

Implementasi *Principal Component Analysis* untuk Mereduksi Faktor-faktor Kandungan Gizi yang Mempengaruhi Jenis Makanan

Onoy Rohaeni¹, Mimah Mutmainah², Ariela Safmi Ramdhani³, Ery Supriyadi R.⁴

Universitas Koperasi Indonesia^{1,2,3,4}

onoyrohaeni@ikopin.ac.id¹

ABSTRAK

Principal Component Analysis (PCA) adalah teknik statistik yang digunakan untuk mengurangi dimensi data multivariat dengan mengubah variabel asli menjadi komponen utama yang lebih sedikit. Penelitian ini bertujuan untuk mereduksi faktor-faktor kandungan gizi yang mempengaruhi jenis makanan dengan membentuk beberapa komponen utama yang terbentuk untuk menggambarkan pola kandungan gizi makanan. PCA telah dilakukan untuk menyederhanakan data kandungan gizi makanan yang terdiri dari enam variabel. Hasil analisis menunjukkan bahwa dua komponen utama dapat menjelaskan 72,2% variasi total data. Komponen utama pertama merepresentasikan pola makanan berbasis karbohidrat, sedangkan komponen utama kedua merepresentasikan pola makanan tinggi energi dan protein. Meskipun nilai KMO rendah, hasil *Bartlett's Test of Sphericity* menunjukkan korelasi signifikan antar variabel, sehingga PCA dapat digunakan sebagai analisis eksploratif untuk memperoleh gambaran awal struktur dan pola kandungan gizi makanan.

Kata Kunci: *PCA, Kandungan gizi makanan, KMO, Bartlett's Test of Sphericity*

I. PENDAHULUAN

Makanan merupakan kebutuhan dasar manusia yang berperan penting dalam menjaga kesehatan dan menunjang aktivitas sehari-hari. Menurut Undang-Undang No. 36 Tahun 2009 tentang Kesehatan, makanan bergizi adalah makanan yang mengandung zat-zat gizi yang diperlukan oleh tubuh untuk pertumbuhan, pemeliharaan, dan perbaikan jaringan tubuh. Setiap jenis makanan memiliki kandungan gizi yang berbeda-beda, seperti energi, karbohidrat, protein, lemak, vitamin, serat, dan gula. Perbedaan kandungan gizi tersebut menyebabkan data makanan memiliki banyak variabel yang saling berkaitan satu sama lain. Kandungan gizi makanan sangat penting untuk kesehatan manusia. Namun, dengan banyaknya jenis makanan yang tersedia, sulit untuk menentukan faktor-faktor kandungan gizi yang paling berpengaruh terhadap jenis makanan. Oleh karena itu, perlu dilakukan analisis untuk mereduksi faktor-faktor kandungan gizi yang paling berpengaruh.

Dalam pengolahan data gizi, sering kali ditemukan hubungan antar variabel, misalnya antara protein dan lemak atau antara karbohidrat dan gula. Banyaknya variabel yang saling berkorelasi membuat data menjadi kompleks dan sulit dianalisis secara langsung menggunakan metode statistik sederhana. Jika seluruh variabel dianalisis tanpa penyederhanaan, maka interpretasi hasil analisis menjadi kurang efektif dan berpotensi menimbulkan informasi yang berulang.

Untuk mengatasi permasalahan tersebut, diperlukan metode analisis yang mampu menyederhanakan data tanpa menghilangkan informasi penting yang terkandung di dalamnya. Salah satu metode statistik multivariat yang dapat digunakannya adalah Analisis Komponen Utama (*Principal Component Analysis/PCA*). Implementasi PCA dalam penelitian ini bertujuan untuk mereduksi faktor-faktor kandungan gizi yang mempengaruhi jenis makanan dengan membentuk beberapa komponen utama yang terbentuk untuk menggambarkan pola

kandungan gizi makanan. Faktor-faktor kandungan gizi yang paling berpengaruh terbagai menjadi dua komponen utama. Komponen utama satu adalah protein, lemak, dan karbohidrat. Komponen utama dua adalah energi, protein, dan lemak. Melalui implementasi PCA ini diharapkan dapat memperoleh gambaran pola utama kandungan gizi makanan yang lebih sederhana, sehingga memudahkan proses analisis dan interpretasi data.

II. TINJAUAN PUSTAKA

A. Analisis Komponen Utama (*Principal Component Analysis*)

Analisis Komponen Utama atau *Principal Component Analysis (PCA)* adalah salah satu metode dalam analisis multivariat yang bertujuan untuk mereduksi dimensi data. PCA bekerja dengan cara mengubah variabel-variabel asli menjadi sejumlah variabel baru yang disebut komponen utama.

Komponen utama merupakan kombinasi linier dari variabel asli dan disusun berdasarkan besarnya variasi data yang dapat dijelaskan. Komponen pertama menjelaskan variasi terbesar dalam data, diikuti oleh komponen-komponen berikutnya. Dengan menggunakan PCA, data yang memiliki banyak variabel dapat disederhanakan menjadi beberapa komponen tanpa kehilangan sebagian besar informasi penting.

Perkembangan *Principal Component Analysis (PCA)* dimulai sejak diperkenalkan pertama kali oleh Karl Pearson pada tahun 1901 dan kemudian dikembangkan secara formal oleh Harold Hotelling pada 1933. Prinsip dasar dari PCA adalah mereduksi dimensi data dengan cara mengubah kumpulan variabel asli yang saling berkorelasi menjadi sekumpulan variabel baru yang tidak berkorelasi, yang disebut *principal components*. Komponen-komponen ini disusun berdasarkan jumlah variansi data yang menjelaskan komponen pertama mewakili arah variasi terbesar dalam data, komponen kedua mewakili variasi terbesar berikutnya yang ortogonal terhadap yang pertama, dan seterusnya.

Secara matematis, PCA memanfaatkan dekomposisi eigen atau *Singular Value Decomposition (SVD)* dari matriks kovarians data. Vektor eigen (*eigenvector*) menggambarkan arah komponen utama, sedangkan nilai eigen (*eigenvalue*) menunjukkan besarnya kontribusi masing-masing komponen terhadap total variansi data. Dengan memilih hanya beberapa komponen utama pertama, PCA memungkinkan perwakilan data yang lebih sederhana namun tetap informatif.

B. Asumsi dan Tahapan *Principal Component Analysis (PCA)*

Asumsi *Principal Component Analysis (PCA)* adalah suatu kondisi yang harus dipenuhi agar hasil analisis PCA dapat diandalkan. Sebelum melakukan PCA, terdapat beberapa asumsi yang harus dipenuhi, antara lain data harus bersifat numerik, berdistribusi normal, dan terdapat korelasi antar variabel. Untuk menguji kelayakan data, digunakan uji *Kaiser-Meyer-Olkin (KMO)* dan *Bartlett's Test of Sphericity*. Nilai KMO menunjukkan kecukupan sampel, sedangkan *Bartlett's Test* menguji apakah terdapat korelasi yang signifikan antar variabel.

Tahapan analisis PCA dimulai dengan membersihkan data dari kesalahan, *missing value*, dan *outlier*; kemudian melakukan standarisasi data untuk memiliki mean nol dan variansi satu, terutama jika variabel memiliki satuan yang berbeda. Selanjutnya dilakukan proses ekstraksi

komponen utama berdasarkan nilai *eigenvalue*, di mana komponen dengan *eigenvalue* lebih dari satu dipertimbangkan sebagai komponen utama. Selain itu, *scree plot* digunakan untuk membantu menentukan jumlah komponen yang paling optimal. Adapun interpretasi dari *scree plot* adalah mencari titik *elbow* (titik belok) pada grafik *scree plot*, memilih komponen utama yang berada di sebelah kiri titik *elbow*, mengabaikan komponen utama yang berada di sebelah kanan titik *elbow*.

C. Interpretasi Hasil Analisis PCA

Interpretasi hasil *Principal Component Analysis* (PCA) adalah suatu proses untuk memahami dan menjelaskan hasil analisis yang dilakukan dengan melihat nilai *eigenvalue*, persentase variasi yang dijelaskan, serta matriks komponen. *Eigenvalue* menunjukkan kontribusi masing-masing komponen dalam menjelaskan variasi data, sedangkan persentase variasi menggambarkan seberapa besar informasi data yang dapat dijelaskan oleh komponen tersebut. Selain itu, interpretasi juga dilakukan melalui matriks komponen hasil rotasi yang menunjukkan hubungan antara variabel asli dengan komponen utama. Variabel dengan nilai loading yang tinggi pada suatu komponen dianggap sebagai variabel dominan. Berdasarkan variabel dominan tersebut, setiap komponen utama dapat diberi makna atau nama sesuai dengan karakteristik gizi yang diwakilinya.

III. METODE PENELITIAN

A. Sumber Data Penelitian

Sumber data dalam penelitian ini adalah data kandungan gizi berbagai jenis makanan. Jumlah objek yang dianalisis dalam penelitian ini adalah 50 jenis makanan, yang masing-masing memiliki nilai kandungan gizi yang berbeda-beda. Data tersebut digunakan sebagai dasar untuk mengidentifikasi pola umum kandungan gizi makanan.

B. Variabel Penelitian

Variabel yang digunakan dalam penelitian ini terdiri dari enam variabel numerik, yaitu:

1. Energi (kkal), menunjukkan jumlah energi yang terkandung dalam makanan.
2. Protein (gram), berfungsi sebagai zat pembangun dan perbaikan jaringan tubuh.
3. Lemak (gram), merupakan sumber energi dan pelarut vitamin tertentu.
4. Karbohidrat (gram), sebagai sumber energi utama bagi tubuh.
5. Serat (gram), berperan dalam menjaga kesehatan sistem pencernaan.
6. Gula (gram), merupakan bagian dari karbohidrat sederhana.

Seluruh variabel tersebut berskala numerik dan dianalisis secara bersamaan menggunakan PCA.

C. Sumber dan Jenis Data

Pengolahan data dalam penelitian ini dilakukan menggunakan perangkat lunak SPSS. Tahapan awal pengolahan data meliputi pemeriksaan data untuk memastikan tidak terdapat kesalahan input, seperti nilai kosong atau kesalahan penulisan angka. Data kemudian disiapkan dalam format yang sesuai, di mana setiap baris merepresentasikan satu jenis makanan dan

setiap kolom merepresentasikan satu variabel kandungan gizi. Seluruh variabel diinput dalam bentuk numerik agar dapat diolah dalam analisis faktor.

Teknik analisis data yang digunakan dalam penelitian ini adalah Analisis Komponen Utama (*Principal Component Analysis*) dengan bantuan menu *Factor Analysis* pada SPSS.

Tahapan analisis data dilakukan sebagai berikut:

1. Uji Kelayakan Data

Uji kelayakan data dilakukan untuk memastikan bahwa data layak dianalisis menggunakan PCA. Uji yang digunakan adalah *Kaiser-Meyer-Olkin* (KMO) untuk mengukur kecukupan sampel, sedangkan untuk menguji adanya korelasi antar variabel adalah uji *Bartlett's Test of Sphericity*.

Data dinyatakan layak dianalisis apabila *Bartlett's Test* menunjukkan nilai signifikansi kurang dari 0,05, yang berarti terdapat korelasi antar variabel.

2. Ekstraksi Komponen Utama

Setelah data dinyatakan layak, dilakukan proses ekstraksi komponen utama menggunakan metode *Principal Component*. Pada tahap ini, komponen utama dibentuk berdasarkan nilai *eigenvalue*, di mana komponen dengan *eigenvalue* lebih dari satu dipertimbangkan sebagai komponen utama.

3. Penentuan Jumlah Komponen Utama

Jumlah komponen utama ditentukan berdasarkan:

- a. Nilai *eigenvalue* (>1)
- b. Persentase variasi kumulatif
- c. *Scree plot*, yang menunjukkan titik penurunan tajam (*elbow*)

Ketiga kriteria tersebut digunakan secara bersama-sama untuk memperoleh jumlah komponen yang paling optimal.

4. Rotasi Komponen

Untuk mempermudah interpretasi hasil, dilakukan proses rotasi komponen menggunakan metode *Varimax*. Rotasi ini bertujuan untuk memperjelas hubungan antara variabel asli dengan komponen utama sehingga struktur data menjadi lebih mudah dipahami.

5. Interpretasi Komponen Utama

Tahap akhir analisis adalah interpretasi komponen utama berdasarkan nilai *factor loading*. Variabel dengan nilai loading yang tinggi dianggap memiliki kontribusi besar terhadap pembentukan komponen. Setiap komponen kemudian diberi makna atau nama sesuai dengan karakteristik variabel dominan yang membentuknya.

D. Alur Penelitian

Secara umum, alur dalam penelitian ini dapat dirangkum sebagai berikut:

1. Pengumpulan dan persiapan data kandungan gizi makanan
2. Pengujian kelayakan data (KMO dan *Bartlett's Test*)
3. Ekstraksi dan penentuan jumlah komponen utama
4. Rotasi dan interpretasi komponen
5. Penarikan kesimpulan berdasarkan hasil analisis PCA

IV. HASIL PENELITIAN DAN PEMBAHASAN

A. Gambaran Umum Data Penelitian

Data yang dianalisis dalam penelitian ini merupakan data kandungan gizi dari 50 jenis makanan dengan enam variabel, yaitu energi, protein, lemak, karbohidrat, serat, dan gula. Keenam variabel tersebut dipilih karena mewakili komponen gizi utama yang umum digunakan dalam penilaian makanan.

Sebelum dilakukan analisis lebih lanjut, data terlebih dahulu dimasukkan ke dalam perangkat lunak SPSS dan dipastikan seluruh variabel telah didefinisikan dalam bentuk numerik. Setiap baris data merepresentasikan satu jenis makanan, sedangkan setiap kolom menunjukkan kandungan gizi tertentu. Data ini kemudian dianalisis menggunakan metode Analisis Komponen Utama (*Principal Component Analysis/PCA*) untuk melihat pola umum kandungan gizi makanan.

B. Uji Kelayakan Data (*KMO dan Bartlett's Test*)

Langkah awal dalam analisis PCA adalah menguji kelayakan data. Uji kelayakan ini bertujuan untuk memastikan bahwa data yang digunakan memang layak dianalisis menggunakan PCA. Dua uji yang digunakan adalah *Kaiser-Meyer-Olkin* (KMO) dan *Bartlett's Test of Sphericity* yang hasil ujinya adalah sebagai berikut:

Tabel 1. *Kaiser-Meyer-Olkin* (KMO) dan *Bartlett's Test of Sphericity*

KMO and Bartlett's Test		
Kaiser-Meyer-Olkin Measure of Sampling Adequacy.		.298
Bartlett's Test of Sphericity	Approx. Chi-Square	277.242
	df	15
	Sig.	.000

Berdasarkan hasil output SPSS, diperoleh nilai KMO sebesar 0,298. Nilai ini menunjukkan bahwa kecukupan sampel tergolong rendah. Artinya, secara statistik struktur data belum terlalu kuat untuk dianalisis menggunakan PCA secara optimal. Hal ini dapat disebabkan oleh jumlah sampel yang relatif terbatas atau korelasi antar variabel yang belum terlalu tinggi.

Namun demikian, hasil *Bartlett's Test of Sphericity* menunjukkan nilai signifikansi sebesar 0,000, yang berarti lebih kecil dari 0,05. Hasil ini menunjukkan bahwa terdapat korelasi yang signifikan antar variabel kandungan gizi. Dengan kata lain, variabel-variabel tersebut saling berhubungan dan tidak berdiri sendiri.

Berdasarkan hasil tersebut, meskipun nilai KMO tergolong rendah, analisis PCA tetap dapat dilakukan karena *Bartlett's Test* menunjukkan bahwa data memiliki korelasi yang cukup untuk dianalisis. Oleh karena itu, PCA dalam penelitian ini digunakan sebagai analisis eksploratif untuk melihat gambaran awal pola kandungan gizi makanan.

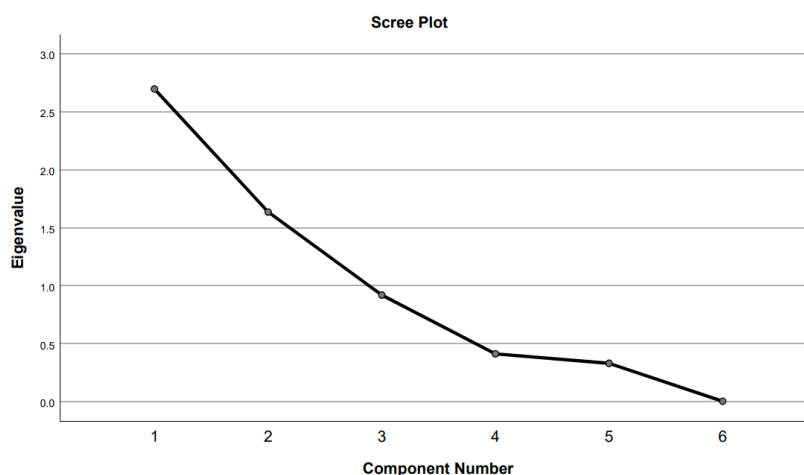
C. Penentuan Jumlah Komponen Utama

Setelah data dinyatakan layak untuk dianalisis, langkah selanjutnya adalah menentukan jumlah komponen utama yang terbentuk. Penentuan jumlah komponen utama dilakukan dengan melihat nilai eigenvalue dan scree plot pada output SPSS.

Tabel 2. Total Variance Explained

Component	Initial Eigenvalues			Extraction Sums of Squared Loadings		
	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %
1	2.696	44.936	44.936	2.696	44.936	44.936
2	1.636	27.263	72.199	1.636	27.263	72.199
3	.920	15.327	87.526			
4	.413	6.884	94.410			
5	.331	5.516	99.926			
6	.004	.074	100.000			

Berdasarkan tabel 2, terdapat dua komponen utama yang memiliki nilai *eigenvalue* lebih besar dari satu. Kedua komponen ini secara bersama-sama mampu menjelaskan sekitar 72,2% variasi total data. Persentase ini menunjukkan bahwa sebagian besar informasi dalam data asli sudah dapat diwakili hanya dengan dua komponen utama.

**Gambar 1. Scree Plot**

Selain itu, *scree plot* juga menunjukkan adanya penurunan nilai *eigenvalue* yang cukup tajam hingga komponen ke-2, kemudian grafik mulai mendatar pada komponen selanjutnya. Pola ini mengindikasikan bahwa dua komponen utama sudah cukup representatif dalam menjelaskan struktur data, sedangkan komponen setelahnya hanya memberikan tambahan informasi yang relatif kecil.

Berdasarkan kedua kriteria tersebut yaitu nilai *eigenvalue* dan *scree plot*, maka ditetapkan bahwa jumlah komponen utama yang digunakan dalam penelitian ini adalah dua komponen.

D. Interpretasi Komponen Utama

Interpretasi pada setiap komponen dilakukan dengan melihat nilai *factor loading* pada *Rotated Component Matrix*. Nilai *factor loading* menunjukkan seberapa kuat hubungan antara variabel asli dengan komponen utama. Ringkasan *Rotated Component Matrix*, disajikan pada tabel 3.

Tabel 3. Rotated Componen Matrix

	Component	
	1	2
Energi	.536	.824
Protein	-.295	.814
Lemak	.231	.879
Karbo	.913	.044
Serat	.557	.086
Gula	.795	.046

Extraction Method: Principal Component Analysis.
Rotation Method: Varimax with Kaiser Normalization.

a. Rotation converged in 3 iterations.

Berdasarkan tabel 3, penamaan komponen dilakukan sebagai berikut:

1. Komponen Utama 1 (pertama). Komponen utama pertama memiliki kontribusi terbesar dalam menjelaskan variasi data. Variabel yang memiliki nilai *loading* tinggi pada komponen ini adalah karbohidrat, gula, dan serat. Hal ini menunjukkan bahwa komponen pertama sangat dipengaruhi oleh kandungan karbohidrat dan turunannya. Berdasarkan variabel dominan tersebut, komponen utama pertama dapat diinterpretasikan sebagai komponen makanan berbasis karbohidrat. Makanan dengan skor tinggi pada komponen ini cenderung memiliki kandungan karbohidrat dan gula yang tinggi.
2. Komponen Utama 2 (kedua). Komponen utama kedua didominasi oleh variabel energi, protein, dan lemak. Variabel-variabel ini berperan penting dalam penyediaan energi dan zat pembangun bagi tubuh. Oleh karena itu, komponen utama kedua dapat diinterpretasikan sebagai komponen makanan tinggi energi dan protein. Makanan dengan skor tinggi pada komponen ini cenderung memiliki kandungan energi, protein, dan lemak yang relatif tinggi dibandingkan makanan lainnya.

E. Pembahasan Hasil Analisis

Hasil analisis PCA menunjukkan bahwa enam variabel kandungan gizi dapat disederhanakan menjadi dua komponen utama yang mewakili pola gizi makanan. Penyederhanaan ini memudahkan proses analisis dan interpretasi karena informasi yang sebelumnya tersebar pada banyak variabel kini dapat dilihat melalui dua dimensi utama.

Komponen pertama menggambarkan pola makanan yang didominasi oleh karbohidrat dan gula, sedangkan komponen kedua menggambarkan pola makanan yang kaya energi dan protein. Pembagian ini menunjukkan bahwa secara umum makanan dapat dikelompokkan berdasarkan sumber energi utamanya.

Meskipun nilai KMO yang diperoleh tergolong rendah, hasil PCA tetap memberikan gambaran awal yang cukup informatif mengenai struktur data kandungan gizi makanan.

V. KESIMPULAN

Berdasarkan hasil analisis dan pembahasan yang telah dilakukan, dapat disimpulkan bahwa Analisis Komponen Utama (*Principal Component Analysis, PCA*) mampu menyederhanakan data kandungan gizi makanan yang terdiri dari enam variabel, yaitu energi, protein, lemak, karbohidrat, serat, dan gula. Melalui penerapan PCA, data yang semula

kompleks dapat direduksi menjadi beberapa komponen utama yang lebih mudah dipahami dan dianalisis.

Hasil analisis menunjukkan bahwa terbentuk dua komponen utama yang secara bersama-sama mampu menjelaskan sekitar 72,2% variasi total data. Komponen utama pertama merepresentasikan pola makanan berbasis karbohidrat, yang didominasi oleh variabel karbohidrat, gula, dan serat. Sementara itu, komponen utama kedua merepresentasikan pola makanan tinggi energi dan protein, yang didominasi oleh variabel energi, protein, dan lemak.

Meskipun nilai *Kaiser-Meyer-Olkin* (KMO) yang diperoleh tergolong rendah, hasil *Bartlett's Test of Sphericity* menunjukkan adanya korelasi yang signifikan antar variabel. Hal ini menunjukkan bahwa PCA dalam penelitian ini tetap dapat digunakan sebagai analisis eksploratif untuk memperoleh gambaran awal mengenai struktur dan pola kandungan gizi makanan.

DAFTAR PUSTAKA

- Hidayat, R., Santoso, B., & Wijayanti, L. (2022). Penerapan principal component analysis dalam pemetaan karakteristik mutu pangan olahan. *Jurnal Teknologi dan Industri Pangan*, 33(1), 45–54. <https://doi.org/10.6066/jtip.2022.33.1.45>.
- Johnson, R. A., & Wichern, D. W. (2018). *Applied Multivariate Statistical Analysis*. 7th Edition. Pearson.
- Lestari, D., Rahmawati, A., & Nugroho, S. (2023). Pemahaman label gizi dan pengaruhnya terhadap pemilihan makanan kemasan pada masyarakat perkotaan. *Jurnal Gizi Indonesia*, 11(2), 97–106. <https://doi.org/10.14710/jgi.11.2.97-106>.
- Nugraha, F., & Laksmiwati, H. (2023). Uji kelayakan analisis faktor menggunakan KMO dan Bartlett pada data kesehatan masyarakat. *Jurnal Biostatistika dan Kependudukan*, 12(1), 33–42.
- Pramesti, D. A., & Wibowo, A. (2024). Reduksi kompleksitas informasi nilai gizi menggunakan principal component analysis pada produk pangan kemasan. *Jurnal Informatika dan Analisis Data*, 5(1), 1–11. <https://doi.org/10.30865/jiad.v5i1.2024>.
- Pratama, R. Y., & Nugroho, B. A. (2022). Pemanfaatan data label gizi sebagai basis analisis kuantitatif produk makanan kemasan. *Jurnal Pangan dan Agroindustri*, 10(3), 161–170.
- Rencher, A. C., & Christensen, W. F. (2012). *Methods of Multivariate Analysis*. 3rd Edition, Wiley.
- Tobing, R.S., dkk. (2026). Penerapan Principal Component Analysis (PCA) untuk Reduksi Dimensi dan Pemetaan Karakteristik Nutrisi pada Produk Makanan Kemasan di Indonesia. *Jurnal Matematika, Ilmu pengetahuan Alam*. 4(1), 20 – 30 <https://doi.org/10.62383/algoritma.v4i1.891>