

INTISARI

ROBUST PRINCIPAL COMPONENT ANALYSIS (ROBPCA) **UNTUK DATA BERDIMENSI BESAR DENGAN OUTLIERS**

Oleh:

Tri Rokhmatun Sobiroh

11/319872/PA/14299

Principal Component Analysis (PCA) merupakan salah satu analisis multivariate yang berguna untuk mereduksi data berdimensi besar yang memiliki korelasi sehingga membentuk data dengan variabel yang lebih sedikit tetapi tetap mampu menjelaskan keragaman dari data. Perkembangan PCA selanjutnya dipengaruhi adanya kebutuhan suatu model PCA yang robust terhadap data pencilan / *outlier*. Oleh karena itu dikenalkanlah suatu konsep *robust PCA*. Tujuan dari metode *robust PCA* adalah untuk memperoleh komponen utama yang tidak terlalu banyak terpengaruh dengan keberadaan *outlier*.

Pada skripsi ini akan dipelajari mengenai ROBPCA yang dikembangkan oleh Hubert dkk (2005) dengan menggabungkan konsep *Projection Pursuit (PP)* dan estimator kovarian yang robust yang dalam hal ini digunakan *Minimum Covariance Determinant (MCD)*. Hasil analisis selanjutnya dibandingkan dengan *Classic Principal Component Analysis (PCA)*. Dalam studi kasus ini diperoleh kesimpulan bahwa ROBPCA lebih baik dari CPCA karena metode ROBPCA mampu menghasilkan jumlah komponen utama yang lebih sedikit daripada metode CPCA dan telah dapat menjelaskan 84,79% dari total variansi sampel.

Kata kunci : *outlier*, CPCA, ROBPCA, komponen utama, *robust*

ABSTRACT

ROBUST PRINCIPAL COMPONENT ANALYSIS (ROBPCA) FOR HIGH DIMENSIONAL DATA WITH OUTLIERS

By

Tri Rokhmatun Sobiroh

11/319872/PA/14299

Principal Component Analysis (PCA) is a multivariate analysis that is useful for reducing the high dimensional data of which have a correlation to form the data with less variables but still able to explain the diversity of the data. Development of PCA influenced by robust PCA models for data with outliers. Therefore introduced a robust concept of PCA. The purpose of robust PCA method is to obtain the main component that is not too much affected by the presence of outliers.

In this thesis will learn about ROBPCA developed by Hubert et al (2005) to incorporate the concept of Projection Pursuit (PP) and a robust covariance estimator in this case used the Minimum Covariance Determinant (MCD). Then, the results of the analysis compared by the Classic Principal Component Analysis (PCA). In the case study, concluded that ROBPCA better than CPCA because ROBPCA method is able to produce the amount of less principal components than CPCA method and able to explain 84.79% of the total variance of the sample.

Key words : outlier, CPCA, ROBPCA, principal component, robust