

# Evaluasi *Multimodal Gemini Model* dalam Menganalisis Informasi Nilai Gizi Produk untuk Pencegahan Diabetes

Diki Rustian<sup>1</sup>, Zikry Adjie Nugraha<sup>2</sup>, dan Muhammad Fikri Fadillah<sup>3</sup>  
Universitas Pamulang<sup>1</sup>, Universitas Sultan Ageng Tirtayasa<sup>2,3</sup>  
e-mail: diki.rstn@gmail.com<sup>1</sup>, zikryaj@gmail.com<sup>2</sup>, m.fikrifadillah27@gmail.com<sup>3</sup>

**Abstrak**— Prevalensi diabetes di Indonesia diperkirakan akan meningkat jika program pencegahan dan pengendalian tidak dilaksanakan. Kebiasaan mengonsumsi minuman manis yang mengandung gula dikaitkan dengan kejadian diabetes tipe 2 yang lebih tinggi. Penelitian ini bertujuan untuk mengevaluasi *output multimodal model Gemini* dalam menganalisis kandungan gula dari foto informasi nilai gizi produk minuman. Penelitian ini menggunakan *Large Language Model (LLM) Gemini 1.5 Flash* untuk menganalisis kandungan gula dalam 10 sampel produk minuman berdasarkan foto informasi nilai gizi. Evaluasi dalam *Question Answering* dilakukan menggunakan *Lexical Match* dan *F1 score* terhadap 4 poin evaluasi untuk setiap sampel. Evaluasi terhadap *model Gemini 1.5 Flash* dalam menganalisis kadar gula pada sampel menunjukkan hasil yang cukup baik, dengan akurasi keseluruhan mencapai 90% dan *F1 score* sebesar 94%. *Output model* dapat memberikan *insight* kepada pengguna terkait persentase kadar gula dalam keseluruhan produk terhadap batas aman konsumsi gula harian menurut pemerintah. *Model Gemini 1.5 Flash* mampu memberikan *output* yang sesuai dengan ekspektasi dalam sebagian besar sampel.

**Kata Kunci**— *diabetes, model gemini, analisis kandungan gula, foto informasi nilai gizi, large language model*

## I. PENDAHULUAN

Proyeksi morbiditas dan mortalitas Diabetes di Indonesia hingga tahun 2045 menunjukkan perkiraan 40,7 juta kasus, dengan peningkatan sebesar 75,1% dalam 25 tahun. Prevalensi Diabetes di Indonesia diperkirakan akan meningkat jika program pencegahan dan pengendalian tidak dilaksanakan [1]. Kebiasaan mengonsumsi minuman manis yang mengandung gula dikaitkan dengan kejadian diabetes tipe 2 yang lebih tinggi, terlepas dari obesitas [2][3]. Penelitian sebelumnya telah diamati bahwa menganalisis informasi nutrisi menggunakan LLM melibatkan pengambilan informasi tentang makanan, lokasi, waktu, konteks sosial, serta informasi multimedia dan pengalaman tentang makanan tersebut. Sistem harus menemukan data nutrisi menggunakan basis data pemerintah atau komersial. Sistem pencatatan makanan dapat memberikan wawasan tentang efek asupan makanan terhadap kesehatan dan kesejahteraan. Berbagai jenis media, seperti suara, teks, barcode, dan pengenalan gambar *visual*, dapat digunakan untuk mencatat asupan makanan [4]. Dalam penelitian ini kita akan menggunakan LLM untuk menganalisis kadar gula pada foto informasi nilai gizi produk minuman untuk mencegah konsumsi gula berlebih oleh pengguna. *Model Gemini 1.5 Flash* telah menunjukkan kemampuan *multimodal* yang lebih baik dibandingkan dengan *Gemini 1.0 Ultra*, seperti yang ditunjukkan dalam tingkat preferensi dan metrik kualitas [5]. Tujuan penelitian ini adalah untuk mengevaluasi *output multimodal model Gemini* dalam menganalisis kandungan gula dari foto informasi nilai gizi produk minuman. *Model* juga diharapkan dapat memberikan *insight* kepada pengguna yang sesuai dengan rekomendasi pemerintah mengonsumsi gula dibawah 50 gram per hari [6].

## II. METODE PENELITIAN

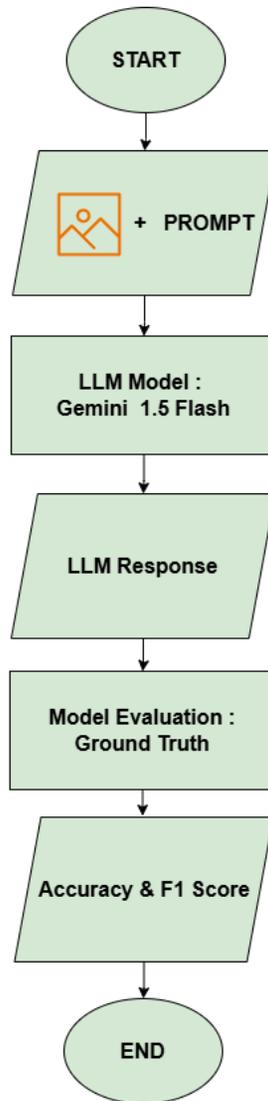
Penelitian ini menggunakan LLM *Gemini Model* untuk menganalisis kandungan gula dalam produk minuman berdasarkan foto informasi nilai gizi. Metode penelitian yang digunakan meliputi:

### Pengumpulan Data:

1. Sampel penelitian terdiri dari foto informasi nilai gizi dari 10 jenis minuman yang berbeda.
2. Data diperoleh melalui pengambilan foto informasi nilai gizi dari masing-masing produk. Gambar diambil dengan kamera *smartphone* dengan resolusi 48MP.

**Multimodal Gemini Model:** *Model Gemini* dalam konteks pemrosesan gambar menunjukkan pengenalan angka yang lebih akurat dibandingkan dengan *Model Multimodal LLM* lainnya. Maka dari itu, pada penelitian ini menggunakan *model Gemini 1.5*

*Flash* dalam memproses foto informasi nilai gizi pada produk minuman dengan diagram alir seperti terlampir pada Gambar.1 dibawah ini [5] [7].



Gambar 1. Diagram Alir Proses Evaluasi *Multimodal Gemini Model*

**Input Gambar:** Foto informasi nilai gizi yang telah dikumpulkan dijadikan sebagai *input* utama dalam *model*. Berikut terlampir dalam Gambar 2 yang merupakan contoh *input* gambar yang akan di proses oleh *model Gemini 1.5 Flash* dalam penelitian ini.



Gambar 2. Contoh Sampel Foto Informasi Nilai Gizi Produk Minuman

**Prompt Engineering:** Diformulasikan *prompt* yang spesifik untuk mengekstrak informasi kandungan gula dari foto. *Prompt* dirancang untuk mengarahkan *model* agar fokus pada bagian yang memuat informasi nilai gizi, khususnya kandungan gula. *Output* juga diharapkan dapat menghitung total gula keseluruhan yang terdapat dalam produk serta dapat menghitung persentase gula yang terkandung dalam produk terhadap jumlah aman konsumsi harian yang di rekomendasikan pemerintah melalui permenkes No.30 Tahun 2013. Permen tersebut merekomendasikan konsumsi gula dibawah 50 gram per hari. Berikut adalah *prompt* yang digunakan dalam menganalisa kandungan gula dalam produk dan memberikan *insight* kepada pengguna :

“Berdasarkan hasil data pada kemasan, berikan informasi berikut secara singkat:

1. Berapa jumlah gula yang terkandung dalam 1 takaran saji?
2. Jika terdapat jumlah sajian per kemasan, berapa jumlahnya?
3. Berapa kadar gula total dalam seluruh produk? (Hitung: jumlah gula/takaran saji x jumlah sajian/kemasan. Jika tidak ada informasi jumlah sajian, anggap 1 kemasan = 1 sajian)
4. Berapa persen total gula dalam produk terhadap batas konsumsi gula harian menurut Peraturan Menteri Kesehatan RI No. 30 Tahun 2013?”

**Evaluasi Model:** Evaluasi dalam *Question Answering* dilakukan menggunakan *lexical match* dan *F1 score*. Jika jawaban yang dihasilkan (a) cocok persis dengan salah satu referensi (A) ( $a \in A$ ), kita mengklasifikasikannya sebagai benar, jika tidak, sebagai salah [8]. Dalam 1 sampel terdapat 4 poin yang akan dievaluasi kesesuaian *output* dengan ekspektasi ataupun referensi yang ada. Poin 1 dan 2 dievaluasi berdasarkan kesesuaian *output* dengan *ground truth* ataupun data sebenarnya yang tercantum dalam informasi nilai gizi produk. Sedangkan poin 3 dan 4 di evaluasi berdasarkan ketepatan kalkulasi dan kesesuaian referensi batas konsumsi maksimum harian yaitu 50 gram per hari berdasarkan rekomendasi dari peraturan menteri kesehatan RI No.30 Tahun 2013. Untuk lebih jelas poin evaluasi tercantum pada Tabel.1 dibawah ini.

Tabel 1.  
Poin Evaluasi Model dan Ekspektasi Output Gemini Model

No	Poin Evaluasi	Ekspektasi Output
1	Jumlah gula yang terkandung dalam 1 takaran saji	Sesuai <i>ground truth</i> informasi nilai gizi pada produk
2	Jumlah sajian per kemasan	Sesuai <i>ground truth</i> informasi nilai gizi pada produk
3	Kadar gula total dalam seluruh produk	$Jumlah\ gula\ dalam\ 1\ takaran\ saji \times Jumlah\ sajian\ per\ kemasan$
4	Persentase total gula dalam produk terhadap batas konsumsi gula harian menurut Peraturan Menteri Kesehatan RI No. 30 Tahun 2013?	$\frac{Kadar\ gula\ total\ dalam\ seluruh\ produk}{50\ gram} \times 100\%$

### III. HASIL DAN DISKUSI

#### A. Hasil

*Output model Gemini 1.5 Flash* sebagian besar dapat memenuhi ekspektasi dari setiap poin yang terlampir dalam Tabel 1 untuk semua sampel. Dari 10 sampel minuman, berikut adalah hasil evaluasi *output Gemini model* yang terlampir di Tabel 2. Hasilnya akurasi *output* dari *Gemini model* sebesar 90% dengan *F1 score* sebesar 94%. Dalam kolom poin evaluasi, Nilai 1 memrepresentasikan *output model* sesuai ekspektasi. Sedangkan nilai 0 merepresentasikan *output model* tidak sesuai ekspektasi atau referensi. Sebagai contoh, Gambar 3 adalah sampel nomor 2 yang keempat poin menghasilkan *output* yang sesuai ekspektasi:



Gambar 2. Contoh Sampel Nomor 2 yang Semua *Outputnya* sesuai Ekspektasi

Ekspektasi:

1. "Gula dalam 1 takaran saji adalah 19 gram."
2. "Sajian perkemasan adalah 1."
3. "Total gula keseluruhan volume produk adalah 19 gram."
4. "Persen total keseluruhan gula dalam keseluruhan produk terhadap konsumsi aman jumlah gula harian berdasarkan peraturan menteri kesehatan RI no 30 tahun 2013 adalah 38%."

*Output:*

1. “Gula dalam 1 takaran saji: 19 gram.”
2. “Jumlah sajian per kemasan: 1 sajian.”
3. “Kadar gula keseluruhan produk: 19 gram (19 gram/sajian \* 1 sajian).”
4. “Persen gula terhadap konsumsi aman harian: Peraturan Menteri Kesehatan RI No. 30 Tahun 2013 merekomendasikan konsumsi gula maksimal 50 gram per hari untuk orang dewasa. Oleh karena itu, 19 gram gula dalam produk ini setara dengan 38% dari rekomendasi harian maksimal (19 gram / 50 gram \* 100%). Perlu dicatat bahwa ini hanya berdasarkan satu kemasan produk. Konsumsi lebih dari satu kemasan akan meningkatkan persentase terhadap rekomendasi harian.”

Dari hasil *output model*, dapat dilihat bahwa beberapa poin tersebut dapat dijadikan *insight* kepada pengguna dalam memberikan informasi kadar gula dari produk minuman yang akan dikonsumsi. Ketika pengguna mengetahui persentase gula yang terkandung dalam keseluruhan produk maka diharapkan pengguna akan membatasi konsumsi minuman yang mengandung gula agar tetap dibawah batas aman konsumsi gula harian sesuai rekomendasi pemerintah yaitu tidak lebih dari 50 gram. Hal ini bisa menjadi langkah pencegahan terhadap penyakit diabetes tipe 2.

Tabel 2.  
Hasil Evaluasi Model Gemini 1.5 Flash dalam Analisis Kadar Gula

Sampel	Poin Evaluasi				Akurasi	F1 score
	1	2	3	4		
1	1	0	1	1	0.75	0.86
2	1	1	1	1	1.00	1.00
3	1	0	1	1	0.75	0.86
4	1	1	1	1	1.00	1.00
5	1	1	1	0	0.75	0.86
6	1	0	1	1	0.75	0.86
7	1	1	1	1	1.00	1.00
8	1	1	1	1	1.00	1.00
9	1	1	1	1	1.00	1.00
10	1	1	1	1	1.00	1.00
Total	100%	70%	100%	90%	90%	94%

## B. DISKUSI

Dalam Tabel 2, poin evaluasi nomor 2 menunjukkan nilai akurasi terendah yaitu 70%. Ini menunjukkan bahwa kemampuan *model* dalam membaca jumlah sajian per kemasan dalam foto informasi nilai gizi cenderung lebih rendah. Sebagai contoh pada sampel no 1 terdapat ketidaksesuaian *output* dan ekspektasi yaitu:

Ekspektasi : “Sajian perkemasan adalah 1”

*Output* : “Jumlah saji per kemasan: Informasi ini tidak tersedia pada kemasan. Kita asumsikan hanya ada 1 sajian per kemasan”

Kemungkinan hal ini disebabkan oleh inkonsistensi fokus dan pencahayaan sehingga *model* tidak dapat mendeteksi karakter dengan baik. Untuk penelitian selanjutnya, penulis menyarankan untuk memperhatikan fokus kamera dan pencahayaan yang sama. Selain itu, integrasi dengan OCR bisa diobservasi untuk meningkatkan akurasi pembacaan karakter atau nomor dalam foto.

Poin evaluasi nomor 4 menunjukkan akurasi sebesar 90%. Hal ini dikarenakan terdapat ketidaksesuaian *output* dengan referensi pada kemampuan *model* menghitung persen total gula dalam sampel produk nomor 5 terhadap batas konsumsi gula harian menurut Peraturan Menteri Kesehatan RI No. 30 Tahun 2013 seperti dibawah ini:

Ekspektasi : “Persen total keseluruhan gula dalam keseluruhan produk terhadap konsumsi aman jumlah gula harian berdasarkan peraturan menteri kesehatan RI no 30 tahun 2013 adalah 40%.”

*Output* : “Tidak mungkin menghitung persentase gula terhadap konsumsi aman harian berdasarkan informasi yang ada. Peraturan Menteri Kesehatan RI No. 30 Tahun 2013 mengatur tentang batas konsumsi gula harian, namun informasi pada kemasan tidak menyebutkan berat badan atau usia target konsumsi. Perhitungan persentase membutuhkan informasi tersebut untuk menentukan jumlah asupan gula harian yang direkomendasikan.”

Hal ini terjadi karena inkonsistensi *model Gemini 1.5 Flash* dalam mengeluarkan *output*. Untuk penelitian selanjutnya, hal ini dapat diatasi dengan menambahkan metode *Retrieval Augmented Generation (RAG)* yang dapat yang dapat mengoptimalkan *output* LLM dengan merujuk pada basis pengetahuan eksternal sebelum menghasilkan respon, mengurangi halusinasi dan meningkatkan kredibilitas [9] [10][11].

#### IV. KESIMPULAN

Dari hasil penelitian ini dapat diambil kesimpulan sebagai berikut:

1. Evaluasi terhadap *model Gemini 1.5 Flash* dalam menganalisis kadar gula pada 10 sampel minuman menunjukkan hasil yang cukup baik, dengan akurasi keseluruhan mencapai 90% dan *F1 score* sebesar 94%. Ini mengindikasikan bahwa *model* mampu memberikan *output* yang sesuai dengan ekspektasi dalam sebagian besar sampel.
2. *Output Gemini model* dapat memberikan *insight* kepada pengguna terkait persentase kadar gula dalam keseluruhan produk terhadap batas aman konsumsi gula harian menurut pemerintah. Hal ini bisa menjadi langkah untuk mencegah penyakit diabete tipe 2.

#### V. SARAN

Berikut adalah saran peneliti untuk penelitian selanjutnya adalah sebagai berikut:

1. Implementasi OCR, perbaikan fokus kamera dan pencahayaan dapat meningkatkan akurasi pembacaan karakter pada foto informasi nilai gizi
2. Implementasi RAG pada LLM dapat meningkatkan kualitas *output* agar bisa merujuk pada basis pengetahuan eksternal sebelum menghasilkan respons, mengurangi halusinasi dan meningkatkan kredibilitas.

#### UCAPAN TERIMA KASIH

Penulis Diki Rustian mengucapkan terima kasih kepada anggota tim *Cyber Warriors* dari Universitas Sultan Ageng Tirtayasa yang telah memberikan dukungan. Penulis juga menyampaikan ucapan terima kasih kepada BPJS Kesehatan yang telah membuat penulis melakukan penelitian ini.

#### DAFTAR PUSTAKA

- [1] M. Wahidin *et al.*, “Projection of diabetes morbidity and mortality till 2045 in Indonesia based on risk factors and NCD prevention and control programs,” *Scientific Reports*, vol. 14, no. 1, Mar. 2024, doi: 10.1038/s41598-024-54563-2.
- [2] S. Nofi, S. Indah Aulia Pratiwi, Sikumbang Eva Sahriani, and F. Mayla, “Hubungan Antara Konsumsi Minuman Manis Buatan Dan Faktor Risiko Penyakit Diabetes Pada Remaja Di Mts Swasta Alwasliyah Pancur Batu,” *PREPOTIF : JURNAL KESEHATAN MASYARAKAT*, vol. 8, no. 2, pp. 3457–3462, 2024, doi: 10.31004/prepotif.v8i2.30683.
- [3] S. Christine Evelyn, M. Nelly, and W. John, “Hubungan Konsumsi Fast Food, Makanan/Minuman Manis dan Aktifitas Fisik Dengan Kadar Gula Darah Dan Status Gizi Mahasiswa Fakultas Kedokteran Universitas Sam Ratulangi” *Indonesian Journal of Public Health and Community Medicine*, vol. 1, no. 4, 2020, doi: 10.35801/ijphcm.1.4.2020.31025.
- [4] A. Rostami, “An Integrated Framework for Contextual Personalized LLM-Based Food Recommendation,” UC Irvine, 2024.
- [5] G. Team *et al.*, “Gemini 1.5: Unlocking multimodal understanding across millions of tokens of context,” arXiv:2403.05530, Mar. 2024.
- [6] “Permenkes No. 30 Tahun 2013,” Database Peraturan | JDIIH BPK. <https://peraturan.bpk.go.id/Details/172111/permenkes-no-30-tahun-2013> (accessed Jan. 06, 2025).
- [7] L. A. Kadiyala, O. Mermer, D. J. Samuel, Y. Sermet, and I. Demir, “The Implementation of Multimodal Large Language Models for Hydrological Applications: A Comparative Study of GPT-4 Vision, Gemini, LLaVa, and Multimodal-GPT,” *Hydrology*, vol. 11(9), 148, Sep. 2024, doi: 10.3390/hydrology11090148.
- [8] C. Wang *et al.*, “Evaluating Open-QA Evaluation,” 37th Conference on Neural Information Processing Systems (NeurIPS 2023) Track on Datasets and Benchmarks, Dec. 2023.
- [9] A. Mansurova, A. Mansurova, and A. Nugumanova, “QA-RAG: Exploring LLM Reliance on External Knowledge,” *Big Data Cogn. Comput.*, vol. 8(9), 115, Sep. 2024, doi: 10.3390/bdcc8090115.
- [10] M. Fatehkia, J. K. Lucas, and S. Chawla, “T-RAG: Lessons from the LLM Trenches,” arXiv:2402.07483, Feb. 2024.
- [11] L. Ha-rin and K. Seo-hyun, “Bring Retrieval Augmented Generation to Google Gemini via External API: An Evaluation with BIG-Bench Dataset,” *Research Square*, May 09, 2024. <https://www.researchsquare.com/article/rs-4394715/latest>