AHMED FARAHAT, University of Waterloo

An Efficient Greedy Method for Unsupervised Variable Selection

The selection of relevant variables is a crucial task in data analysis. This work proposes a novel method for unsupervised variable selection, which efficiently selects variables in a greedy manner. We first define an effective criterion for unsupervised variable selection which measures the reconstruction error of the data matrix based on the selected variables, and then present a novel algorithm for greedily minimizing the reconstruction error based on the variables selected so far. Experiments on real datasets demonstrate the effectiveness of the proposed algorithm in comparison to the state-of-the-art methods for unsupervised variable selection.