@ntu.edu.tw]