PREDIKSI KEPADATAN NUTRISI BUAH-BUAHAN DENGAN MODEL REGRESI LINEAR BERGANDA DAN ORDINARY LEAST SQUARE (OLS)

Catherine Vanya Pangemanan¹

1. Universitas Koperasi Indonesia Kawasan Pendidikan Tinggi Jl. Raya Jatinangor No.KM. 20, RW.5, Cibeusi, Kec. Jatinangor, Kabupaten Sumedang, Jawa Barat 45363 Email: catherinevannya742@gmail.com

ABSTRAK

Kepadatan nutrisi makanan, khususnya buahbuahan, menjadi kritis dalam mencegah malnutrisi dan penyakit terkait gizi. Penelitian ini menggunakan Regresi Linier Berganda untuk memprediksi kepadatan nutrisi pada 572 jenis buah berdasarkan kandungan nutrisi seperti protein, serat, vitamin A, vitamin C, vitamin E, kalsium, magnesium, kalium, dan zat besi. Analisis menunjukkan bahwa setiap kenaikan satu unit dalam kandungan protein, serat, dan vitamin C menghasilkan peningkatan kepadatan nutrisi yang diprediksi sebesar 1.002, 1.003, dan 1.004 unit, masing-masing. Model Regresi Linier Berganda yang dikembangkan memiliki Adjusted R-squared sebesar 1.00, menunjukkan bahwa model secara efektif menjelaskan variabilitas data. Uji-F ANOVA menunjukkan bahwa semua variabel prediktor berkontribusi secara signifikan terhadap kepadatan nutrisi (F(34, 537) = 2275.28, p <0.001).

Penelitian ini menjadi penguat dari penelitian terdahulu [1,2,4,5]. Selain itu untuk penelitian berikutnya, penelitian ini dapat menjadi dasar karena penelitian ini menekankan manfaat kepadatan nutrisi dalam meningkatkan kesehatan dan mengurangi risiko penyakit kronis. Studi ini menyoroti pentingnya buah-buahan sebagai sumber utama kepadatan nutrisi yang dapat meningkatkan kesehatan populasi.

Kata kunci: Kepadatan nutrisi, buah-buahan, regresi linier berganda, Ordinary Least Square (OLS)

1. PENDAHULUAN

Institute of Medicine dan WHO [3,7] menyebutkan bahwa buah-buahan menjadi salah satu jenis pangan dengan kepadatan nutrisi yang tinggi. Berdasarkan indeks NRFI kepadatan nutrisi yang tinggi ini berkorespondensi dengan banyaknya vitamin, mineral, dan serat [2] dalam suatu jenis pangan, dalam hal ini adalah buah. Namun 'banyak', pada konteks ini bisa bermakna memiliki banyak ragam kandungan vitamin, mineral dan serat dalam satu buah atau jenis vitamin, mineral dan seratnya tidak terlalu beragam namun memiliki kadar kandungan yang banyak.

Mengetahui kepadatan nutrisi pangan sangat penting untuk mencegah malnutrisi dan obesitas. Maka penelitian ini terlaksana untuk memodelkan prediksi kepadatan nutrisi pada berbagai jenis buah, berdasarkan kandungan vitamin, mineral, dan serat seperti protein, serat, vitamin A, vitamin C, vitamin E, kalsium, magnesium, kalium, dan zat besi menggunakan Regresi Linier Berganda dengan metode Ordinary Least Square. Selain itu penelitian ini bertujuan untuk memastikan bagaimana makna 'banyak' pada konteks kandungan vitamin, mineral dan serat pada jenis pangan berkepadatan nutrisi tinggi, salah satunya adalah buah-buahan.

2. TINJAUAN PUSTAKA

Kepadatan nutrisi mengukur jumlah nutrisi seperti vitamin, mineral, dan serat dalam suatu jumlah kalori makanan. Nutrient Rich Foods Index menilai kepadatan nutrisi makanan berdasarkan kandungan nutrisi utama yaitu: protein, serat, vitamin A, vitamin C, vitamin E, kalsium, magnesium, kalium, dan zat besi [2].

Makanan berkepadatan nutrisi tinggi, seperti buahbuahan, menawarkan lebih banyak nutrisi per kalori dibandingkan makanan berkepadatan rendah [1].

Institute of Medicine [3] dan WHO [7] merekomendasikan konsumsi makanan padat nutrisi, termasuk buah-buahan, sayuran, bijibijian, dan protein rendah lemak. USDA dan HHS juga menekankan pentingnya pemenuhan kebutuhan nutrisi dari makanan berkualitas untuk mengatasi gizi buruk dan [2]. Sebagai buktinya, diet tinggi kepadatan nutrisi meningkatkan kesehatan dan mengurangi risiko penyakit kronis [4] dan program intervensi yang mempromosikan makanan padat nutrisi juga efektif meningkatkan perbaikan asupan nutrisi pada populasi rentan [5].

Penelitian terhadap dataset [6] menggunakan Regresi Linier Berganda dengan metode OLS berhipotesis bahwa kita dapat memprediksi kepadatan nutrisi berbagai jenis buah-buahan berdasarkan kandungan vitamin, mineral, dan serat seperti protein, serat, vitamin A, vitamin C, vitamin E, kalsium, magnesium, kalium, dan zat besi. Ini bertujuan untuk membuktikan apakah benar buah memiliki kepadatan nutrisi yang kaya akan kandungan tersebut.

3. METODE PENELITIAN

Bahasa pemrograman Python akan membantu penelitian ini membersihkan dataset kemudian membangun model Regresi Linier Berganda untuk memprediksi Y = 'Kepadatan Nutrisi' pada buah-buahan dengan metode *Ordinary Least Square*. Setelah model dibangun maka model akan memulai tahap uji keberartian menggunakan Uji ANOVA, uji koefisien determinasi (R^2) , hingga *future selection* melalui Uji-T.

No.	Nama Fitur	Variabel	No.	Nama Fitur	Variabel
1	Makanan	X1	18	Vitamin B2 (Riboflavin) (mg)	X18
2	Nilai kalori	X2	19	Vitamin B3 (Niacin) (mg)	X19
3	Lemak (g)	X3	20	Vitamin B5 (Asam Pantotenat) (mg)	X20
4	Lemak Jenuh (g)	X4	21	Vitamin B6 (mg)	X21
5	Lemak tak jenuh tunggal (g)	X5	22	Vitamin C (mg)	X22
6	Lemak tak jenuh ganda (g)	X6	23	Vitamin D (mg)	X23
7	Karbohidrat (g)	X7	24	Vitamin E (mg)	X24
8	Gula (g)	X8	25	Vitamin K (mg)	X25
9	Protein (g)	X9	26	26 Kalsium (mg)	
10	Serat Makanan (g)	X10	27	Tembaga (mg)	X27
11	Kolesterol (mg)	X11	28	Besi (mg)	X28
12	Sodium (g)	X12	29	Magnesium (mg)	X29
13	Air (g)	X13	30	Mangan (mg)	X30
14	Vitamin A (mg)	X14	31	Fosfor (mg)	X31
15	Vitamin B1 (Tiamin) (mg)	Vitamin B1 (Tiamin) (mg) X15 32 Kalium (mg)		X32	
16	Vitamin B11 (Asam Folat) (mg)	X16	33	Selenium (mg)	X33
17	Vitamin B12 (mg)	3717	34	Seng (mg)	X34
1/	Vitamin B12 (mg)	X17	35	Kenadatan Nutrisi	v

Tabel 1 Variabel X dan Y

3.1. Deskripsi Dataset

Penelitian ini menggunakan dataset 'Himpunan Data Nutrisi Pangan' hasil akuisisi web-scrapping situs Yazio.com [6] memanfaatkan library Python berupa BeautifulSoup, pandas, Matplotlib, dan seaborn. Database Makanan Komprehensif [6] memberikan informasi nutrisi terperinci untuk berbagai macam makanan yang biasa dikonsumsi di seluruh dunia. Dataset ini bertujuan untuk mendukung perencanaan diet, analisis nutrisi, dan tujuan pendidikan dengan menyediakan data ekstensif tentang kandungan makro dan mikronutrien makanan. Dataset memiliki 37 fitur dan 572 baris data. Worksheet yang dipilih penelitian ini adalah 'Kandungan Gizi Pada Buah-Buahan'.

No.	Nama Fitur	ur Penjelasan				
1	Makanan	Nama atau jenis makanan.				
2	Nilai kalori	Total energi yang disediakan oleh makanan, biasanya diakar dalam kilokalori (kital) per 100 gram.				
3	Lemak (g)	Jumlah total lemak gram per 100 gram, termasuk kerusakan yang mengikutinya.				
4	Lemak Jenuh (g)	Jumlah lemak jenuh (lemak yang biasanya meningkatkan kadar kolesterol dalam darah) gram per 100 gram.				
5	Lemak tak jenuh tunggal (g)	Furnish lemak tak jemih tunggal (dianggap lemak jantung sehat) gram per 100 gram.				
6	Lemak tak jenuh ganda (g)	Jumlah lemak tak jeouh ganda (termasuk lemak esensial yang dibutuhkan tubuh Anda tetapi tidak dapat diproduksi sendiri) gram per 100				
7	Karbohidrat (g)	Total forbobidest gram per 100 gram, termasuk gula.				
8	Gula (g)	Total gula gram per 100 gram, subset dari karbohidrat.				
9	Protein (g)	Total protein gram per 100 gram, penting untuk perbaikan dan pertumbuhan tubuh.				
10	Serat Makanan (g)	Kandungan serat gram per 100 gram, penting untok kesehatan pencernaan.				
11	Kolesterol (mg)	Kandungan kolesterol dalam miligram per 100 gram, berkaitan dengan kesehatan jantung.				
12	Sodium (g)	Kandungan natrium dalam miligram per 100 gram, penting untuk keseimbangan cairan dan fungsi saraf.				
13	Air (g)	Kadar air gram per 100 gram, yang mempengaruhi kepadatan energi makanan.				
14	Vitamin A (mg)	Jumlah Vitamin A dalam mikrogram per 100 gram, penting untuk penglihatan dan fungsi kekebalan tubuh.				
15	Vitantin B1 (Tiamin) (mg)	Penting untuk metabatisme glukosa.				
16	Vitamin B11 (Asam Folat)	Penting untuk fungsi sel dan pertumbuhan jaringan, terutama penting dalam kehamilan.				
17	Vitamin B12 (mg)	Penting untuk fungsi etak dan pembentukan darah.				

Tabel 2 Keterangan Fitur Pada Dataset Nutrisi Pangan

No.	Nama Fitur	er Penjelstan					
18	Vitamin B2 (Riboflavin) (rog)	Diperlukan untuk produksi energi, fungsi sel, dan metabolisme lemak.					
19	Vitamin B3 (Nizcin) (mg)	Mendukung kesebatan sistem pencernaan, kulit, dan saraf.					
20	Vitamin B5 (Asam Pantotenat) (mg)	Dipertukan untuk membuat sel darah, dan membantu mengubah makanan menjadi energi.					
21	Vitamin B6 (mg)	Penting untuk perkembangan otak normal dan menjaga sistem saraf dan kelebalan tubuh tetap sehat.					
22	Vitamin C (mg)	Penting untuk perbaikan sensua jaringan tubuh.					
23	Vitamin D (mg)	Penting untuk penyerapan kalsium, meningkatkan pertumbuhan tulang dan kesehatan.					
24	Vitamin E (mg)	Bertindak sebagai antioksidan, membantu melindungi sel dari kerusakan yang disebabkan oleh radikal bebas.					
25	Vitamin K (mg)	Diperlukan untuk pembekuan darah dan kesehatan tulang					
26	Kalsium (mg)	Penting untuk membangun dan memelihara tulang dan gigi yang kuat.					
27	Tembaga (mg)	Membantu pembentukan kolagen, meningkatkan penyerapan zat besi dan berperan dalam produksi energi.					
28	Besi (mg)	Penting untak pembentukan sel darah merah.					
29	Magnesium (mg)	Penring uatuk banyak proses dalam tubuh termasuk pengaturan fungsi otot dan saraf, kadar gula darah, dan tekanan darah dan membuat protein, tulang, dan DNA.					
30	Mangan (rog)	Terlibat dalam pembentukan tulang, faktor pembekuan darah, dan enzim yang berperan dalam metabelisme lemak dan karbohidrat, penyerapan kalsism, dan regulasi gula darah.					
31	Fosfor (mg)	Membantu pembentukan tulang dan gigi dan diperlukan bagi tubuh untuk membuat protein untuk pertumbuhan, pemeliharaan, dan perbaikan sel dan jaringan.					
32	Kalium (mg)	Membantu mengatur keseimbangan cairan, kontraksi otot, dan sinyal saraf.					
33	Scientum (mg)	Penting untuk reproduksi, fungsi kelenjar tiroid, produksi DNA, dan melindungi tubuh dari kerusakan yang disebabkan oleh radikal bebar dan infeksi.					
34	Sing (mg)	Diperlukan agar sistem kekebalan tubuh berfungsi dengan balk dan berperan dalam pembelahan sel, pertumbuhan sel, penyembuhan luka dan pemecahan karbohidrat.					

Tabel 3 Keterangan Fitur Pada Dataset Nutrisi Pangan

3.2. Pemodelan Regresi Linier Berganda dengan OLS

Merupakan teknik statistik yang digunakan untuk memahami hubungan antara satu variabel dependen (Y) dan dua atau lebih variabel independen (X). Metode ini berusaha untuk menemukan garis regresi terbaik yang dapat meminimalkan jumlah kuadrat perbedaan antara nilai-nilai observasi sebenarnya dan nilai-nilai yang diprediksi oleh model. OLS memperkirakan koefisien regresi (beta) untuk setiap variabel X. Tujuan utamanya adalah untuk menemukan koefisien yang menyesuaikan garis regresi sedekat mungkin dengan data sebenarnya (fit).

Model Regresi Linier Berganda:

$$\hat{Y} = \beta_0 + \beta_1 X_1 + \dots + \beta_n X_n \tag{1}$$

Rumus Ordinary Least Square:

$$RSS = \Sigma (Y_i - \hat{Y}_i)^2 \tag{2}$$

$$RSS = \Sigma (Y_i - \hat{b}_0 - \hat{b}_1 X_i)^2 \tag{3}$$

3.3. Uji-F ANOVA

Uji keberartian model ini bertujuan menentukan apakah variabel predictor (*independen*) X_1 sampai dengan X_{34} tersebut signifikan, secara perhitungan statistik, dapat memengaruhi performa variabel yang diprediksi (*dependen*) Y.

Pasangan Hipotesis Uji-F:

 H_0 : $\beta=0$: $variabel\ X$ tidak memiliki pengaruh terhadap variabel\ Y H_1 : $\beta\neq0$: $variabel\ X$ memiliki pengaruh terhadap variabel\ Y

Statistik Uji-F:

$$F_{hitung} = \frac{r^2(n-k-1)}{k(1-r^2)}$$
 (4)

3.4. Uji Koefisien Determinasi Model Regresi (R^2)

Evaluasi dilakukan untuk memastikan model memberikan prediksi yang akurat dan dapat diandalkan. Salah satu metrik evaluasi umum adalah R-squared (koefisien determinasi), yang mengukur seberapa baik variabilitas dalam data yang dijelaskan oleh model.

3.5. Uji-T

Menguji signifikansi koefisien model regresi yang menunjukkan bahwa variabel prediktor terkait memiliki kontribusi yang berarti terhadap model.

Pasangan Hipotesis Uji-T:

 H_0 : $\beta=0$: tidak terdapat pengaruh yang signifikan antara variabel X dengan variabel Y. H_1 : $\beta\neq0$: terdapat pengaruh yang siginifikan antara variabel X dengan variabel Y.

4. HASIL PENELITIAN DAN PEMBAHASAN

Prediksi kepadatan nutrisi dari dataset [6] menggunakan Regresi Linier Berganda dan Ordinary Least Square melalui tahap dan menghasilkan informasi berikut.

4.1. Data Cleansing

Dataset dibersihkan dari kolom 'Unnamed: 0.1', 'Unnamed: 0', yang bertipe data integer. Karena dua kolom tersebut tidak berarti apa-apa bagi dataset sehingga dihilangkan saja menggunakan fungsi 'column to drop'.

4.2. Memodelkan Regresi Linier Berganda dengan OLS

Gunakan library ScikitLearn untuk melakukan perhitungan Ordinary Least Square (kuadrat terkecil). Didapatkan hasil seperti pada Tabel 4. Untuk nilai koefisien konstantanya adalah 0.0009.

No.	Variabel X	Koefisien	No.	Variabel X	Koefisien
1	Nilai Kalori	-6.2E-05	11	Natrium	0.0018
2	Lemak	1.0002	12	Air	-3.352E-06
3	Lemak Jenuh	0.0003	13	Vitamin A	1.0005
4	Lemak Tunggal Tak Jenuh	0.0003	14	Vitamin B1	-0.0009
5	Lemak Jamak Tak Jenuh	0.0005	15	Vitamin B11	-0.0137
6	Karbohidrat	1.0003	16	Vitamin B12	-0.0581
7	Gula	0.000014	17	Vitamin B2	0.0005
8	Protein	1.0004	18	Vitamin B3	-0.0003
9	Serat Pangan	1	19	Vitamin B5	0.0003
10	Kolesterol	2.24E-07	20	Vitamin B6	0.0009

Tabel 4 Model Regresi Linier Berganda Hasil OLS

No.	Variabel X	Koefisien	No.	Variabel X	Koefisien
21	Vitamin C	1	31	Kalium	-3.438E-06
22	Vitamin D	-0.0002	32	Selenium	0.000008099
23	Vitamin E	0.0005	33	Seng	-0.0001
24	Vitamin K	-0.0001			
25	Kalsium	1			
26	Tembaga	1.42E-06			
27	Zat Besi	1			
28	Magnesium	2.36E-05			
29	Mangan	-1.5E-05			
30	Fosfor	5.17E-06			

Tabel 5 Model Regresi Linier Berganda Hasil OLS

4.3. Menguji Keberartian Model Regresi Linear Berganda OLS dengan Uji-F (ANOVA)

F-statistik (F hitung	g) Peluang F-statistik
4.08E+08	0.00

Kriteria uji untuk F-statistik atau F_{hitung} berdasarkan hasil Peluang F-statistik pada Tabel 6 adalah: Peluang $F_{hitung} < \alpha$, dengan $\alpha = 5\%$,

ENTHISIAS JOURNAL

dengan Uji-T

maka H_0 ditolak dan H_1 diterima. Artinya variabel X memiliki pengaruh terhadap variabel Y.

4.4. Menguji koefisien determinasi (R2) Model Regresi Linear Berganda OLS

Pada Tabel 7 nilai masing-masing R-square dan Adjusted R-Square adalah sebesar 1.00, yang berarti hal tersebut berkorespondensi dengan kinerja model yang sangat baik untuk memprediksi Kepadatan Nutrisi buah-buahan.

djusted R-Square
1.00

4.5. Menguji Signifikansi Koefisien Model Regresi Linier Berganda OLS

Berikut adalah beberapa variabel X terpilih hasil Uji-T, yang memiliki pengaruh signifikan terhadap variabel Y (Kepadatan Nutrisi). Dilihat dari kriteria Uji-T, bahwa $p-value\ t_{hitung} < \alpha$, dengan $\alpha = 5\%$, menyebabkan H_0 ditolak dan H₁ diterima. Sebelum variabel-variabel X ini juga, terdapat konstanta dengan $t_{hitung} = 0.371$ $dan |P| > t_{hitung} = 0.711.$

No.	Variabel :	X	t	P > t	Kriteria Uji-T
1	Lemak	X3	1476.37	0.00	p-value < 0.05, Terima H1
2	Karbohidrat	X7	2277.27	0.00	p-value < 0.05, Terima H1
3	Protein	X9	1985.79	0.00	p-value < 0.05, Terima H1
4	Serat Pangan	X10	2613.25	0.00	p-value < 0.05, Terima H1
5	Vitamin A	X14	3575.61	0.00	p-value < 0.05, Terima H1
6	Vitamin C	X22	2.12E+04	0.00	p-value < 0.05, Terima H1
7	Kalsium	X26	2.65E+04	0.00	p-value < 0.05, Terima H1
8	Besi	X28	5137.85	0.00	p-value < 0.05, Terima H1

Tabel 8 Uii-T

Seluruh variabel X pada Tabel 8 ini adalah variabel X yang akan digunakan sebagai prediktorprediktor pada model prediksi variabel Y (Kepadatan Nutrisi) ini. Maka model prediksi yang dihasilkan sesuai persamaan (1) adalah:

$$\hat{Y} = 0.0009 + 1.0002X_3 + 1.0003X_7 + 1.0004X_9 + 1X_{10} + 1.0005X_{14} + 1X_{22} + 1X_{26} + 1X_{28}$$
 (5)

Dalam persamaan prediksi kepadatan nutrisi dari model regresi linear berganda, setiap variabel prediktor memiliki peran penting. Variabel Lemak (X_3) , Karbohidrat (X_7) , dan Protein (X_9) masingmasing dengan koefisien 1.0002, 1.0003, dan

1.0004, menunjukkan bahwa peningkatan satu unit dalam kandungan lemak, karbohidrat, dan protein makanan berkontribusi langsung pada peningkatan kepadatan nutrisi yang diprediksi sebesar masingmasing 1.0002, 1.0003, dan 1.0004 unit. Variabel Serat Pangan (X_{10}) dengan koefisien 1 memberikan kontribusi positif terhadap kepadatan nutrisi yang diprediksi. Variabel Vitamin A (X_{14}) , Vitamin C (X_{22}) , Kalsium (X_{26}) , dan Zat Besi (X_{28}) masing-masing dengan koefisien 1 menunjukkan bahwa peningkatan kandungan vitamin dan mineral ini juga berkontribusi peningkatan langsung pada kepadatan nutrisi yang diprediksi.

KESIMPULAN

Penelitian menggunakan model prediksi Regresi Linier Berganda metode Ordinary Least Square ini menghasilkan persamaan $\hat{Y} = 0.0009 +$ $1.0002X_3 + 1.0003X_7 + 1.0004X_9 + 1X_{10} +$ $1.0005X_{14} + 1X_{22} + 1X_{26} + 1X_{28}$. Model regresi linier berganda tersebut membuktikan bahwa memprediksi Kepadatan Nutrisi beragam jenis buah dapat dilakukan dengan variabel Lemak, Karbohidrat, Protein, Serat Pangan, Vitamin A, Vitamin C, Kalsium dan Zat Besi. Hasil ini mengonfirmasi penelitian yang menyebutkan bahwa makanan dengan kepadatan gizi tinggi mengandung protein, serat, vitamin A, vitamin C, kalsium, dan zat besi [1,2,4,5] kecuali vitamin E, magnesium dan kalium.

Model regresi linier berganda OLS ini berhasil menunjukkan bahwa model prediksi yang dibuat secara sangat baik mampu menjelaskan variabilitas kepadatan nutrisi dari buah-buahan berdasarkan variabel-variabel vang dimasukkan ke dalam model. Nilai R-squared dan Adjusted R-squared yang mendekati 1.000 menunjukkan bahwa model mampu menjelaskan semua variasi yang diamati dalam data. F-statistic yang sangat besar dan Prob (F-statistic) yang mendekati nol menunjukkan bahwa model secara keseluruhan signifikan secara statistik. Log-Likelihood yang tinggi (1129.2) menunjukkan bahwa model kemungkinan yang tinggi untuk menghasilkan data yang diamati.

DAFTAR PUSTAKA (12pt, bold)

[1]. Drewnowski, A. (2005). Concept of a nutritious food: Toward a nutrient density

- score. American Journal of Clinical Nutrition, 82(4), 721-732.
- [2]. Fulgoni, V. L., Keast, D. R., Drewnowski, A. (2009). Development and validation of the nutrient-rich foods index: A tool to measure nutritional quality of foods. *Journal of Nutrition*, 139(8), 1549-1554.
- [3]. Institute of Medicine. (2005). Dietary Reference Intakes for Energy, Carbohydrate, Fiber, Fat, Fatty Acids, Cholesterol, Protein, and Amino Acids. Washington, DC: The National Academies Press.
- [4]. Miller, W. C., Koceja, D. M., & Hamilton, E. J. (2009). A meta-analysis of the past 25 years

- of weight loss research using diet, exercise, or diet plus exercise intervention. *International Journal of Obesity*, 31(8), 160-164.
- [5]. Nicklas, T. A., O'Neil, C. E., & Fulgoni, V. L. (2011). Diet quality is inversely related to cardiovascular risk factors in adults. *Journal of Nutrition*, *141*(12), 2217-2223.
- [6]. Utsav Dey. (2024). Food Nutrition Dataset [Data set]. Kaggle. https://doi.org/10.34740/KAGGLE/DSV/88 20139
- [7]. World Health Organization. (2020). Healthy diet.