

ABSTRACT

Robust Principal Component Regression with Wild Bootstrap for Handling Outliers, Multicollinearity, and Heteroskedasticity in Chronic Hepatitis B Data

By

Joseph Dedek Parhusip

This study aims to analyze the performance of the Robust Principal Component (RPC) method combined with Wild Bootstrap in handling outliers, multicollinearity, and heteroskedasticity in chronic hepatitis B data. The data were obtained from the World Health Organization (WHO) and consist of several epidemiological indicators. The analysis methods include Principal Component Analysis (PCA), Least Trimmed Squares (LTS), and Wild Bootstrap using Wu and Liu multipliers. The results show that the dataset contains outliers, strong multicollinearity, and heteroskedasticity. The RPC-Wild Bootstrap method produces more stable parameter estimates, with RPC Boot Wu showing lower standard error and RMSE compared to RPC Boot Liu. Therefore, the RPC-Wild Bootstrap method is effective in producing more stable and reliable parameter estimates for complex real-world data.

Keywords: Robust Principal Component, Wild Bootstrap, Outliers, Multicollinearity, Heteroskedasticity.

ABSTRAK

IMPLEMENTASI *ROBUST PC* DENGAN *WILD BOOTSTRAP* DALAM MENGATASI *OUTLIER*, MULTIKOLINEARITAS, DAN HETEROSKEDASTISITAS PADA DATA HEPATITIS B KRONIS

Oleh

Joseph Dedek Parhusip

Penelitian ini bertujuan untuk menganalisis kinerja metode *Robust Principal Component (RPC)* yang dikombinasikan dengan *Wild Bootstrap* dalam menangani pencilan (*outlier*), multikolinearitas, dan heteroskedastisitas pada data hepatitis B kronis. Data diperoleh dari *World Health Organization (WHO)* dan terdiri atas beberapa indikator epidemiologis. Metode analisis yang digunakan meliputi *Principal Component Analysis (PCA)*, *Least Trimmed Squares (LTS)*, dan *Wild Bootstrap* dengan pengali Wu dan Liu. Hasil penelitian menunjukkan bahwa dataset mengandung pencilan, multikolinearitas yang kuat, dan heteroskedastisitas. Metode *RPC Wild Bootstrap* menghasilkan estimasi parameter yang lebih stabil, dengan *RPC Boot Wu* menunjukkan nilai *standard error* dan *RMSE* yang lebih rendah dibandingkan *RPC Boot Liu*. Oleh karena itu, metode *RPC Wild Bootstrap* efektif dalam menghasilkan estimasi parameter yang lebih stabil dan andal untuk data dunia nyata yang kompleks.

Kata kunci : *Robust Principal Component, Wild Bootstrap, Pencilan (Outlier), Multikolinearitas, Heterokedastisitas.*