

**CHENNAI – PONDICHERRY**

**Ranking and Clustering Software Cost Estimation Models through a Multiple Comparisons Algorithm**

**Abstract**

Software Cost Estimation can be described as the process of predicting the most realistic effort required to complete a software project. Due to the strong relationship of accurate effort estimations with many crucial project management activities, the research community has been focused on the development and application of a vast variety of methods and models trying to improve the estimation procedure. From the diversity of methods emerged the need for comparisons to determine the best model. However, the inconsistent results brought to light significant doubts and uncertainty about the appropriateness of the comparison process in experimental studies. Overall, there exist several potential sources of bias that have to be considered in order to reinforce the confidence of experiments. In this paper, we propose a statistical framework based on a multiple comparisons algorithm in order to rank several cost estimation models, identifying those which have significant differences in accuracy, and clustering them in nonoverlapping groups. The proposed framework is applied in a large-scale setup of comparing 11 prediction models over six datasets. The results illustrate the benefits and the significant information obtained through the systematic comparison of alternative methods.