

Padmore Nana Prempeh

UTEP

TITLE: Adaptive estimation of the Tuning Parameters of the Structured Screen-and-Select Variable Selection (S3VS) Framework

Abstract:

This report evaluates the Structured Screen-and-Select Variable Selection (S3VS) method for variable selection in ultra-high dimensional data. The S3VS method is tested with two variable selection techniques, LASSO and NLP-MOM, and is applied to simulated datasets with varying sample sizes and numbers of variables. The performance of the method is assessed based on metrics like True Positive Rate (TPR) and False Discovery Rate (FDR). We introduce novel approaches for determining leading variables and leading sets, including the Fixed Correlation Threshold and Percentile-based Correlation Threshold. The results demonstrate that these methods are particularly effective in high-dimensional settings, where traditional methods often struggle, and offer flexible and robust solutions for variable selection in complex datasets. The findings also suggest that the S3VS method scales well to ultra-high dimensional data, making it a promising tool for future applications in genomics and other fields dealing with large-scale data.