

BAB I

PENDAHULUAN

1.1. Latar Belakang

Principal Component Analysis (PCA) atau analisis komponen utama merupakan salah satu analisis multivariat yang digunakan untuk mereduksi dimensi data dari yang berukuran besar dan saling berkorelasi menjadi dimensi yang lebih kecil dan tidak saling berkorelasi. Namun walaupun dimensi data menjadi lebih kecil, kita tidak akan kehilangan banyak informasi karena variasi data tetap dipertahankan minimal 80% (Johnson dan Wichern, 1992: 359).

Analisis komponen utama dilakukan untuk memperoleh komponen-komponen utama yang dapat menjelaskan sebagian besar variasi data. Komponen pertama memiliki persentase terbesar daripada komponen-komponen lainnya dan akan menjelaskan variasi terbesar data, komponen utama kedua memiliki persentase terbesar kedua dari variasi data, dan seterusnya.

Perkembangan analisis komponen utama dimulai sejak diperkenalkan pertama kali oleh Pearson pada tahun 1901. Sejalan dengan perkembangan teknologi komputer dan kemajuan di bidang matematika, analisis komponen utama hingga kini masih terus mengalami perkembangan. Perkembangan selanjutnya, diperkenalkan generalisasi dari analisis komponen utama oleh Loève pada tahun 1963. Perkembangan analisis komponen utama selanjutnya dipengaruhi adanya kebutuhan model analisis komponen utama yang robust terhadap data yang mengandung *outlier* (pencilan). Analisis komponen utama

klasik sangat rentan terhadap kehadiran pencilan karena didasarkan pada matriks varians kovarians yang juga sangat sensitif terhadap keberadaan data pencilan, terutama jika data mengandung pencilan yang ekstrim (Suryana, 1998).

Pendeteksian pencilan merupakan langkah yang sangat penting dalam suatu analisis data, karena pencilan dapat mempengaruhi keseimbangan data. Mendeteksi pencilan pada sampel data univariat dapat dilakukan melalui gambar atau uji statistik sederhana. Pada sampel bivariat, pendeteksian pencilan sudah mulai sulit dan uji statistik yang digunakan akan lebih rumit. Hal ini menjadi lebih sulit lagi jika sampel data yang digunakan berdimensi tinggi atau multivariat. Untuk mengatasi pencilan ini dibutuhkan metode penaksir robust yang tangguh terhadap pencilan sehingga analisis komponen utama robust merupakan metode komponen utama yang tidak lagi dipengaruhi pencilan.

Menurut Jackson (1991: 365), pada kasus analisis komponen utama terdapat dua cara yang sering digunakan para peneliti untuk memperoleh estimasi robust, pertama dengan menghasilkan estimasi robust pada matriks korelasi atau varians kovarians yang digunakan pada analisis komponen utama. Setelah itu, analisis komponen utama klasik dapat dilakukan. Kedua dengan cara menghasilkan estimasi robust pada akar dan vektor eigennya, yang beberapa hasilnya dapat digunakan untuk memperoleh estimasi robust pada matriks korelasi atau varians kovarians.

Karena kini sudah banyak metode lain, maka alternatif tersebut tidak harus selalu digunakan. Pada tahun 1985, Rousseeuw memperkenalkan metode *Minimum Volume Ellipsoid* (MVE) (Rousseeuw, Van Aelst, 2009: 71) sebagai

penaksir robust. MVE merupakan penaksir robust dengan *high breakdown* pertama pada lokasi dan *scatter* multivariat yang sering digunakan. MVE terkenal karena kepekaannya terhadap pencilan yang membuatnya dapat diandalkan untuk mendeteksi pencilan. Metode ini sangat berguna dalam mendeteksi pencilan pada data multivariat, karena memenuhi jumlah batas maksimum pencilan yaitu 50%.

Di Indonesia masalah ibu hamil dan kesehatan bayi baru lahir masih merupakan masalah kesehatan masyarakat yang perlu memperoleh perhatian khusus karena masih banyak kasus kematian bayi yang baru lahir. Ibu hamil yang menderita status gizi kurang mempunyai resiko melahirkan bayi dengan kematian saat persalinan, pendarahan, pasca persalinan yang sulit karena lemah dan mudah mengalami gangguan kesehatan kesehatan ibu ketika hamil. Oleh karena itu, dilakukan pengamatan terhadap status gizi ibu hamil dan keadaan bayi saat dilahirkan. Pencilan yang terdapat di dalam data tersebut mungkin saja dapat mempengaruhi variasi data.

Berdasarkan penjelasan tersebut, penulis tertarik dengan konsep yang diperkenalkan oleh Rousseeuw dan diterapkan pada data yang penulis peroleh, sehingga tugas akhir ini diberi judul **PENERAPAN METODE *MINIMUM VOLUME ELLIPSOID* PADA ANALISIS KOMPONEN UTAMA ROBUST.**

1.2. Rumusan Masalah

Berdasarkan latar belakang yang telah penulis kemukakan, rumusan masalah yang akan dikaji dalam tugas akhir ini adalah:

1. Bagaimana cara mendeteksi pencilan pada analisis komponen utama menggunakan metode *minimum volume ellipsoid* ?

2. Bagaimana cara menangani pencilan menggunakan analisis komponen utama robust?
3. Bagaimana penerapan analisis komponen utama robust dengan pendeteksian pencilan menggunakan metode *minimum volume ellipsoid* pada data Bayi Baru Lahir di Puskesmas Mergangsan, Yogyakarta untuk memperoleh sub ruang yang optimal?
4. Bagaimana keuntungan menggunakan *Weighted PCA*?

1.3. Tujuan penulisan

Berdasarkan rumusan masalah yang telah diuraikan, tujuan penulisan tugas akhir ini adalah untuk:

1. Mengetahui cara mendeteksi pencilan pada analisis komponen utama menggunakan metode *minimum volume ellipsoid*.
2. Mengetahui cara menangani pencilan menggunakan analisis komponen utama robust.
3. Menerapkan analisis komponen utama robust dengan pendeteksian pencilan menggunakan metode *minimum volume ellipsoid* pada data Bayi Baru Lahir di Puskesmas Mergangsan, Yogyakarta untuk memperoleh sub ruang yang optimal.
4. Bagaimana keuntungan menggunakan *Weighted PCA*?

1.4. Manfaat Penulisan

1. Manfaat Teoritis

Dapat memperkaya ilmu pengetahuan dan memahami pengembangan konsep statistika multivariat, khususnya analisis komponen utama.

2. Manfaat Praktis

Dapat menerapkan dan mengembangkan metode-metode robust pada data yang mengandung pencilan baik pada analisis komponen utama maupun analisis multivariat lainnya.

1.5. Batasan Masalah

Dalam tugas akhir ini penulis membatasi masalah pada data yang akan digunakan yaitu data dengan jumlah variabel kurang dari jumlah observasi ($p < n$). Selain itu, pendeteksian pencilan yang dibandingkan adalah metode MVE dengan pendekatan jarak Mahalanobis.

1.6. Sistematika Penulisan

BAB I : PENDAHULUAN

Bab ini membahas latar belakang masalah, rumusan masalah, tujuan masalah, manfaat penulisan, batasan permasalahan, dan sistematika serta metodologi penulisan.

BAB II : LANDASAN TEORI

Bab ini berisi teori-teori pendukung masalah yang dikaji dalam tugas akhir ini dan diambil dari berbagai sumber yang relevan.

BAB III : PEMBAHASAN

Bab ini menjelaskan materi analisis komponen utama robust dan pendeteksian pencilan dengan metode *Minimum Volume Ellipsoid* yang akan penulis kaji berdasarkan teori-teori pada Bab II.

BAB IV : STUDI KASUS

Bab ini membahas penerapan materi analisis komponen utama robust menggunakan metode *Minimum Volume Ellipsoid* pada data Bayi Baru Lahir di Puskesmas Mergangsan, Yogyakarta.

BAB V : KESIMPULAN DAN SARAN

Bab ini menjelaskan tentang kesimpulan dan saran dari keseluruhan materi dan pengolahan data dalam penyusunan tugas akhir ini.

1.7. Metode Penulisan

Metode penulisan dalam tugas akhir ini menggunakan berbagai sumber, yaitu beberapa buku penunjang, jurnal, dan data untuk diaplikasikan menggunakan metode MVE.

