

Comparação de Modelos de IA para Extração de Dados em Glicosímetros

Genivaldo Braynner Teixeira do Carmo¹, Julyanne Maria dos Santos Correia¹, Ronaldo Rodrigues da Silva Filho¹, Pablo Azevedo Sampaio¹, and Robson Wagner Albuquerque de Medeiros¹

¹Departamento de Computação, Universidade Federal Rural de Pernambuco, Brasil

Resumo

A diabetes é uma condição crônica que requer monitoramento constante dos níveis de glicose no sangue, sendo essencial o uso de glicosímetros para a obtenção dessas informações. Este trabalho tem como objetivo comparar três modelos de Inteligência Artificial, Gemini, GPT-4o e Llava 1.5, para identificar qual deles extrai, de forma mais eficaz, os dados de glicose, data e hora dos glicosímetros. Utilizando técnicas de engenharia de *prompt*, busca-se aprimorar a precisão e eficiência na extração desses dados, otimizando o monitoramento e contribuindo para a melhor gestão da saúde de pacientes diabéticos.

Palavras-chave: Inteligência Artificial — Engenharia de *Prompt* — Monitoramento de Glicose.

1 Introdução

A diabetes é uma das doenças crônicas mais prevalentes no mundo, afetando milhões de pessoas de diferentes faixas etárias [1] [2]. Trata-se de uma condição que interfere na forma como o corpo metaboliza a glicose, a principal fonte de energia para as células [3]. Independentemente do tipo, a diabetes exige um controle constante dos níveis de glicose no sangue para garantir a estabilidade metabólica e prevenir complicações a longo prazo [4].

O controle adequado da diabetes é vital, não apenas para evitar complicações imediatas, como hipoglicemia (quando o nível de glicose está muito baixo) [5] ou hiperglicemia

(quando o nível está muito alto) [6], mas também para prevenir uma série de problemas de saúde que podem se desenvolver ao longo do tempo. Doenças cardiovasculares [7], danos aos nervos (neuropatia), problemas renais (nefropatia) e complicações oculares (retinopatia diabética) são apenas algumas das consequências potenciais de um controle inadequado [8]. Estudos mostram que o monitoramento rigoroso da glicemia, juntamente com intervenções oportunas, pode reduzir drasticamente o risco dessas complicações, tornando o cuidado diário indispensável [9].

Nesse cenário, os glicosímetros desempenham um papel central. Esses dispositivos eletrônicos permitem que os pacientes meçam rapidamente seus níveis de glicose, utilizando uma pequena amostra de sangue [10]. A facilidade de uso e a portabilidade dos glicosímetros revolucionaram o cuidado com a diabetes, proporcionando aos pacientes uma ferramenta eficiente para monitorar sua condição em tempo real [11]. No entanto, apesar de serem amplamente acessíveis, o gerenciamento dos dados obtidos pelos glicosímetros pode ser uma tarefa cansativa e repetitiva para o paciente, já que ele precisa anotar manualmente cada medição ao longo do dia para monitorar seus níveis de glicose. Essa rotina, embora essencial, pode se tornar desgastante e propensa a erros, comprometendo a precisão do controle glicêmico. A integração de tecnologias avançadas, como a Inteligência Artificial (IA), pode fornecer uma solução para otimizar a análise e o uso desses dados, transformando a gestão da diabetes.

Este trabalho propõe uma abordagem inovadora para o uso de IAs na extração de dados de glicosímetros, focando na comparação de três modelos distintos: Gemini, GPT-4o e Llava 1.5. O objetivo é determinar qual desses modelos se destaca na precisão e eficiência ao extrair informações críticas dos glicosímetros, como os níveis de glicose, data e hora das medições. Para isso, serão aplicadas técnicas de engenharia de *prompt* [12], que visam estruturar as consultas de IA de maneira a maximizar a qualidade das respostas. Através dessa análise, espera-se não apenas identificar o melhor modelo de IA para essa função, mas também contribuir para o avanço do monitoramento de diabetes, tornando o processo mais automatizado, preciso e acessível.

A estrutura deste trabalho está organizada da seguinte forma: na próxima seção, será apresentado um breve contexto sobre engenharia de *prompt*, explicando a importância dessa técnica para o desenvolvimento do projeto, o conceito de IA Generativa, destacando sua relevância para o tema e os tipos de glicosímetros disponíveis no mercado e à justificativa da escolha do dispositivo utilizado neste estudo. A terceira seção explora os trabalhos relacionados, que serviram como base para esta pesquisa. Em seguida, na quarta seção, descreverei a metodologia utilizada, detalhando cada etapa do processo com os *prompts* usados. Na quinta seção, será demonstrada a implementação dos *prompts* usados e também do GPT-4o no aplicativo Android. Na sexta seção, serão discutidos os

resultados obtidos com o projeto. Na sétima seção, apresentarei as conclusões. Por fim, na oitava seção, serão abordadas sugestões e possíveis direções para trabalhos futuros.

2 Conceitos Preliminares

Esta seção apresenta de forma breve alguns conceitos fundamentais que são essenciais para o entendimento deste trabalho. Esses conceitos fornecerão a base teórica necessária para compreender as metodologias e os resultados apresentados ao longo do trabalho.

2.1 Inteligências Artificiais Generativas

As Inteligências Artificiais Generativas estão revolucionando o campo da tecnologia, sendo cada vez mais reconhecidas por sua capacidade única de criar novos conteúdos a partir dos dados em que foram treinadas [13]. Ao contrário das abordagens tradicionais de IA, que seguem instruções pré-definidas para resolver problemas específicos, as IAs generativas, como o nome sugere, podem gerar textos, imagens, músicas, entre outros tipos de conteúdo, de forma autônoma e adaptável. Isso abre uma série de novas possibilidades, pois essas IAs são capazes de produzir resultados que muitas vezes vão além do que foi fornecido nos dados originais.

Uma das tecnologias mais notáveis nesse campo são as Redes Gerativas Adversariais (GANs) [14], que trabalham com dois sistemas competindo entre si: o gerador e o discriminador. O gerador cria novos exemplos de dados, enquanto o discriminador tenta determinar se esses exemplos são reais ou gerados. Esse ciclo de competição constante permite que o gerador melhore continuamente, criando dados cada vez mais realistas. Além disso, temos os modelos de Transformadores, como o GPT, que são especializados na geração de linguagem natural. O GPT, por exemplo, pode gerar textos coerentes e relevantes a partir de comandos simples, sendo amplamente utilizado para tarefas que envolvem linguagem.

Essas IAs geram grande interesse devido às suas aplicações práticas em várias áreas. Elas já são usadas na criação de arte digital, na geração automática de conteúdo para marketing, no desenvolvimento de histórias, e até na programação de código. Contudo, sua utilidade vai muito além dessas aplicações criativas. Um exemplo importante é a possibilidade de utilizar IA generativa no campo da saúde, especialmente no controle de doenças como a diabetes. O monitoramento dos níveis de glicose no sangue é um processo essencial para pacientes diabéticos, mas pode se tornar uma tarefa repetitiva e cansativa. A extração manual dos dados dos glicosímetros, seguida de anotações diárias, pode ser propensa a erros ou falhas.

É aqui que a IA generativa pode fazer uma grande diferença. Ao utilizar essa tecnologia, é possível automatizar a extração dos dados de glicosímetros de forma eficiente e precisa. Em vez de depender de entradas manuais, a IA pode coletar e interpretar os dados, organizando-os de maneira útil e clara. Por exemplo, ao utilizar modelos de IA, os dados de glicose, data e hora das medições podem ser extraídos de maneira rápida e precisa, permitindo que o paciente ou o profissional de saúde identifique facilmente padrões de variação glicêmica ao longo do tempo. Isso não só melhora o controle diário da diabetes, mas também facilita a tomada de decisões sobre o tratamento, tornando o processo muito mais prático.

Além disso, a IA generativa pode oferecer uma visão mais detalhada dos dados. Não se trata apenas de registrar os números de glicose, mas de interpretar esses números e fornecer informações valiosas sobre as tendências do paciente, como o impacto de certos alimentos, atividades ou medicamentos nos níveis de glicose. Isso traz uma camada extra de precisão e eficiência no monitoramento, que pode ser crucial para o controle da doença.

Portanto, ao explorar as capacidades das IAs generativas no contexto dos glicosímetros, podemos transformar o gerenciamento da diabetes. Ao reduzir a carga manual e oferecer dados mais detalhados, essa tecnologia pode proporcionar um controle mais eficaz, melhorando a qualidade de vida dos pacientes. Mesmo que ainda existam desafios a serem superados, como o custo elevado e o acesso limitado à tecnologia, o potencial dessas IAs na área da saúde é inegável [15].

2.2 Engenharia de *Prompt*

A engenharia de *prompt* (*Prompt Engineering*) surge como uma ferramenta essencial no desenvolvimento e refinamento de modelos de IA generativa, permitindo que eles gerem respostas mais precisas, coerentes e livres de vieses. O poder dessa técnica reside na qualidade dos *prompts*, que são as instruções textuais que orientam o modelo em sua tarefa de gerar informações. Quanto mais refinado e específico o *prompt*, maior a chance de obter resultados que atendam às expectativas do usuário.

A técnica de engenharia de *prompt* envolve um processo contínuo de ajuste e experimentação. Isso significa que, ao longo de diversas iterações, os *prompts* são modificados para alcançar maior clareza, relevância e precisão nas respostas geradas pela IA. Entre as principais estratégias utilizadas estão a definição clara de contexto, o fornecimento de exemplos específicos, a delimitação de restrições, e a elaboração de perguntas ou comandos que direcionem a IA a focar em determinados aspectos do problema.

Por exemplo, ao fornecer um *prompt* mais detalhado com instruções claras, o modelo é capaz de entender melhor o que é esperado dele, evitando ambiguidades ou confusões que poderiam resultar em respostas menos precisas. Além disso, a especificidade de termos

utilizados no *prompt* pode fazer toda a diferença: termos vagos ou genéricos tendem a gerar respostas amplas e irrelevantes, enquanto um *prompt* específico resulta em saídas mais focadas e úteis [16].

Outra técnica importante é a introdução de "condicionamentos" no *prompt*, onde se oferece à IA parâmetros que limitam ou expandem as respostas [17]. Isso é particularmente útil para minimizar vieses, garantindo que o modelo mantenha coerência sem se desviar de temas importantes ou sensíveis. Além disso, o ajuste do tom ou da linguagem pode influenciar diretamente na qualidade da resposta, garantindo que a IA responda de acordo com a formalidade ou o estilo desejado.

Para este trabalho, a engenharia de *prompt* se mostrou fundamental para explorar o potencial das IAs e garantir sua aplicação eficiente. Através do uso dessas técnicas, conseguimos moldar a maneira como as IAs capturam e processam os dados de glicosímetros, maximizando a precisão e relevância das informações extraídas, aplicando-as em diferentes contextos e demandas com eficácia.

2.3 Tipos de Glicosímetros

Os glicosímetros desempenham um papel crucial no gerenciamento da Diabetes Mellitus Tipo 1 (DM1), pois permitem que os pacientes monitorem seus níveis de glicose no sangue de maneira prática e precisa. Com os avanços tecnológicos, diferentes tipos de glicosímetros foram desenvolvidos, cada um oferecendo características e funcionalidades específicas que se ajustam às variadas necessidades dos pacientes.

Esses dispositivos podem ser classificados em diversas categorias, dependendo do método utilizado para medir a glicose. Os glicosímetros convencionais, por exemplo, funcionam através da coleta de uma pequena amostra de sangue, geralmente obtida por meio de uma picada no dedo, para realizar a medição. Esses dispositivos são amplamente aceitos devido à sua precisão e custo acessível. Por outro lado, os monitores contínuos de glicose (CGMs, do inglês *Continuous Glucose Monitors*) acompanham os níveis de glicose em tempo real, fornecendo leituras constantes ao longo do dia e alertando o paciente sobre possíveis alterações anormais nos níveis de glicose. Este tipo de glicosímetro é especialmente valioso para pacientes que necessitam de um monitoramento mais rigoroso. Além disso, existem glicosímetros não invasivos, ainda em desenvolvimento ou disponíveis em versões limitadas, que medem a glicose através da pele ou do fluido intersticial, eliminando a necessidade de amostras de sangue.

Cada tipo de glicosímetro possui vantagens e desvantagens que devem ser cuidadosamente avaliadas com base no perfil e nas necessidades de cada paciente. A escolha do glicosímetro adequado pode ter um impacto significativo no controle da glicemia e, conseqüentemente, na qualidade de vida dos pacientes com DM1.

Para este trabalho, utilizarei como base um glicosímetro convencional. O objetivo é extrair informações como a data, a hora e o valor da glicemia a partir das leituras realizadas por esse dispositivo. Essas informações serão capturadas e processadas utilizando técnicas avançadas de Inteligência Artificial, que serão posteriormente integradas a um aplicativo móvel para auxiliar no controle glicêmico de pacientes com Diabetes Mellitus Tipo 1 (DM1). A escolha de um glicosímetro convencional se justifica pela sua ampla utilização e confiabilidade na medição dos níveis de glicose, proporcionando dados consistentes e precisos que são essenciais para a gestão adequada da doença.

Abaixo está representado o glicosímetro convencional que será utilizado como base para este estudo. Este dispositivo permitirá a extração de dados essenciais como a glicemia, a data e a hora das medições, elementos cruciais para o desenvolvimento e a integração da funcionalidade no aplicativo móvel proposto.



Figura 1: Glicosímetro convencional utilizado para a extração de dados de glicemia, data e hora. [18]

3 Trabalhos Relacionados

Nesta seção, serão discutidos trabalhos que utilizam técnicas semelhantes às empregadas neste projeto, com foco na extração de dados de dispositivos e no uso de Inteligência Artificial (IA) para otimizar a eficiência do processo. Um estudo é o de H. H. Shomee e A. Sams (2021) [19], que desenvolveram um sistema de detecção e reconhecimento de placas de veículos utilizando um modelo de aprendizado profundo baseado na arquitetura YOLO.

Esse sistema automatiza a extração de informações de placas e aplica redes adversárias generativas (GANs) para aprimorar a qualidade das imagens. Embora o foco desse estudo esteja na detecção de placas, a ideia de usar IA para otimizar a extração de dados se alinha diretamente ao objetivo deste projeto, que busca automatizar a coleta de dados de glicosímetros.

Adicionalmente, Margaret Mitchell et al. (2012) [20] propuseram o sistema Midge, um modelo de geração de descrições de imagens com base em detecções de visão computacional. Esse sistema utiliza a co-ocorrência de palavras com base na sintaxe para filtrar detecções ruidosas e gerar descrições detalhadas. Apesar de focado na geração de descrições de imagens, o método de extração de informações relevantes pode ser adaptado ao contexto da extração de dados de glicosímetros, especialmente em termos de processar grandes volumes de dados.

Por fim, Narek Tumanyan et al. (2023) [21] apresentaram uma estrutura inovadora que explora modelos de difusão de texto-para-imagem para realizar tradução de imagem para imagem guiada por texto. Essa abordagem permite modificar imagens mantendo seu layout semântico original, o que demonstra como modelos generativos podem ser aplicados de forma flexível no controle e manipulação de dados visuais complexos. Embora o estudo tenha foco em imagens, o conceito de controle preciso sobre os dados pode ser aplicado à extração e análise de informações de dispositivos médicos, como glicosímetros.

4 Metodologia

Para a avaliação e comparação dos modelos de IA (GPT-4o, Gemini e Llava 1.5), foi implementado um processo detalhado que inclui a geração de *prompts* específicos para a extração dos dados de glicosímetros, incluindo valores de glicose, data e hora. O principal objetivo foi determinar qual modelo apresentaria a melhor eficiência, considerando não apenas a precisão dos resultados, mas também a eficiência no uso de tokens, uma vez que o uso excessivo de tokens pode impactar o custo e a velocidade do processo. A seguir, descrevo os passos adotados e os *prompts* gerados para cada modelo:

4.1 Preparação dos Dados

Foram selecionadas 100 imagens de diferentes marcas¹ de glicosímetros, incluindo G-Tech, Accu-Chek, OneTouch, Oncall plus II e GlicoCheck, para garantir uma ampla diversidade de dispositivos. O processo de preparação envolveu ajustes na qualidade das imagens para evitar ruídos visuais, redimensionamento padronizado e validação de que as informações

¹https://github.com/BraynnerTeixeira/test_imagens

exibidas (valor da glicose, data e hora) estivessem legíveis e claras para a análise pelos modelos de IA.

4.2 Modelos Avaliados

Três modelos de IA foram utilizados para a extração dos dados:

- GPT-4o: Um modelo de linguagem natural de última geração, treinado para compreender e gerar texto. O GPT-4o foi selecionado por ser amplamente reconhecido por seu desempenho consistente em uma ampla gama de contextos, além de sua facilidade de integração com sistemas e APIs.
- Gemini: Um modelo similar ao GPT-4o, focado em otimizar a extração de informações. Foi escolhido por seu potencial em gerar resultados precisos em diferentes cenários de análise de dados, e por ser uma alternativa de alta performance que oferece capacidades comparáveis ao GPT-4o, mas com foco em extração e processamento de dados mais específicos.
- Llava 1.5: Um modelo mais específico, com certas limitações, como a necessidade de usar a língua inglesa para as interações, mas que é facilmente implantado em qualquer máquina, tornando-se uma opção prática e acessível em diferentes ambientes de implementação.

4.3 Geração e Avaliação dos *Prompts*

Para cada modelo, diferentes *prompts* foram gerados com o objetivo de extrair os dados dos glicosímetros da forma mais eficiente possível. A escolha dos *prompts* seguiu os princípios da engenharia de prompt, que visa adaptar os comandos para maximizar a precisão e relevância das respostas, com o menor uso de tokens possível.

- GPT-4o:
 1. Extraia os valores de glicoses desse glicosímetro com a sua respectiva data e hora.
 2. Analise a imagem do glicosímetro e diga qual o valor da glicose, data e hora.
 3. Valor glicose com data e hora.
- Gemini:
 1. Dada a imagem, qual o valor da glicose com a data e a hora?.
 2. Extraia desse glicosímetro: A glicose, a data e hora.

3. Valor glicose com data e hora.

- Llava 1.5:

1. *Extract from the image that was sent: Glucose, date and time.*
2. *What is the glucose value, date and time shown on the glucometer?.*
3. *What is the glucose value with the date and time?.*

Diferentemente dos outros modelos, o Llava 1.5 precisou ser utilizado em inglês, pois suas respostas em português apresentaram uma menor precisão. Esse aspecto foi considerado na avaliação geral do modelo.

4.4 Critérios de Avaliação

A escolha do melhor *prompt* para cada modelo foi baseada em dois critérios principais:

- Precisão dos dados: O *prompt* deveria retornar corretamente o valor da glicose, a data e a hora.
- Eficiência no uso de tokens: O *prompt* que utilizasse o menor número de tokens foi selecionado, uma vez que isso reduz os custos e o tempo de processamento sem comprometer a qualidade da resposta.

5 Avaliação dos modelos de IA

A implementação dos *prompts* foi uma etapa crucial no processo de integração dos modelos de IA ao sistema, sendo necessário testar diferentes abordagens para otimizar a extração dos dados de glicose, data e hora dos glicosímetros. Além de avaliar a precisão na extração das informações, foi fundamental quantificar a eficiência no uso de *tokens* para cada modelo e *prompt*. Essa quantificação permitiu avaliar o custo computacional, considerando que o uso excessivo de *tokens* poderia aumentar o custo operacional e impactar negativamente o desempenho do sistema. Assim, ao comparar o número de *tokens* utilizados por modelo e *prompt*, foi possível identificar quais abordagens ofereciam o melhor equilíbrio entre economia de recursos e exatidão dos dados extraídos.

5.1 GPT-4o

Para o modelo GPT-4o, foram testados três *prompts* diferentes, com o objetivo de extrair as informações de glicose, data e hora de forma precisa e eficiente. A seguir estão os *prompts* avaliados:

- *Prompt 1*: "Extraia os valores de glicose desse glicosímetro com a sua respectiva data e hora." Resultado: Este *prompt* apresentou alta precisão e retornou corretamente todos os valores solicitados. No entanto, o uso de tokens foi elevado, o que poderia comprometer a eficiência do sistema a longo prazo.
- *Prompt 2*: "Analisar a imagem do glicosímetro e diga qual o valor da glicose, data e hora." Resultado: Assim como o primeiro *prompt*, este também obteve sucesso na extração dos dados, mas novamente o uso de tokens foi elevado.
- *Prompt 3*: "Valor glicose com data e hora." Resultado: Este foi o *prompt* mais eficiente, utilizando menos tokens e retornando os dados corretamente. Devido à sua eficiência, tanto em termos de precisão quanto de economia de tokens, este *prompt* foi o selecionado para ser utilizado na versão final da aplicação.

5.2 Gemini

Para o modelo Gemini, a estratégia seguiu um processo similar ao utilizado no GPT-4o, com foco em encontrar o melhor *prompt* que equilibrasse precisão e economia de tokens:

- *Prompt 1*: "Dada a imagem, qual o valor da glicose com a data e a hora?" Resultado: O *prompt* foi bem-sucedido, retornando os dados corretamente, mas com alto uso de tokens.
- *Prompt 2*: "Extraia desse glicosímetro: A glicose, a data e hora." Resultado: Assim como o primeiro, este *prompt* retornou as informações corretamente, mas também com um uso elevado de tokens.
- *Prompt 3*: "Valor glicose com data e hora." Resultado: Este *prompt*, por utilizar menos tokens e continuar a fornecer os dados precisos, foi o escolhido para a versão final.

5.3 Llava 1.5

Diferente dos outros modelos, o Llava 1.5 apresentou algumas limitações, especialmente em relação à precisão dos dados extraídos. Foi necessário utilizar os *prompts* em inglês, pois as respostas em português não apresentaram o desempenho esperado. Os *prompts* avaliados foram:

- *Prompt 1*: "*Extract from the image that was sent: Glucose, date and time.*" Resultado: Este *prompt* só retornou o valor da glicose, sem a data e a hora.

- *Prompt 2*: ” *What is the glucose value, date and time shown on the glucometer?*” Resultado: Semelhante ao primeiro, este *prompt* também só retornou o valor da glicose, sem as demais informações.
- *Prompt 3*: ” *What is the glucose value with the date and time?*” Resultado: Este foi o *prompt* que obteve os melhores resultados, embora tenha falhado em algumas imagens. Em certas ocasiões, retornou a glicose, data e hora corretamente.

5.4 Critérios de Escolha

A escolha dos *prompts* para a implementação final foi baseada em dois critérios principais:

1. Precisão dos dados: Os *prompts* que retornaram o valor da glicose juntamente com a data e a hora de forma precisa foram priorizados.
2. Eficiência no uso de tokens: Como o uso excessivo de tokens impacta diretamente o custo e o tempo de processamento da aplicação, foram escolhidos os *prompts* que utilizaram menos tokens, sem comprometer a qualidade dos dados extraídos.

Dessa forma, os *prompts* selecionados para GPT-4o e Gemini foram aqueles que combinaram precisão e menor uso de tokens, enquanto para o Llava 1.5, apesar das limitações, foi escolhido o *prompt* que apresentou os melhores resultados no idioma inglês.

5.5 Análise Comparativa dos Modelos na Leitura do Glicosímetro

A avaliação dos modelos de IA foi aprofundada para entender não apenas a eficácia dos *prompts*, mas também a capacidade intrínseca de cada modelo em interpretar e extrair corretamente os dados apresentados nos glicosímetros. Nesta subseção, comparamos a eficiência dos modelos GPT-4o, Gemini e Llava 1.5 na tarefa específica de leitura dos valores de glicose, data e hora a partir das imagens, e discutimos a adequação de cada um para aplicações de monitoramento de glicose.

Todos os modelos foram testados sob as mesmas condições, e a medição foi padronizada, considerando todas as marcas de glicosímetros mencionadas.

5.5.1 Capacidade de Reconhecimento e Interpretação de Imagens

O GPT-4o, demonstrou uma elevada capacidade de interpretar imagens de glicosímetros, extraíndo com precisão os valores de glicose, data e hora. O modelo foi consistente mesmo em imagens com variações de iluminação, ângulos ou pequenas obstruções, indicando uma robusta compreensão visual e contextual das informações apresentadas.

Tabela 1: GPT-4o

Marcas	Precisão da Glicose	Data e Hora
G-Tech	100%	100%
Accu Chek	100%	98%
OneTouch	100%	100%
Oncall plus II	100%	97%
GlicoCheck	100%	100%

O modelo Gemini, apresentou uma eficácia satisfatória na interpretação das imagens, conseguindo extrair os dados. Contudo, pode-se notar que em imagens com qualidade inferior ou condições menos ideais, o tempo de resposta foi maior. Mesmo com a redução no número de tokens no *prompt* otimizado (“Valor glicose com data e hora.”), o modelo levou mais tempo para processar essas imagens, o que indica uma sensibilidade maior às condições de captura.

Tabela 2: Gemini

Marcas	Precisão da Glicose	Data e Hora
G-Tech	100%	100%
Accu Chek	100%	100%
OneTouch	100%	100%
Oncall plus II	100%	99%
GlicoCheck	100%	100%

Já o Llava 1.5, enfrentou dificuldades significativas na interpretação das imagens dos glicosímetros. O modelo frequentemente não conseguia extrair todas as informações necessárias, mostrando limitações na capacidade de reconhecimento visual, especialmente em relação à data e hora.

Tabela 3: Llava 1.5

Marcas	Precisão da Glicose	Data e Hora
G-Tech	85%	1%
Accu Chek	70%	0%
OneTouch	75%	3%
Oncall plus II	60%	0%
GlicoCheck	72%	0%

5.5.2 Comparação de Desempenho dos *Prompts* do GPT-4o

Nesta subseção, será analisado os gráficos que mostram o desempenho das requisições de três diferentes *prompts* utilizados com o modelo GPT-4o. Os gráficos apresentam o

tempo de resposta em diferentes requisições, o que nos permite avaliar a consistência e a eficiência de cada *prompt*.

Para realizar as medições, utilizou-se o Postman, uma ferramenta de teste de APIs que registrou o tempo exato de cada requisição até a extração dos dados.

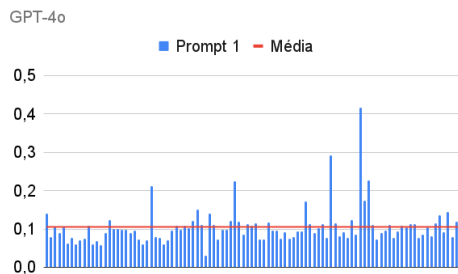


Figura 2: GPT-4o - *Prompt 1*

O primeiro gráfico (Figura 2) exibe as requisições do *Prompt 1*, que, embora tenha retornado os dados de forma precisa, apresentou tempos de resposta elevados em várias requisições, com um pico que ultrapassou 0,4s. A média geral do tempo de resposta, representada pela linha vermelha, ficou em torno de 0,105s, indicando uma variação considerável nas requisições. Esses picos podem impactar a experiência do usuário em aplicações que exigem rapidez na resposta, como no monitoramento contínuo de glicose.

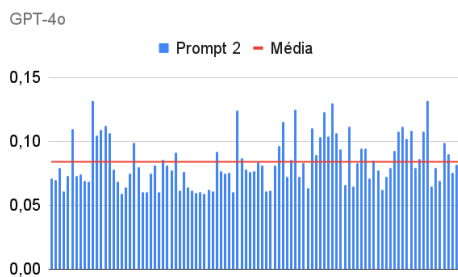


Figura 3: GPT-4o - *Prompt 2*

O gráfico do *Prompt 2* (Figura 3) mostra uma melhora em relação ao primeiro, com a maioria das requisições abaixo de 0,10s. No entanto, a média geral ainda permanece abaixo dessa marca, o que indica que, apesar de mais estável que o *Prompt 1*, o *Prompt 2* ainda apresenta variações que podem comprometer a eficiência do sistema em cenários que demandam alta performance. A distribuição dos tempos de resposta, apesar de mais controlada, ainda apresenta picos que ultrapassam os 0,10s.

O *Prompt 3* (Figura 4), por sua vez, apresentou o melhor desempenho em termos de tempo de resposta. Como mostrado no gráfico, a maioria das requisições ficaram abaixo da média de 0,00067 segundos. Isso faz com que o *Prompt 3* seja o mais eficiente, com

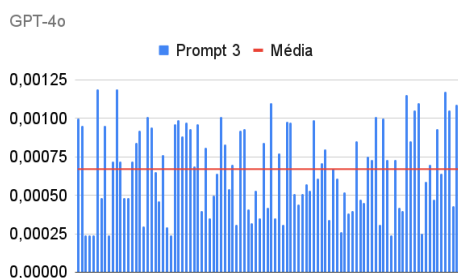


Figura 4: GPT-4o - *Prompt 3*

tempos de resposta mais rápidos e consistentes em comparação com os outros *prompts*. Essa eficiência torna o *Prompt 3* ideal para aplicações que precisam de respostas rápidas e consistentes, como o monitoramento de glicose em tempo real.

Com base nos gráficos apresentados, é evidente que o *Prompt 3* é o mais adequado para a aplicação proposta, já que oferece o melhor equilíbrio entre precisão e tempo de resposta. A eficiência do *Prompt 3* em termos de consistência e baixa variação nas requisições destaca-se em relação aos *Prompts 1* e *2*, que, apesar de precisos, apresentam variações significativas nos tempos de resposta.

5.5.3 Comparação de Desempenho dos *Prompts* do Gemini

Os gráficos a seguir ilustram o desempenho do modelo Gemini em termos de tempo de resposta para três diferentes *prompts*, medidos em segundos. Esses dados nos permitem avaliar a eficiência de cada *prompt* e identificar qual foi o mais adequado para a extração precisa e eficiente dos dados.

Para realizar as medições com o modelo Gemini, foi utilizado o playground do Google AI Studio.

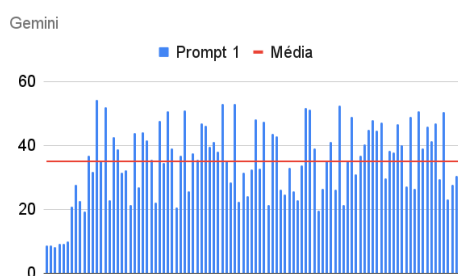


Figura 5: Gemini - *Prompt 1*

O gráfico do *Prompt 1* (Figura 5) revela tempos de resposta relativamente altos e inconsistentes, medidos em segundos. Em várias requisições, o tempo de resposta ultrapassou 40 segundos, com uma média geral bastante elevada em comparação aos outros

prompts. As variações entre as requisições foram significativas, indicando que esse *prompt* pode não ser o mais confiável para aplicações que exigem respostas rápidas e consistentes.

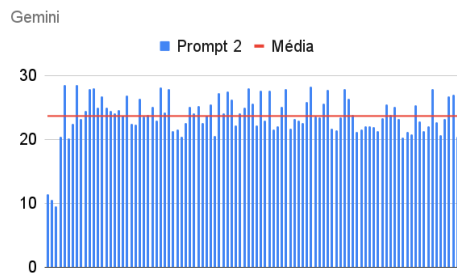


Figura 6: Gemini - *Prompt 2*

No gráfico do *Prompt 2* (Figura 6), observamos uma melhora no tempo de resposta em relação ao *Prompt 1*. A maioria das requisições ficou abaixo de 30 segundos, e a média geral estabilizou-se pouco acima de 20 segundos. Embora tenha apresentado uma performance mais consistente, ainda existem alguns picos que superam os 25 segundos, sugerindo que esse *prompt*, apesar de mais eficiente que o primeiro, ainda não é o ideal para otimizar a velocidade de resposta.

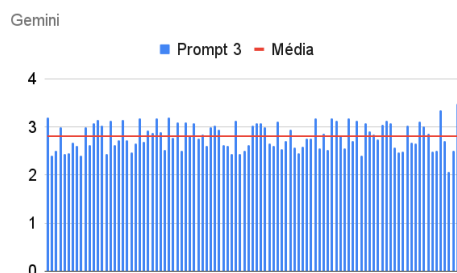


Figura 7: Gemini - *Prompt 3*

O *Prompt 3* (Figura 7) se destaca como o mais eficiente, com tempos de resposta notavelmente baixos e consistentes, medidos em segundos. A maior parte das requisições ficou em torno de 3 segundos, com uma média muito próxima desse valor. Esse gráfico indica que o *Prompt 3* foi o mais estável e rápido, apresentando tempos de resposta significativamente menores do que os outros dois *prompts*. Essa performance faz com que o *Prompt 3* seja o mais adequado para ser utilizado em uma aplicação de monitoramento contínuo, onde a rapidez e consistência são essenciais.

É possível observar que o *Prompt 3* foi o mais eficiente em termos de tempo de resposta para o modelo Gemini, considerando os resultados em segundos. Enquanto os *Prompts 1* e *2* apresentaram variações significativas e tempos de resposta mais elevados, o *Prompt 3* demonstrou tempos baixos e consistentes, tornando-o a escolha ideal para o monitora-

mento de glicose em tempo real. Essa estabilidade sugere que o *Prompt 3* oferece o melhor equilíbrio entre precisão e eficiência, atendendo às necessidades da aplicação proposta.

5.5.4 Comparação de Desempenho dos *Prompts* do Llava 1.5

Os gráficos a seguir mostram o tempo de resposta para três diferentes *prompts* utilizados com o modelo Llava 1.5. Cada gráfico nos permite avaliar a consistência e a eficiência de cada *prompt*, com base no tempo de processamento medido em segundos.

O modelo Llava 1.5 foi avaliado utilizando o Google Colab para as medições.

No gráfico do *Prompt 1* (Figura 8), observamos tempos de resposta relativamente altos, com valores chegando a quase 6 segundos. A média de resposta ficou em torno de 4,64 segundos, o que indica um desempenho abaixo do esperado para uma aplicação que requer rapidez na extração de dados. Além disso, há uma grande variação nas requisições, tornando este *prompt* inconsistente e menos confiável para uma aplicação prática.

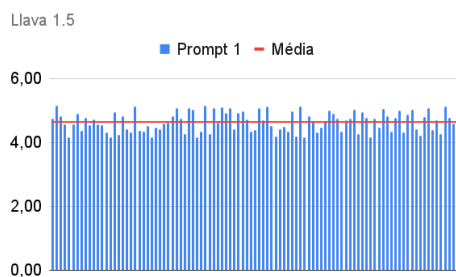


Figura 8: Llava 1.5 - *Prompt 1*

O gráfico do *Prompt 2* (Figura 9) mostra uma leve melhora em relação ao *Prompt 1*. A maioria das requisições teve um tempo de resposta entre 3 e 4 segundos, com a média também estabilizada em torno de 3,40 segundos. Embora tenha apresentado menos variação e uma performance mais estável que o primeiro *prompt*, o *Prompt 2* ainda não conseguiu atingir o nível de eficiência ideal para uma aplicação que exige rapidez.

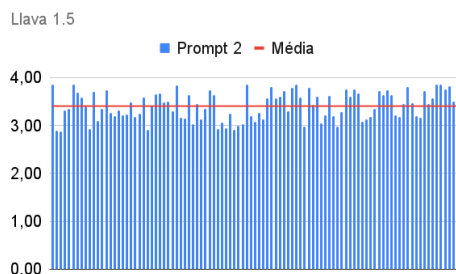


Figura 9: Llava 1.5 - *Prompt 2*

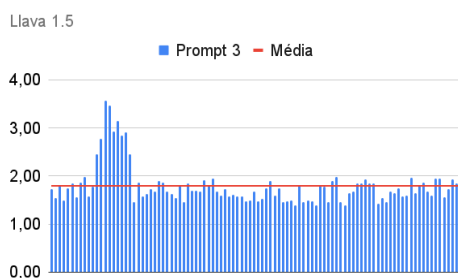


Figura 10: Llava 1.5 - *Prompt 3*

O *Prompt 3* (Figura 10) foi claramente o mais eficiente, com a maioria dos tempos de resposta estabilizados em torno de 2 segundos. A média ficou consistentemente abaixo de 1,79 segundos, com menos variações significativas entre as requisições. Esse gráfico demonstra que o *Prompt 3* não apenas teve os tempos de resposta mais rápidos, mas também foi o mais consistente, tornando-o o melhor candidato para a implementação em uma aplicação de monitoramento contínuo de glicemia.

5.5.5 Resultados das Médias de Tempo de Resposta

A seguir, será apresentado a análise dos gráficos que mostram a média de tempo de resposta para cada *prompt* dos modelos GPT-4o, Gemini, e Llava 1.5, expressa em segundos. Esses dados são essenciais para avaliar a eficiência de cada modelo em termos de velocidade, complementando a análise de precisão já mencionada anteriormente. Além disso, é importante observar o erro padrão de cada modelo, que indica a variação dos tempos de resposta em relação à média, e nos ajuda a entender a consistência das respostas.

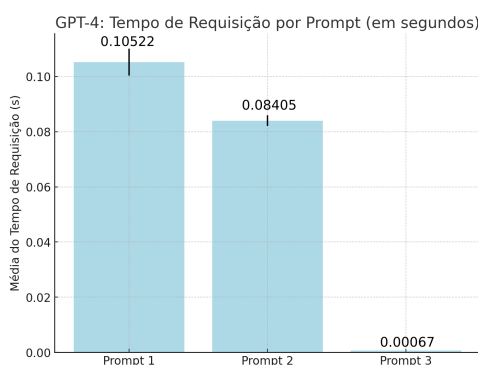


Figura 11: Tempo médio dos *prompts* do GPT-4o

Analisando o erro padrão entre os *prompts* do GPT-4o (Figura 11), podemos observar que o *Prompt 3* é não só o mais rápido, mas também o mais consistente, apresentando a menor variação nos tempos de resposta. Isso significa que, além de ser eficiente em termos de velocidade, ele também mantém uma estabilidade muito maior em suas execuções. Em

contraste, o *Prompt 1*, que foi o mais lento, apresentou a maior variação nos tempos, o que o torna menos confiável em cenários que exigem alta consistência. O *Prompt 2* teve um desempenho intermediário, com uma variação menor que o *Prompt 1*, mas ainda assim inferior à consistência do *Prompt 3*. Portanto, podemos concluir que o *Prompt 3* é a escolha ideal não apenas por sua rapidez, mas também por sua alta previsibilidade e estabilidade em diferentes execuções.

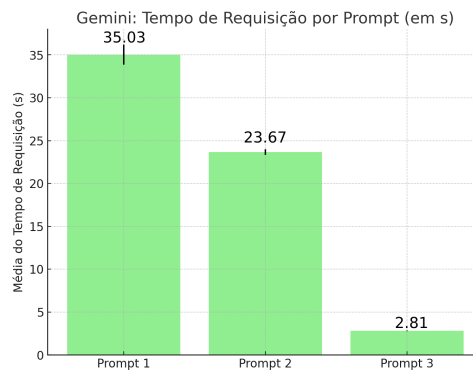


Figura 12: Tempo médio dos *prompts* do Gemini

No gráfico da figura 12, vemos que o modelo Gemini mostrou uma variação muito maior nos tempos de requisição. O *Prompt 1* teve o maior tempo médio, em torno de 35 segundos, enquanto o *Prompt 2* apresentou uma melhora significativa, com uma média de cerca de 23 segundos. O *Prompt 3*, assim como no GPT-4o, foi o mais rápido, com um tempo de resposta muito inferior aos outros, cerca de 3 segundos. Esses dados revelam que o Gemini tem uma variação mais acentuada entre os *prompts*, sendo o *Prompt 3* o mais eficiente e o *Prompt 1* o mais lento, mostrando que a escolha do *prompt* é crucial para otimizar o desempenho.

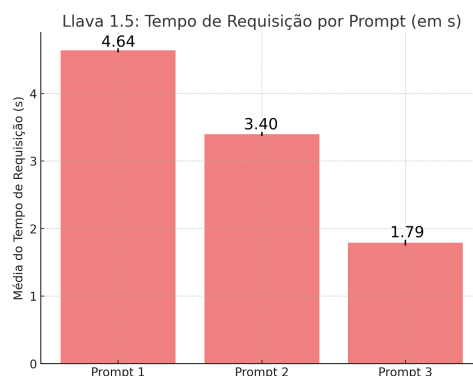


Figura 13: Tempo médio dos *prompts* do Llava 1.5

Analisando o modelo Llava 1.5 (Figura 13), ele também apresentou diferenças notáveis entre os tempos de resposta dos *prompts*. O *Prompt 1* teve o maior tempo médio, em

torno de 4,6 segundos, enquanto o *Prompt 2* mostrou uma redução para aproximadamente 3,4 segundos. Novamente, o *Prompt 3* foi o mais rápido, com um tempo médio de 1,8 segundos. Esses resultados indicam que, embora o Llava 1.5 tenha uma performance inferior comparada ao GPT-4o, o *Prompt 3* é consistentemente a melhor escolha para tempos de resposta mais rápidos.

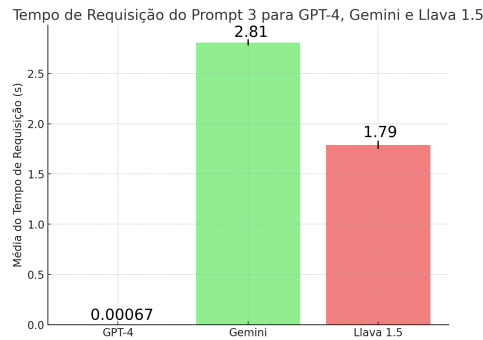


Figura 14: Tempo médio dos *prompts 3* de cada modelo

Comparando o melhor *prompt* de cada modelo (Figura 14), podemos observar uma clara diferença de desempenho. O GPT-4o se destacou com o tempo de requisição mais rápido, com uma média de 0,00067 segundos, tornando-o a opção mais eficiente em termos de processamento. Em contrapartida, o Gemini apresentou o maior tempo de resposta, com uma média de 2,81 segundos, mostrando que, apesar de sua precisão, sua velocidade é consideravelmente inferior, o que pode limitar seu uso em aplicações que exigem agilidade. O Llava 1.5 apresentou um desempenho intermediário, com uma média de 1,79 segundos, sendo mais rápido que o Gemini, mas ainda atrás do GPT-4o. Esses resultados indicam que, para aplicações que exigem respostas rápidas e processamento eficiente, o GPT-4o é a escolha mais indicada, enquanto o Llava 1.5 é uma alternativa viável e o Gemini é o menos adequado em termos de tempo de resposta.

5.5.6 Precisão na Extração dos Dados

1. GPT-4o: Alcançou alta precisão na extração dos valores de glicose, data e hora. Sua habilidade em compreender o layout típico dos glicosímetros e interpretar corretamente os dígitos e informações exibidas foi fundamental para sua eficácia.
2. Gemini: Também demonstrou alta precisão na extração dos valores de glicose, assim como na obtenção correta das informações de data e hora. Sua capacidade de interpretar os dados exibidos no glicosímetro de maneira precisa reforça sua eficácia para o monitoramento contínuo, oferecendo uma solução confiável para a análise automatizada dos níveis de glicose.

3. Llava 1.5: A precisão foi consideravelmente inferior em comparação com os outros modelos. O Llava 1.5 frequentemente falhava em extrair corretamente os valores de glicose e tinha dificuldade significativa em identificar a data e a hora, limitando sua utilidade para aplicações que demandam alta confiabilidade.

5.5.7 Impacto da Qualidade das Imagens

É importante destacar que a qualidade das imagens dos glicosímetros pode variar significativamente, dependendo do dispositivo utilizado para a captura (como diferentes modelos de smartphones), das habilidades do usuário e das condições ambientais. Um modelo de IA utilizado nesta aplicação deve ser capaz de lidar com essas variações para garantir a confiabilidade dos dados. O GPT-4o demonstrou maior resiliência a essas variações, mantendo a precisão e a velocidade de resposta. O Gemini mostrou-se mais sensível, com tempos de resposta mais longos em condições adversas, enquanto o Llava 1.5 teve uma eficiência insatisfatória nessas situações.

5.5.8 Impacto da Conexão de Internet na Performance dos Modelos

Ao longo dos testes, foi avaliado como uma conexão de internet instável influenciou o desempenho dos modelos GPT-4o, Gemini, e Llava 1.5 durante os testes. Embora todos os modelos tenham apresentado variações nos tempos de resposta devido à oscilação da conexão, essas flutuações se mantiveram, em sua maioria, próximas à média calculada em cada caso. No entanto, o modelo Gemini foi o que apresentou a maior discrepância em termos de variação de tempo.

Nos testes com o Gemini, especialmente com os *Prompts* 1 e 2, os tempos de resposta oscilaram de maneira mais acentuada. O *Prompt* 1 variou entre 8 segundos e 50 segundos em diferentes tentativas, demonstrando uma vulnerabilidade maior ao impacto da instabilidade da rede. Isso sugere que o Gemini é mais suscetível a alterações na qualidade da conexão em comparação com os outros modelos testados.

Em contraste, tanto o GPT-4o quanto o Llava 1.5 mantiveram suas variações de tempo mais próximas das médias registradas, mesmo sob condições de conexão adversas. Embora tenham ocorrido algumas flutuações, essas oscilações foram menos pronunciadas e não afetaram significativamente o desempenho geral dos modelos.

6 Resultados

Os resultados obtidos a partir da implementação e testes dos modelos de Inteligência Artificial selecionados para a extração de dados de glicosímetros. Foram analisados a eficiência dos modelos GPT-4o, Gemini e LLaVA 1.5, utilizando diferentes *prompts* para

avaliar a precisão e a capacidade de cada um em extrair informações essenciais, como o valor da glicose, a data e a hora das medições. Os resultados demonstram as diferenças de capacidade de extração entre os modelos, destacando as situações em que cada um obteve maior ou menor sucesso na tarefa proposta.

6.1 *Prompts*

No caso do GPT e do Gemini, o *prompt* "Valor glicose com data e hora" mostrou-se o mais eficaz, retornando todas as informações de maneira correta, precisa e com uso de menos tokens. Essa técnica foi essencial para ajustar os modelos às especificidades das imagens dos glicosímetros, garantindo a extração de dados relevante.

Por outro lado, o LLaVA 1.5 mostrou-se menos eficiente. O melhor resultado foi obtido com o *prompt* em inglês "What is the glucose value with the date and time?", mas mesmo assim, o modelo não conseguiu extrair corretamente a data e a hora, e em alguns casos, o valor da glicose também foi identificado de forma incorreta.

6.2 IAs

Os critérios avaliados foram a precisão na leitura dos valores de glicose e a capacidade de extrair corretamente todos os parâmetros esperados, como glicose, data e hora.

Entre os modelos avaliados, o GPT-4o e o Gemini apresentaram a melhor eficiência em ambos os critérios.

A precisão do GPT-4o foi equivalente à do Gemini, demonstrando uma capacidade igualmente consistente de interpretar e extrair os dados com exatidão. Ao capturar as leituras do glicosímetro, o GPT-4o conseguiu identificar os valores de glicose com alta precisão, minimizando possíveis erros que poderiam comprometer o controle da diabetes. Essa confiabilidade é essencial para garantir que os usuários tenham dados corretos e possam tomar decisões seguras sobre sua saúde.

Além disso, o GPT-4o se destacou por sua capacidade de retornar os dados de forma mais simplificada em formato JSON, facilitando o processamento e a manipulação das informações em um aplicativo mobile. Essa simplicidade no tratamento dos dados, aliada à sua maior resiliência a oscilações de conexão de internet, contribuiu para uma integração mais eficiente com o sistema. Enquanto o Gemini apresentou variações significativas em tempos de resposta com uma conexão instável, o GPT-4o manteve sua consistência, tornando o desenvolvimento mais ágil e menos propenso a erros.

O modelo Gemini, apresentou um funcionamento bastante sólido. Por fim, o LLaVA 1.5 ocupou a terceira posição, mostrando uma performance que, apesar de funcional, foi menos eficiente em comparação com os outros dois modelos.

7 Conclusão

Neste trabalho, foi realizada a análise detalhada e a escolha do melhor *prompt* para cada modelo de IA (GPT-4o, Gemini e Llava 1.5), com o objetivo de extrair de forma precisa os dados dos glicosímetros, como os valores de glicose, data e hora. Para cada modelo, foram testadas diversas abordagens de *prompts*, levando em consideração não apenas a precisão dos dados retornados, mas também a eficiência em termos de tempo de resposta e resiliência a variações de conexão. Isso garantiu que o processo de extração fosse tanto preciso quanto econômico, facilitando a aplicação prática dessas IAs em ambientes de monitoramento contínuo de glicemia.

O GPT-4o se destacou em relação ao Gemini por sua maior resiliência a oscilações de conexão de internet, mantendo tempos de resposta consistentes. Além disso, o GPT-4o ofereceu dados retornados em formato JSON de forma simplificada, facilitando a manipulação e processamento de informações.

Este trabalho demonstra o potencial do uso de modelos avançados de IA no controle de condições crônicas, como a diabetes. A solução desenvolvida provou ser moderna e eficiente, podendo ser expandida com funcionalidades futuras, como alertas automáticos para situações de risco (hipoglicemia/hiperglicemia) e sugestões personalizadas com base no histórico de glicose do usuário, tornando a gestão da saúde ainda mais proativa e personalizada.

8 Trabalhos futuros

Com base nos resultados alcançados neste projeto, há diversas possibilidades de expansão e aprimoramento que podem ser exploradas em trabalhos futuros. Um dos principais focos de desenvolvimento seria a melhoria da robustez dos modelos de IA em condições de conexão de internet instável, especialmente no caso do Gemini, que apresentou maior discrepância nos tempos de resposta. Soluções como cache local de dados ou o uso de técnicas de otimização de rede poderiam ser exploradas para garantir uma maior consistência no desempenho do modelo, independentemente da qualidade da conexão.

Outra direção interessante seria a ampliação do escopo dos dados extraídos dos glicosímetros. Além de coletar informações básicas como glicose, data e hora, futuros estudos podem investigar a aplicação de IA Generativa para identificar padrões glicêmicos ao longo do tempo e fornecer previsões personalizadas sobre variações de glicose. Isso poderia incluir alertas automáticos para situações de risco, como hipoglicemia ou hiperglicemia, melhorando ainda mais o acompanhamento do paciente.

A integração com outras variáveis de saúde, como níveis de atividade física ou consumo alimentar, também pode ser considerada. Por meio do uso de aprendizado de máquina,

seria possível recomendar ajustes na dieta ou na dosagem de insulina com base no histórico individual do usuário, fornecendo uma gestão ainda mais personalizada da diabetes.

Por fim, a implementação de tecnologias de segurança e privacidade de dados será crucial para garantir que as informações sensíveis dos usuários sejam protegidas, especialmente em aplicações de saúde que dependem da coleta e armazenamento de grandes quantidades de dados pessoais.

Referências

- [1] International Diabetes Federation. *IDF Diabetes Atlas*. International Diabetes Federation, 10th edition, 2021.
- [2] Paraskevi Salpea Belma Malanda Suvi Karuranga Nigel Unwin Stephen Colagiuri Leonor Guariguata Ayesha A. Motala Katherine Ogurtsova Jonathan E. Shaw Dominic Bright Rhys Williams On behalf of the IDF Diabetes Atlas Committee 1 Pouya Saedi, Inga Petersohn. Global and regional diabetes prevalence estimates for 2019 and projections for 2030 and 2045: Results from the international diabetes federation diabetes atlas, 9th edition. *Diabetes Research and Clinical Practice*, 2019.
- [3] J.W. Sohn and W.K. Ho. Cellular and systemic mechanisms for glucose sensing and homeostasis. *Pflugers Archiv - European Journal of Physiology*, 472:1547–1561, 2020.
- [4] William H. Polonsky, Lawrence Fisher, Charles H. Schikman, Deborah A. Hinnen, Christopher G. Parkin, Zhihong Jelsovsky, Bettina Petersen, Matthias Schweitzer, and Robin S. Wagner. Structured self-monitoring of blood glucose significantly reduces a1c levels in poorly controlled, noninsulin-treated type 2 diabetes: Results from the structured testing program study. *Diabetes Care*, 34(2):262–267, 2011.
- [5] Marcia Nery. Hipoglicemia como fator complicador no tratamento do diabetes melito tipo 1. *Arquivos Brasileiros de Endocrinologia Metabologia*, 2008.
- [6] Guillermo E. Umpierrez Ketan Dhatariya, Leonor Corsino. Management of diabetes and hyperglycemia in hospitalized patients. *MDText.com, Inc.*, 2000.
- [7] C. Di Mario, S. Genovese, G.A. Lanza, et al. Role of continuous glucose monitoring in diabetic patients at high cardiovascular risk: an expert-based multidisciplinary delphi consensus. *Cardiovascular Diabetology*, 21:164, 2022.
- [8] National Institute of Diabetes and Digestive and Kidney Diseases. Blood glucose control studies for type 1 diabetes: Dcct edic, 2023. Accessed: 2024-09-09.

- [9] Jay S. Skyler. Effects of glycemic control on diabetes complications and on the prevention of diabetes. *Clinical Diabetes*, 22(4):162–166, 2004.
- [10] Ruth S. Weinstock, Grazia Aleppo, Timothy S. Bailey, Richard M. Bergenstal, William A. Fisher, Deborah A. Greenwood, and Laura A. Young. The role of blood glucose monitoring in diabetes management. *ADA Clinical Compendia*, 2020(3):No Pagination Specified, 2020.
- [11] Danielle Bruen, Colm Delaney, Larisa Florea, and Dermot Diamond. Glucose sensing for diabetes monitoring: Recent developments. *Sensors*, 17(8), 2017.
- [12] L. Giray. Prompt engineering with chatgpt: A guide for academic writers. *Annals of Biomedical Engineering*, 51(11):2629–2633, 2023.
- [13] Kuldeep Singh Kaswan, Jagjit Singh Dhatteval, Kiran Malik, and Anupam Baliyan. Generative ai: A review on models and applications. *2023 International Conference on Communication, Security and Artificial Intelligence (ICCSAI)*, 2023.
- [14] H. Alqahtani, M. Kavakli-Thorne, and G. Kumar. Applications of generative adversarial networks (gans): An updated review. *Archives of Computational Methods in Engineering*, 28:525–552, 2021.
- [15] A. Romagnoli, F. Ferrara, R. Langella, et al. Healthcare systems and artificial intelligence: Focus on challenges and the international regulatory framework. *Pharmaceutical Research*, 41:721–730, 2024.
- [16] W. Cain. Prompting change: Exploring prompt engineering in large language model ai and its potential to transform education. *TechTrends*, 68:47–57, 2024.
- [17] Himath Ratnayake and Can Wang. A prompting framework to enhance language model output. In Tongliang Liu, Geoffrey Webb, Longbing Yue, and Dacheng Wang, editors, *AI 2023: Advances in Artificial Intelligence*, volume 14472 of *Lecture Notes in Computer Science*. Springer, Singapore, 2024.
- [18] Accu-Chek. Monitor de glicemia accu-chek guide. <https://www.accu-chek.com.br/monitores-de-glicemia/guide>. Accessed: September 14, 2024.
- [19] Homaira Huda Shomee and Ataher Sams. License plate detection and recognition system for all types of bangladeshi vehicles using multi-step deep learning model. In *2021 Digital Image Computing: Techniques and Applications (DICTA)*, pages 01–07, 2021.

- [20] Margaret Mitchell, Jesse Dodge, Amit Goyal, Kota Yamaguchi, Karl Stratos, Xufeng Han, Alyssa Mensch, Alexander Berg, Tamara Berg, and Hal Daumé III. Midge: Generating image descriptions from computer vision detections. In *Proceedings of the 13th Conference of the European Chapter of the Association for Computational Linguistics*, pages 747–756, 2012.
- [21] Narek Tumanyan, Michal Geyer, Shai Bagon, and Tali Dekel. Plug-and-play diffusion features for text-driven image-to-image translation. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 1921–1930, 2023.