

Implementasi algoritma *principal component analysis* dalam reduksi faktor yang mempengaruhi produksi susu sapi perah

Anindya Luthfiani Susetyo^{1*}, Dian Maharani²

^{1,2} Program Studi Matematika, Universitas Islam Negeri Maulana Malik Ibrahim Malang
e-mail: *200601110085@student.uin-malang.ac.id

Kata Kunci:

reduksi; statistika; koperasi; susu; analisis komponen utama

Keywords:

reduction; statistic; cooperation; milk; principal component analysis

ABSTRAK

Banyak faktor yang dapat mempengaruhi produksi susu sapi yaitu pakan yang diberikan, teknologi yang digunakan peternak, dan jumlah sapi yang dimiliki peternak. Karena banyaknya faktor eksternal yang dapat mempengaruhi jumlah produksi susu, maka perlu dilakukan reduksi variabel yang kurang berpengaruh secara signifikan agar pihak manajemen dapat fokus pada faktor yang memiliki pengaruh paling besar, dalam hal reduksi variabel, metode statistik *principal component analysis* (PCA) dapat digunakan. Metode ini digunakan untuk menyederhanakan variabel-variabel yang

ada di dalam dataset dengan cara mereduksi dimensinya sehingga memudahkan dalam menginterpretasikan data. Reduksi variabel dengan menggunakan PCA menghasilkan dua komponen baru dengan menggunakan tiga metode yaitu *cumulative sum of proportions*, *eigenvalue* dan *scree plot*. Setelah dilakukan analisis PCA dengan menggunakan tiga metode, diperoleh dua komponen baru yaitu PC1 dan PC2. Satu variabel yang masuk dalam komponen baru PC1 adalah X6 karena memiliki nilai lebih dari 0,5 dan terdapat dua variabel yang masuk dalam komponen baru PC2, yaitu X4 dan X5.

ABSTRACT

Many factors can affect the milk production of cows, namely the feed given, the technology used by farmers, and the number of cows owned by farmers. Due to the large number of external factors that can affect the amount of milk production, it is necessary to reduce variables that have less significant influence so that management can focus on the factors that have the most influence, in terms of variable reduction, the *principal component analysis* (PCA) statistical method can be used. This method is used to simplify the variables in the dataset by reducing their dimensions so that it is easier to interpret the data. Variable reduction using PCA produces two new components using three methods, namely *cumulative sum of proportions*, *eigenvalue* and *scree plot*. After PCA analysis using three methods, two new components were obtained, namely PC1 and PC2. One variable included in the new PC1 component is X6 because it has a value of more than 0.5 and there are two variables included in the new PC2 component, namely X4 and X5.

Pendahuluan

Peternakan merupakan bidang usaha di Indonesia yang bergerak dalam pembibitan serta budidaya hewan ternak. Terdapat banyak macam hewan yang dapat



This is an open access article under the [CC BY-NC-SA](https://creativecommons.org/licenses/by-nc-sa/4.0/) license.

Copyright © 2023 by Author. Published by Universitas Islam Negeri Maulana Malik Ibrahim Malang.

dibudidayakan sebagai hewan ternak, salah satunya adalah sapi. Hewan ini banyak digunakan sebagai hewan ternak tidak hanya di Indonesia karena beragam manfaat yang dimiliki utamanya adalah olahan daging dan susu yang dimanfaatkan sebagai sumber pangan manusia. Susu segar hasil produksi susu sapi perah memiliki nutrisi tinggi yang dapat digunakan sebagai asupan bergizi bagi tubuh manusia. Kadar protein yang dimiliki oleh susu segar sekitar 3,5% dan kadar lemak sekitar 3,0% - 3,8% yang baik bagi perkembangan dan pertumbuhan anak-anak (Susilawati et al., 2021).

Salah satu tempat usaha yang mendukung produksi sapi adalah koperasi. Koperasi adalah suatu sarana dengan tujuan untuk meningkatkan kualitas kesejahteraan bagi para anggota koperasi khususnya bagi masyarakat dengan ekonomi menengah ke bawah (Handayani & Khasanah, 2021). Banyak koperasi yang menjadikan susu sapi perah sebagai produk usaha utama, salah satunya adalah Koperasi Agro Niaga Jabung Syariah Jawa Timur yang terletak di Jawa Timur. Koperasi ini berperan sebagai pengumpul susu sapi perah yang didapatkan dari peternak sapi di sekitar koperasi yang juga merupakan anggota koperasi.

Di Indonesia sendiri jumlah sapi potong lebih banyak daripada jumlah sapi perah. Menurut data dari Badan Pusat Statistik (BPS) Indonesia, tahun 2022 Indonesia memiliki sapi potong dengan jumlah sebesar 37 juta ekor dan sapi perah sebesar 500 ribu ekor. Keduanya memiliki ketimpangan jumlah yang cukup besar. Jumlah sapi perah yang sedikit dapat menjadi salah satu faktor yang menyebabkan jumlah produksi susu sapi perah rendah di Indonesia. Negara Indonesia memiliki angka produksi susu sapi perah sebesar 2,8 juta ton pada tahun 2022. Jika dibandingkan dengan Amerika Serikat yang memiliki jumlah hasil produksi sebesar 100 juta ton susu tentunya angka ini memiliki perbedaan yang cukup signifikan.

Faktor lain yang dapat mempengaruhi produksi susu sapi adalah pakan ternak yang diberikan, teknologi yang digunakan oleh peternak, dan jumlah sapi yang dimiliki oleh peternak (Setiawan, 2017). Karena terdapat banyaknya faktor eksternal yang dapat berpengaruh terhadap jumlah produksi susu sapi perah, perlu dilakukan pengurangan variabel yang kurang berpengaruh signifikan agar dapat dilakukan fokus manajemen terhadap faktor yang memiliki pengaruh paling besar.

Pembahasan

Dalam kasus reduksi variabel dapat digunakan metode statistika yaitu *principal component analysis* (PCA). PCA adalah sebuah metode yang pertama kali dikemukakan oleh Beltrami dan Jordan (1874) yang digunakan untuk menganalisis data yang bersifat multivariat. Metode ini menjelaskan komposisi varian kovarian dari Kumpulan variabel yang diperoleh melalui kombinasi linier dari variabel asal (Robert & Brown, 2004). Definisi lain menyebutkan bahwa PCA merupakan metode yang digunakan untuk menganalisis dalam bidang keilmuan statistika dengan peubah ganda yang dapat digunakan untuk mengurangi peubah asal menjadi beberapa variabel baru dengan sifat orthogonal namun tetap mempertahankan varian dari variabel asalnya (Nugroho, 2008).

Pendapat lain menyatakan bahwa metode PCA merupakan salah satu metode dalam analisis multivariat yang digunakan dengan tujuan melakukan kajian terhadap

struktur matriks ragam-peragam atau *covariance matrix* melalui linear variabelnya. Metode ini digunakan untuk menyederhanakan variabel-variabel pada *dataset* dengan cara mengurangi dimensinya agar lebih mudah untuk melakukan interpretasi terhadap data tersebut (Badri & Sari, 2021).

Beberapa hal yang perlu diperhatikan ketika menggunakan metode PCA, antara lain adalah sebagai berikut:

1. PCA bertujuan untuk mengurangi variabel pada data. Ketika dilakukan analisis diusahakan agar fokus pada jumlah minimal yang akan diperlukan untuk menjelaskan jumlah maksimal dari varian total yang diwakili dalam Kumpulan variabel asli.
2. Error varian mewakili jumlah proporsi yang cukup kecil dari total varian.
3. Hasil algoritma PCA digunakan untuk proses pengembangan skala.

Normalisasi Data

Sebelum *dataset* dilakukan proses data, akan dilakukan pra proses data, yaitu normalisasi data. Proses normalisasi data dilakukan pada semua variabel tanpa terkecuali. Dengan dilakukannya proses normalisasi data, variabel-variabel akan dilakukan pengukuran ulang pada nilai-nilai variabel sehingga memiliki arah dan rentang yang sama. Selain itu, tahap ini juga membuat proses data dapat dengan mudah dilakukan dan tidak menyebabkan peningkatan yang besar dalam beban kerja dan kebutuhan daya ketika proses data dilakukan (Inggi Permana, 2022). Adapun variabel-variabel yang digunakan pada normalisasi data adalah sebagai berikut:

Tabel 1. 1 Variabel Penelitian

No.	Variabel	Keterangan
1.	X_1	Umur Peternak
2.	X_2	Populasi Sapi
3.	X_3	Populasi Induk Sapi
4.	X_4	Produksi Susu per Induk
5.	X_5	Nilai Kualitas Susu
6.	X_6	Take Home Pay (THP)

Uji Nilai Matriks Anti-Correlation

Tahap uji nilai matriks anti-correlation dilakukan dengan tujuan untuk mengetahui nilai dari measure of sampling adequacy (MSA). Nilai dari MSA harus bernilai lebih dari 0.5 atau sebesar 0,05. Dikatakan memiliki korelasi apabila hasil dari uji nilai matriks anti-correlation lebih dari 0.05, maka kemudian dapat dilakukan analisis dengan metode principal component analysis (PCA) (Ghozali, 2017).

Variabel – variabel pada dataset setelah dilakukan perhitungan dengan menggunakan bantuan software didapatkan hasil, yaitu:

Tabel 1. 2 Hasil Matrik Anti-Correlation

	X_1	X_2	X_3	X_4	X_5	X_6
X_1	1.00000 ^a	0.05035	0.03717	0.01906	-0.00081	0.021719
X_2	0.05035	1.00000 ^a	0.916975	0.055031	-0.03607	0.542307
X_3	0.03717	0.916975	1.00000 ^a	0.020618	-0.08101	0.56902
X_4	0.01906	0.055031	0.0206182	1.00000 ^a	0.83683	0.48683
X_5	-0.0081	-0.03607	-0.08017	0.836837	1.00000 ^a	0.523059
X_6	0.02171	0.542307	0.569029	0.486839	0.52305	1.00000 ^a

Tabel 1.2 menunjukkan nilai MSA pada kolom dengan tanda “a” memperlihatkan bahwa masing-masing variabel pada dataset tidak memiliki nilai MSA kurang dari 0,5 sehingga hal ini sesuai dengan syarat dari uji nilai matriks anti-correlation yang berarti bahwa variabel-variabel pada dataset dapat dilakukan analisis dengan menggunakan PCA.

Uji Nilai Kaiser Mayer Olkin (KMO)

Pengujian pada nilai KMO bertujuan untuk mengetahui apakah data yang digunakan layak untuk dilakukan analisis dengan menggunakan metode PCA. Angka yang ditunjukkan pada hasil uji nilai KMO memiliki artinya masing-masing. Berikut adalah interpretasi dari nilai KMO:

Tabel 1. 3 Interpretasi Nilai KMO

Nilai KMO	Interpretasi
0.90 – 1.00	Data Sangat Baik
0.80 – 0.90	Data Baik
0.70 – 0.80	Data Agak Cukup
0.60 – 0.70	Data Lebih dari Cukup
0.50 – 0.60	Data Cukup
0.00 – 0.50	Data Tidak Layak

Nilai KMO dengan angka yang relatif kecil perlu dilakukan pertimbangan ulang untuk dilanjutkan analisis atau tidak karena korelasi yang dimiliki antar variabel tidak dapat dijelaskan.

Tabel 1. 4 Hasil Nilai KMO

Variabel	Nilai KMO	MSA
Umur (X_1)	KMO = 0.61	0.58
Jumlah Populasi Sapi (X_2)		0.61
Jumlah Induk Sapi (X_3)		0.55
Produksi Susu per Induk (X_4)		0.63
Kualitas Susu Sapi (X_5)		0.53
Take Home Pay (X_6)		0.75

Pada Tabel 1.4 setelah analisis dilakukan ditunjukkan nilai KMO sebesar 0,61 yang lebih besar dari 0,5 . Hal ini berarti bahwa Kumpulan variabel pada dataset dapat dilakukan analisis perhitungan lebih lanjut. Kemudian pada kolom ketiga ditunjukkan nilai dari MSA masing-masing variabel yang memiliki nilai lebih dari 0,5 yang mana semua variabel pada data penelitian dapat diikuti dalam pengujian selanjutnya.

Uji Nilai Bartlett

Pengujian pada nilai bartlett dilakukan untuk mengetahui homogenitas pada data. Uji nilai bartlett digunakan apabila terdapat banyak koefisien korelasi yang dihasilkan memiliki nilai kurang dari 0,5. Sampel data yang diuji bertujuan untuk melihat apakah data tersebut layak untuk dilakukan analisis berdasarkan nilai KMO dan MSA pada hasil yang didapatkan.

Tabel 1. 5 Nilai Bartlett Test

Bartlett Test of Homogeneity of Variances	
<i>Bartlett's K-squared</i>	181124
<i>Df</i>	5
<i>P-Value</i>	$2.2e - 16$

Tabel 1.5 menunjukkan hasil perhitungan yang dilakukan dengan bantuan software dengan nilai uji bartlett sebesar $2.2e-16$. Nilai ini sesuai dengan syarat uji nilai bartlett yaitu kurang dari 0,5 sehingga dapat disimpulkan bahwa data merupakan data yang homogen.

Reduksi Variabel

Implementasi metode PCA untuk melakukan reduksi variabel untuk menentukan jumlah komponen utama umumnya terdapat tiga metode yang dapat digunakan, yaitu:

- Jumlah komponen utama yang ditentukan dari jumlah kumulatif proporsi varians total. Metode ini dapat digunakan pada matriks varian-kovarians dan matriks korelasi.

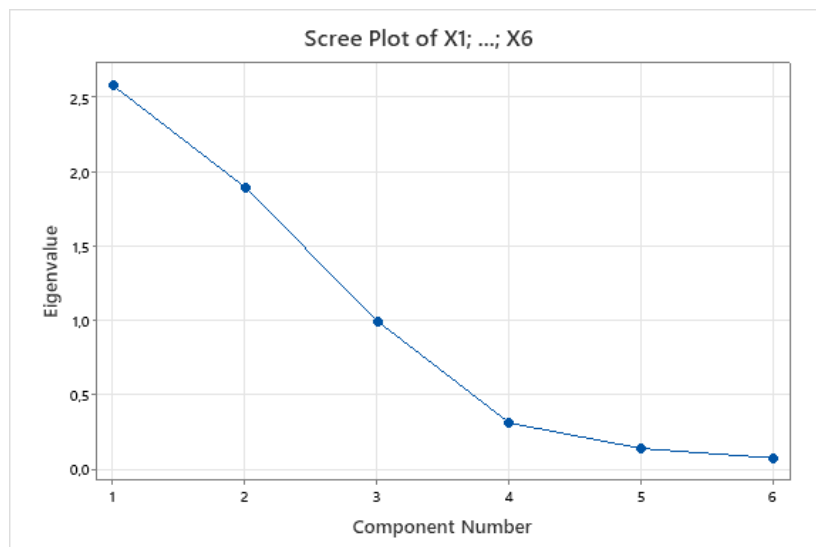
Tabel 1. 6 Hasil Eigen Analysis

<i>Eigenanalysis of the Correlation Matrix</i>						
Eigenvalue	2,5806	1,8961	0,9973	0,3113	0,1396	0,0752
Proportion	0,430	0,316	0,166	0,052	0,023	0,013
Cumulative	0,430	0,746	0,912	0,964	0,987	1,000

Berdasarkan hasil analisis pada Tabel 1.6 terlihat bahwa nilai dari kumulatif proporsi mencakup minimal 75% dari data asli adalah dua komponen.

- Metode kedua didasarkan pada nilai eigen dan hanya dapat digunakan pada matriks korelasi. Nilai eigen yang digunakan adalah nilai eigen yang lebih dari satu, sehingga dapat dikatakan dari Tabel 1.6 bahwa terdapat dua komponen yang memenuhi persyaratan.
- Menggunakan scree plot. Metode ini dapat digunakan pada matriks korelasi maupun pada matriks varian-kovarians.

Gambar 1. 1 Scree Plot



Pengambilan kesimpulan yang dilakukan dari hasil visualisasi data menggunakan scree plot bersifat subyektif (Nugroho, 2008). Jika dilihat dari Gambar 1.1 menunjukkan bahwa plot mulai mendatar pada titik kedua. Maka dapat diketahui dari scree plot ini bahwa komponen utama yang terbentuk berjumlah dua.

Kesimpulan

1. Setelah dilakukan analisis PCA dengan menggunakan tiga metode, didapatkan dua komponen baru yaitu PC1 dan PC2 seperti pada Tabel 1.7.

Tabel 1. 7 Komponen Baru

Variable	PC1	PC2	PC3	PC4	PC5	PC6
X_1	0,037	-0,033	-0,998	-0,024	0,005	-0,010
X_2	0,462	-0,443	0,019	0,338	0,293	0,624
X_3	0,456	-0,466	0,035	0,172	-0,058	-0,735
X_4	0,380	0,514	-0,019	0,512	-0,569	0,074
X_5	0,351	0,566	0,000	0,001	0,720	-0,197
X_6	0,556	0,044	0,035	-0,770	-0,262	0,159

Tabel 1.7 menunjukkan satu variabel yang termasuk ke dalam komponen baru PC1 yaitu X_6 karena memiliki nilai lebih dari 0,5 dan terdapat dua variabel yang termasuk ke dalam komponen baru PC2 yaitu X_4 dan X_5 .

2. Terdapat tiga variabel data asli yang direduksi yaitu umur peternak, populasi sapi dan populasi induk sapi.

Saran

Saran untuk penelitian selanjutnya yaitu dapat dengan menambahkan jumlah variabel yang akan dilakukan penelitian dengan menggunakan metode penelitian yang berbeda sehingga hasil yang diperoleh dapat dijadikan bahan perbandingan dengan penelitian terdahulu.

Daftar Pustaka

- Badri, F., & Sari, S. U. R. (2021). Penerapan Metode Principal Component Analysis (PCA) Untuk Identifikasi Faktor-Faktor yang Mempengaruhi Sikap Mahasiswa Memilih Melanjutkan Studi ke Kota Malang. *Building of Informatics, Technology and Science (BITS)*, 3(3), 426–431. <https://doi.org/10.47065/bits.v3i3.1139>
- Ghozali. (2017). *Analisis-Faktor-Yang-Merupakan-Intensi*. 7(1).
- Handayani, W. R., & Khasanah, U. (2021). Peran Nilai Islami Anggota Koperasi dalam Meningkatkan Eksistensi Lembaga. *Iqtishoduna*, 17(1), 81–96. <https://doi.org/10.18860/iq.v17i1.9821>
- Ilmaniati & Putro. (2019). Analisis Komponen Utama Faktor-Faktor Pendahulu Mikro, Kecil, Dan Menengah (Umkm) Di Indonesia. 11(1), 67–78.
- Inggi Permana, F. N. S. (2022). The Effect of Data Normalization on the Performance of the Classification Results of the Backpropagation Algorithm. *IJIRSE: Indonesian*

Journal of Informatic Research and Software Engineering, 2(1), 67–72.

- Nitasari, A. N., Salsabila, F. N., Ramadhanty, D. T., Anggriawan, M. R., Amelia, D., Mardianto, M. F. F., & Ana, E. (2023). Reduksi Faktor-Faktor yang Mempengaruhi Kualitas Air Hujan dengan Metode Analisis Komponen Utama. *Zeta - Math Journal*, 8(1), 7–15. <https://doi.org/10.31102/zeta.2023.8.1.7-15>
- Nugroho, S. (2008). *Statistika Mutivariat Terapan*. In UNIB Press Bengkulu.
- Rendana, M., Pitayati, P. A., & Lestari, M. (2022). Penggunaan Metode Analisis Komponen Utama untuk Menilai Hubungan antara Sebaran Kasus COVID-19 dan Kelompok. *Jurnal Ilmu Kesehatan Masyarakat*, 11(01), 84–92. <https://doi.org/10.33221/jikm.v11i01.1034>
- Robert, B., & Brown, E. B. (2004). *Appllied Multivariate Statistical Analysis*. Title (Issue 1).
- Setiawan, M. S. D. (2017). *Skripsi Analisis faktor-faktor yang mempengaruhi produksi susu sapi perah di Desa Sumogawe Kecamatan Getasan*. 64.
- Susilawati, I., Putranto, W. S., & Khairani, L. (2021). Pelatihan Berbagai Metode Pengolahan Susu Sapi sebagai Upaya Mengawetkan, Meningkatkan Nilai Manfaat, dan Nilai Ekonomi. *Media Kontak Tani Ternak*, 3(1), 27–31.